Journal of Eye Movement Research
7(5):4, 1-15

Attention distribution and cognitive load in a subtitled academic lecture: L1 vs. L2

Jan-Louis Kruger
Macquarie University, Sydney, Australia
North-West University, Vanderbijlpark, South Africa

Esté Hefer
North-West University,
Vanderbijlpark, South Africa

Gordon Matthew
North-West University, Vanderbijlpark,
South Africa

In multilingual classrooms, subtitling can be used to address the language needs of students from different linguistic backgrounds. The way students distribute their visual and cognitive resources during a lecture is important in educational design. Students have to shift their attention between sources of information of varying density and relevance. If there is redundancy between these sources, there will be competition and possible cognitive overload. This paper compares visual attention distribution between subtitles and other sources of information through eye tracking and relates this to academic comprehension and cognitive load as measured through self-report questionnaires and EEG. The study provides promising results for the use of both first and second language subtitles in academic contexts.

Keywords: Subtitling; eye tracking; visual attention distribution; cognitive load; Sesotho first language (L1); English second language (L2)

Introduction

The way students distribute their visual and cognitive attentional resources during an academic lecture is of paramount importance in educational design. When attending to an academic lecture, students constantly have to shift their attention between different sources of information of varying information density and relevance. These sources include verbal communication (spoken and written words) as well as non-verbal communication (e.g. intonation, gestures and facial expressions) of the lecturer and other students, visual aids (e.g. text or graphics on a blackboard, whiteboard or slides), and any other materials that convey important and subject-related information. If there is redundancy between the words spoken by a lecturer, visual information on a slide, and a transcription or translation of the words of the lecturer in subtitles (in the case of a recorded lecture), there will necessarily be competition, and a risk of cognitive overload.

In this study, different eye tracking measures, an EEG (electroencephalograph), and a self-reported task load questionnaire are used to monitor students' eye movements and levels of engagement while watching a recording of an academic lecture. The main focus of this paper is the comparison of visual attention distribution (derived from eye tracking data) between subtitles, slides and the lecturer (as information-rich sources) and the rest of the screen (an information-poor source), as recorded for participants reading subtitles in Sesotho as their first language (L1) and English as their second language (L2).

In particular, the study grapples with the complex relationship between the benefits of dual coding and the limits of the human cognitive system as expressed in cognitive load theory, particularly in the context of educational material in the viewer's second language. Although the study was conducted specifically in an educational context, the findings may also be relevant to other fields in media- and film studies as well as audiovisual translation.

The central aims of the study are to determine the impact of attention distribution and subtitle language on comprehension in an academic context, and to determine the impact of subtitles on cognitive load.

DOI 10.16910/jemr.7.5.4
This article is licensed under a
Creative Commons Attribution 4.0 International license.
ISSN 1995-8692

Language and learning

In South African higher education, multilingual classrooms (classrooms that contain students from different linguistic backgrounds) are the norm. However, despite their language differences, these students share a common ground, namely that they mainly attend academic lectures in English. For most of these students, English is not their first language although it will have been their language of teaching and learning (LoTL) for the major part of their education. Students with an African language as home language in South Africa as well as other regions in Africa typically receive most, if not all, of their primary and secondary education in English. Pretorius and Mampuru emphasise (2007:38) this:

> The African continent is characterised by linguistic diversity but due to its colonial past, the majority of learners in Sub-Saharan Africa do not do their schooling in their home language but through the medium of a former colonial language. If schooling does occur in the home language, it does so for a few years only, before switching to the former colonial language.

This means that English often becomes these students' stronger language (i.e. stronger than their home language) in an academic context (Matjila & Pretorius, 2004; Pretorius & Mampuru, 2007; Hefer, 2011). Nevertheless, there has been some evidence that, in this context, students still comprehend certain materials better when read in their L1 (when their L1 is a language other than English) than when read in English as their L2 (Mahlasela, 2013; Hefer, 2011).

The current study is therefore also situated within this context of language and learning – do students benefit from L1 as opposed to L2 English subtitles, and is there a difference in cognitive load when reading subtitles in the different languages?

Visual attention distribution

Visual attention distribution refers to where viewers focus their attention when presented with different sources of information. The distribution of visual attention between text and graphic elements form a semiotic relationship, and when these elements are mutually beneficial it can lead to the forming of a conceptual idea (Carney & Levin, 2002). Research has shown that poor integration of text and graphic elements can cause hindrance, which can lead to the text or the graphic being considered a "distraction" which diverts the focus from areas which supply important information. It could also result in cognitive overload (Carney & Levin, 2002).

Furthermore, when there is redundancy between two or more sources of information, the competition between the sources may also impact negatively on comprehension due to potential cognitive overload (see, e.g., Diao, Chandler & Sweller, 2007; Mayer, 2002; Mayer, Heiser & Lohn, 2001). The current study elaborates on this issue by investigating the use and impact of subtitles in a video recording of an academic lecture where there is redundancy not only between spoken and written text (lecturer's speech and transcript thereof in the subtitles), but also between spoken text and graphics on slides.

While comprehension tests have been the primary method of investigating academic performance with regards to the integration of text and graphics and other educational design factors in the past (Hegarty, Carpenter & Just, 1991; Pellegrino, Chudowsky & Glaser, 2001), the use of eye tracking technology is proving to be an increasingly valuable tool in this regard, as it offers a precise indication of visual attention distribution during the viewing of stimuli (Paas, Tuovinen, Tabbers & Van Gerven, 2003; Kruger, 2013). Although eye tracking remains an indirect measurement of cognitive load (see Brünken, Plass & Leutner, 2003), it has the potential to provide insight into the cognitive processing of multimodal texts, particularly when used in combination with other measures.

Subtitle reading and eye tracking

Eye tracking has long been used to study reading (see Rayner, 1998 for an overview), and is especially valuable when studying subtitle reading because it offers detailed information about the viewing process (De Linde & Kay, 1999:37), with participants having to look at the on-screen visuals and read the information presented in the subtitles. There is therefore a great demand for visual attention, as information has to be gathered from different sources of information, sources that are often in competition.

In the last decade, a few eye tracking studies appeared that investigate subtitle processing. D'Ydewalle and De Bruycker (2007) investigated the reading of standard interlingual subtitles as well as reversed subtitles (subtitling into the foreign language) and found more regular reading patterns when participants view standard interlingual subtitles. Importantly, they establish that subtitles

DOI 10.16910/jemr.7.5.4
This article is licensed under a
Creative Commons Attribution 4.0 International license.
ISSN 1995-8692

are read automatically by viewers, also in the reversed subtitling condition where the participants do not understand the language in the subtitles.

Perego et al. (2010) did a study to determine the effect of subtitles on memory and comprehension while participants watch a subtitled film. Participants' cognitive performance was measured by a general comprehension test on the content of the film as well as using face-name associations of characters in the film. The results show that participants rely on subtitles to understand the content of the film with all of the participants having no problem reading the subtitles (Perego et al., 2010). The authors conclude that the cognitive processing of subtitles is effective, and it does not impact negatively on the processing of the visuals.

Winke et al. (2013) did an eye tracking study to determine the total amount of time foreign language students spend on subtitles in Chinese, Arabic, Russian and Spanish. All the participants were of intermediate proficiency in all of these languages. The students were given two English documentaries, one about salmon and the other about bears. These videos included foreign language subtitles in the languages given. Each student watched both videos with one being subtitled in a foreign language and the other not being subtitled at all. The results showed that on average 68% of the total time was spend on the subtitles, with results ranging from 63% (Spanish) to 75% (Arabic). The results from this study showed that L1-English speaking students of Spanish and Russian behave differently from L1-Arabic and Chinese students when reading subtitles. It seemed that the Chinese-language students were more accustomed to employing a strategy of reading subtitles while paying less attention to images on the screen when the verbal information was difficult to process (Winke et al., 2013).

Other studies investigating subtitle processing by means of eye tracking include that by Szarkowska et al. (2011) which presents a comparison of subtitle processing by Deaf, hard of hearing and hearing audiences; Ghia (2012) who studied the impact of translation strategies on subtitle reading; Bisson et al. (2012) who looked at the processing of native language and foreign language subtitles; Rajendran et al. (2013) who investigated the impact of text chunking on subtitle reading; and Krejtz et al. (2013) whose study sheds light on the processing of subtitles when the subtitles stay on screen during shot changes.

Cognitive load theory

Cognitive load (CL) is a theoretical construct describing the internal processing of tasks that cannot be observed directly (Mayer, 2002). According to Diao et al. (2007: 237), cognitive load theory (CLT) can be defined as being "concerned with relationships between working and long-term memory and the effects of those relationships on learning and problem solving". Within CLT, the redundancy effect occurs when learners have to mentally coordinate the same information presented simultaneously in different forms (Diao et al., 2007: 239). The mere fact that the information is present in more than one form, as in a subtitled academic lecture, means that the viewer not only has to manage attention distribution, but also has to assign some cognitive capacity to the verification of the information between the different sources. This could result in cognitive overload.

According to literature, CL can be subcategorised into intrinsic, extraneous and germane CL (Mayer, 2002). Intrinsic CL is an inherent quality of the material presented to a participant based on the difficulty thereof. This type of CL cannot be manipulated in an experiment. Extraneous CL is created by the way the information is presented (e.g. a video with or without subtitles) and can therefore be manipulated and is related to the design of the instructional material. Germane CL constitutes the remaining available cognitive resources, or the CL that people use to process and comprehend material and to form and automate schemata.

The higher the intrinsic and extraneous load, the less capacity remains in working memory for germane CL, which can result in cognitive overload. In educational design subtitles are assumed to increase extraneous CL (Brünken et al. 2003; Paas & Van Merriënboer, 1993; Paas, Van Merriënboer & Adam, 1994). It is suggested that since subtitles increase extraneous CL, it results in a reduction in germane CL that is responsible for the formation of schemata, and is therefore detrimental to learning.

In other fields like language acquisition, however, subtitles are regarded to decrease extraneous cognitive load because of the visual support it provides, thereby increasing germane CL and impacting positively on performance and learning (Paas et al., 2003). This is in line with dual coding theory which holds that combining images with verbal information improves information

DOI 10.16910/jemr.7.5.4
This article is licensed under a
Creative Commons Attribution 4.0 International license.
ISSN 1995-8692

processing (Sydorenko, 2010; Paivio, 1986, 1991, 2007), as well as with the information delivery hypothesis which holds that the delivery of the same information by more than one path results in improved learning (see Mayer et al., 2001:190). Due to this complexity it is essential to measure the impact of subtitles on cognitive load under different conditions and in different contexts in order to determine the usefulness of subtitles in multimodal educational design.

Methodology

Participants.

A convenience sampling method was employed to select Sesotho L1-speaking students from the Vaal Triangle Campus of the North-West University, who study through medium English as a second language (L2). Sesotho as L1 was set as criterion for this study because it is one of the official languages of South Africa, and because it is spoken by the majority of students on the campus where the study took place. It therefore provides a common demographic. Sesotho is spoken by 7.6% of the country's population, or 3.8 million people (www.southafrica.info, 2014).

A total of 72 participants were initially tested, but after excluding invalid data sets, 68 participants remained. Data sets were excluded based on whether participants' eye movements had been sufficiently tracked. For this, an eye tracking ratio of 80% was used as cut-off point.

Materials.

The materials used in this study include a biographical questionnaire, a video recording of a first-year Psychology lecture, a comprehension test and a self-report questionnaire on task load and engagement.

The biographical questionnaire was used to collect basic information on participants and to control for confounding variables such as age, field of study, and existing subject knowledge of Psychology.

The primary stimulus shown to participants was a 14 minute segment of a video recording of a first-year Psychology lecture. The lecture was presented in English, and was presented to participants without subtitles, with English subtitles or with Sesotho subtitles, depending on the group to which they had been assigned randomly. The first group (n = 22) watched the recorded lecture without subtitles (Group E); the second group (n = 26) watched the recorded lecture with L2 English subtitles (Group EE); and the third group (n = 20) watched the recorded lecture with L1 Sesotho subtitles (Group ES). The subtitles were produced using Screen's Poliscript™ subtitling software. A maximum of two lines were used, with a maximum of 37 characters per line. Subtitle presentation rate was set at 120 words per minute (wpm). In practice this means that the subtitles present a near-verbatim transcription of the lecturer's words, synchronised with the spoken words according to established subtitle parameters. The recorded lecture also contained presentation slides, which were presented in English to all groups to replicate the original classroom design.

The comprehension test was issued to assess participants' understanding of the content of the lecture, as well as being an objective indirect measure of CL. The test consisted of 20 multiple choice items with an item reliability index of .9, and participants could answer all questions in their own time. The questions consisted only of elements mentioned by the lecturer in the recorded segment, such as definitions and examples. The same test was administered twice: Test 1 was administered directly after participants watched the stimulus in order to measure short-term memory; Test 2 was administered approximately two weeks later in order to measure longer-term memory.

CL measurements.

With regard to measurement, CL can be conceptualised in three dimensions, namely the mental load, mental effort and performance of a participant (Diao, Chandler & Sweller, 2007). Mental load is imposed by the difficulty of the environment in which the task is being completed. Mental effort, in turn, can be defined as the total amount of controlled cognitive processing in which a subject is engaged, while measures of mental effort can provide information on the cognitive costs of learning, performance or both (Kalyuga, 2012). The level of performance can be established by a post-task test where the number of correct answers serves as an indication of performance. The combination of performance and mental effort is then considered to be the best indicator of CL (Diao, Chandler & Sweller, 2007).

According to Brünken et al. (2003) instruments or methods used to determine cognitive load can be classified in terms of causal relations and objectivity. Causal

DOI 10.16910/jemr.7.5.4
This article is licensed under a
Creative Commons Attribution 4.0 International license.
ISSN 1995-8692

relations can be divided into two sub-categories, namely indirect and direct measurements of CL. According to Brünken et al. (2003:55), "The causal relation dimension classifies methods based on the type of relation of the phenomenon observed by the measure and the actual attribute of interest". Objectivity can also be divided into two sub-categories, namely objective and subjective measurements of cognitive load. The objectivity category describes whether the method uses subjective, self-reported data or objective observations of behaviour, physiological conditions, or performance.

Together, the different categories form a matrix of measurements with four categories: subjective-indirect, objective-indirect, subjective-direct and objective-direct. These four categories are used to categorise the measurements used in the current study (Table 1):

Table 1. Categorization of CL measurements used in the current study (based on Brünken et al., 2003)

Category	**Measurement**
subjective-indirect	self-reported mental effort
subjective-direct	self-reported frustration levels and self-reported comprehension effort
objective-indirect	eye tracking measures (RIDT) and comprehension test
objective-direct	EEG measures

Self-reported mental effort, frustration levels and comprehension effort. The self-report questionnaire on task load used in the current study was compiled from questionnaires generally used to determine the mental effort involved in completing specific tasks (Klimesch, Schack & Sauseng, 2005; Mampusti, Ng, Quinto, Teng, Suarez & Trogo, 2011; Nesbit & Hadwin, 2006). In the current study, the self-report questionnaire was administered in order to determine the participants' own perceptions of the effort involved in viewing the lecture. The answers to the respective questions were all presented on a scale from either 1 to 5 or 1 to 7, as outlined in Table 2.

Table 2. Self-report questionnaire on task load

Category	Question	Range
Mental demand	How mentally demanding was the lecture you just watched?	Extremely undemanding to extremely demanding (7-point scale)
Temporal demand	How hurried or rushed was the task?	Extremely unrushed to extremely rushed (7-point scale)
Frustration	Describe your level of discouragement, irritation, stress or annoyance while watching the lecture?	Extremely low to extremely high (7-point scale)
Difficulty level	How easy or difficult was the lecture to understand?	Extremely easy to extremely difficult to understand (7-point scale)
Concentration / engagement	To what extent could you concentrate on the lecture (i.e. without thinking about other things)?	Not at all to all of the time (5-point scale)

Eye tracking measurements. For the purpose of the current study the focus was limited to two eye tracking measures as calculated from basic fixations and saccades, namely %DT (the percentage dwell time in an area of interest) and RIDT (Reading Index for Dynamic Texts; see Kruger & Steyn, 2014). These measures are of particular relevance as they indicate attention distribution and the extent to which subtitles are read.

- *Percentage dwell time (%DT).*

To accurately calculate data for the different sources of information, eye tracking data is grouped according to so-called "areas of interest" (AOIs) as in Figure 1. One eye tracking measure of particular importance in this regard, is "dwell time", which refers to the total amount of time spent looking at and processing a specific object or area. Dwell time is calculated as the sum of the duration of all fixations and saccades that hit the AOI (SMI, 2009b).

DOI 10.16910/jemr.7.5.4
This article is licensed under a
Creative Commons Attribution 4.0 International license.
ISSN 1995-8692



Figure 1. Screenshot from video showing marked areas of interest (AOIs).

The percentage dwell time is then calculated as the dwell time on a specific AOI as a percentage of the total dwell time for the video. This measurement gives an indication of attention distribution in terms of the amount of time spent on the various sources of information as defined for this study.

- *Reading Index for Dynamic Texts (RIDT).*

Although it is an index based on eye movement data, RIDT can also be considered an indirect measure of CL (Table 1) as it is directly related to the task of reading, and gives an indication of the extent to which a subtitle was read and processed, both visually and cognitively. The RIDT score becomes an important indication of cognitive load particularly when viewed in combination with other measures such as performance or EEG.

Subtitles are presented as textual information, but the reading of subtitles cannot be analysed in the conventional sense as the subtitles are embedded in the video it accompanies (i.e. it becomes part of the image). Another reason why conventional reading statistics cannot be applied to subtitles, is because the text is not static, but constantly changes (appears and disappears) in short segments or sentences. Essentially the subtitles become part of the audiovisual material, meaning that eye tracking systems cannot automatically calculate the specific measurements usually associated with the analysis of reading. Based on the visual inspection of reading behaviour of participants when reading subtitles, Kruger and Steyn (2014) developed an index with the potential to provide a reliable measure of the reading and visual processing of subtitles. In very simple terms, the Reading Index for Dynamic Texts (RIDT) is derived from the following measurements: number of unique fixations, average forward saccade length, number of standard words and standard word length. The following equation contains the formula by Kruger and Steyn (2014) used to calculate RIDT for a video *v*, with participant *p* viewing subtitle *s*:

$$\text{RIDT}_{vps} = \frac{\text{number of unique fixations for } p \text{ in } s}{\text{number of standard words in } s} \text{ X } \frac{\text{average forward saccade length for } p \text{ in } s}{\text{standard word length for } v}$$

The formula generates a score that provides an indication of the degree to which a particular subtitle was read by a particular participant. The average score taken for all the subtitles for a particular participant would then give an indication of the overall degree to which that participant read the subtitles for the entire video.

A very low RIDT score indicates very little reading and a high score indicates that the subtitles were read to greater extent. Although the index does not provide a baseline, a score closer to 1 indicates full processing.

Electroencephalography (EEG). EEG is a popular neuroimaging technique that measures electrical activity produced by the brain via electrodes that are placed on the scalp (Antonenko, Paas, Grabner & Van Gog, 2010:428). These measurements vary predictably in response to changing levels of cognitive stimuli (Anderson & Bratman, 2008; Klimesch, 1999). This makes EEG an appropriate choice for assessing cognitive load in educational psychology (Antonenko et al., 2010:428).

At present, it is believed that electrical activity in the brain generates at least four distinct rhythms (O'Brien, 2008). Two of these rhythms, namely Alpha and Theta, have been reported as sensitive to the difficulty of task manipulations (Janisse, 1977; Basar, 1999; Gevins & Smith, 2003).

The measurement of the changes in the Alpha and Theta brain wave rhythms reflects what is happening when participants process information in different situations, even if the participant is unaware of the changes or is unable to verbalize them (O'Brien, 2008; Gevins & Smith, 2003) . Therefore, when a person is frustrated, their mind emits a particular pattern of brain wave rhythms that is detected by an EEG. The algorithms in the associated software interpret this pattern and give a graphical representation which indicates frustration. These algorithms are created through various classifica-

DOI 10.16910/jemr.7.5.4
This article is licensed under a
Creative Commons Attribution 4.0 International license.
ISSN 1995-8692

tion approaches (e.g. Support Vector Machines, etc.), and along with numerous features taken from raw EEG data, are used to create a model of human academic emotion which includes boredom, confusion, engagement and frustration (Klimesch, 1999).

Procedure.

An SMI iViewX™ RED eye tracking system was used to monitor and record participants' eye movements while watching the recorded lecture. The RED system is a dark pupil system using the pupil/corneal reflex method. It has a sampling rate of 50 Hz, and calculates the pupil position, pupil size and relative head movement. Minimum fixation duration was set as 80 ms, with 100 px as maximum dispersion. For EEG data an Emotiv™ Neuro-headset EEG was used to record participants' brain activity while their eye movements were being recorded. The raw data was not used for interpretation; interpretations are based on the categorised output generated by the Emotiv™ software and is presented in terms of five channels: short-term excitement, long-term excitement, frustration, engagement and meditation.

All participants were tested individually. They were seated comfortably in a sufficiently illuminated room, on a stable chair at a distance of 700 mm from the stimulus screen. As soon as participants were seated, the EEG was placed on their heads and checked for valid signal and data recording before starting with the experiment. An instruction page was displayed on the screen prior to the experiment stimuli. The eye tracking and EEG data gathered during the reading of this page was used to check for accurate recording and was set as baselines for analysing the various EEG channels.

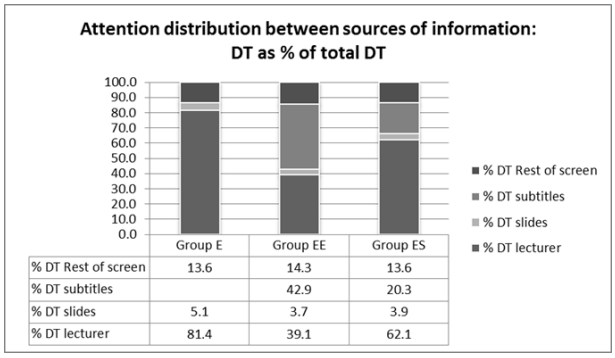
Results

The findings from the different measures are discussed in terms of two research aims, namely to determine the impact of attention distribution and subtitle language on comprehension, and to determine the extent to which subtitles affect cognitive load.

The impact of attention distribution and subtitle language on comprehension.

The addition of subtitles to any video necessarily affects the attention distribution of viewers. It is clear from Figure 2 that the attention allocated to information-rich sources (subtitles, slides and lecturer) remained fairly constant for all three conditions at just more than 85%. Slides received 5.1% of the attention in the unsubtitled condition (Group E) and 3.7% and 3.9% respectively in the presence of English subtitles (Group EE) and Sesotho subtitles (Group ES).

Interestingly, the roughly 80% of remaining visual attention is split differently in the three groups. In Group E the participants predictably look at the lecturer. In Group EE, participants divide their attention almost equally between the lecturer (39.1%) and the English subtitles (42.9%), but in Group ES, participants only allocate about one quarter of their visual attention to the Sesotho subtitles (20.3%) and the rest to the lecturer (62.1%).



	Group E	Group EE	Group ES
% DT Rest of screen	13.6	14.3	13.6
% DT subtitles		42.9	20.3
% DT slides	5.1	3.7	3.9
% DT lecturer	81.4	39.1	62.1

Figure 2. Attention distribution between sources of information: DT as % of total DT

For Group EE and Group ES, the difference in the distribution of attention to subtitles, measured as a percentage of the total dwell time in the subtitle area, reached statistical significance in a Mann-Whitney U-test ($U=106.00$, $z= -3.10$, $p<0.05$).

A statistically significant difference was also found when comparing the extent to which the participants in Group EE and Group ES read the subtitles ($U=119.00$, $z= -3.28$, $p<0.05$). Figure 3 illustrates the difference in subtitle reading as expressed by the RIDT scores.

DOI 10.16910/jemr.7.5.4
This article is licensed under a
Creative Commons Attribution 4.0 International license.
ISSN 1995-8692

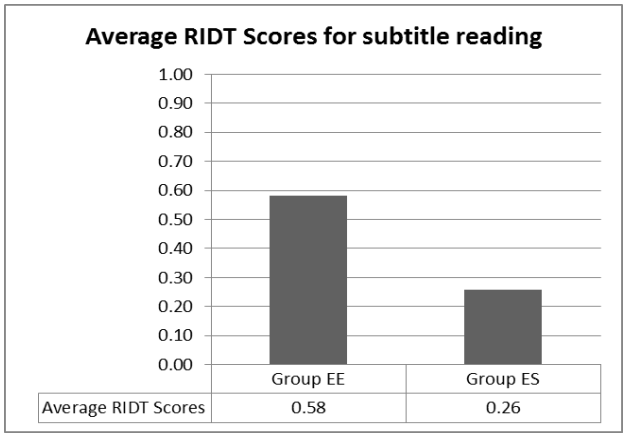


Figure 3. RIDT scores for subtitle reading

In spite of the attention allocated to the subtitles, and contrary to previous research which found that subtitles increase academic performance, the current study found no statistical significance in the difference in performance between the groups. A one-way ANOVA between the three groups revealed no significant difference in either the first comprehension test completed directly after the video or the second comprehension test completed two weeks later. There was a slight difference in comprehension scores in both tests (Figure 4). On average, Group ES (63%) and Group E (64%) scored lower than Group EE (67%) in Test 1, although this difference was not statistically significant ($t(42)= .85$, $p > .05$). Test 2 yielded different results, with Group ES outperforming the two other groups with an average of 58%, scoring better than both Group EE (51%) and Group E (54%) although neither of these differences were statistically significant.

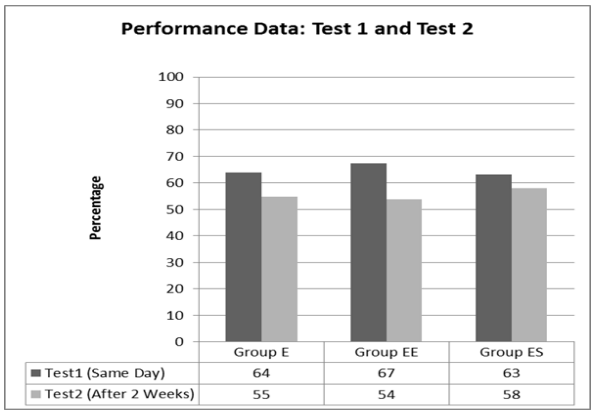


Figure 4. Performance data and information retention between Test 1 and Test 2

In terms of the difference in results between the first and the second comprehension test, it would seem that although all three groups performed worse in the test written two weeks after seeing the video, the group that saw the Sesotho subtitles did not deteriorate as much as the two other groups as can be seen in Figure 4.

T-tests by variables yielded statistically significant differences between Test 1 and Test 2 for Group E ($t(37)= 2.38$, $p < .05$) as well as for Group EE ($t(40)= 3.61$, $p < .001$) but not for Group ES ($t(32)= 1.05$, $p > .05$). This would seem to suggest that the benefits derived from the English subtitles disappeared over a period of two weeks, whereas the opposite was found for the Sesotho subtitles, where Group ES outperformed the two other groups in the second test and retained information better. This has to be read with caution, however, as the Sesotho subtitles group in fact avoided reading the subtitles to a large extent, but it could be an indication that checking information in Sesotho in the subtitles resulted in longer term benefits in information retention.

These findings may not provide evidence for any dramatic comprehension gains from the use of subtitles in an academic context, but the fact that neither English nor Sesotho subtitles resulted in a drop in performance at least dispels the fear that split attention resulting from the introduction of subtitles results in cognitive overload. The suggestion from the data that the presence of L1 subtitles may be beneficial for information retention over a longer period, even when subtitles are not read in full, also provides an interesting avenue for further research.

The effect of subtitles on cognitive load (CL).

The impact subtitles have on attention distribution, particularly the fact that there is an inevitable split in visual attention between the subtitles, lecturer and slides, may have an impact on perceptions of CL in subjective scales (self-report questionnaire on task load), as well as on CL measured by means of objective measurements like EEG and RIDT.

Self-reported questionnaire findings. The self-reported questionnaire was used as a subjective measure of CL. The average values for the different categories (mental demand, frustration, difficulty level and concentration/engagement) are given in Figure 5 below.

DOI 10.16910/jemr.7.5.4
This article is licensed under a
Creative Commons Attribution 4.0 International license.
ISSN 1995-8692

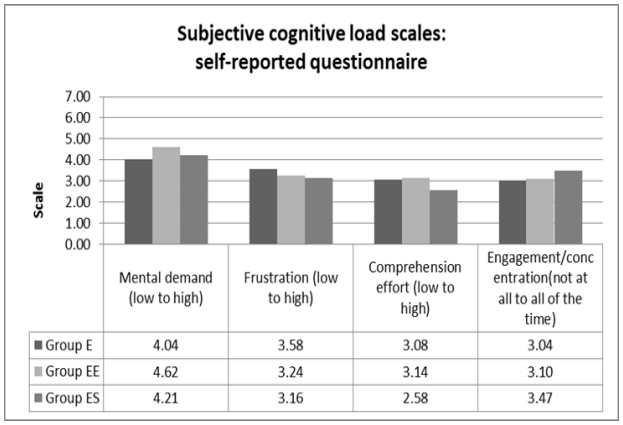


Figure 5. Average values for subjective self-reported questionnaire

Although no statistically significant effects could be identified in any of the scales by means of Kruskal Wallis ANOVAs, it is interesting that Group E experienced the highest frustration levels and the lowest mental load levels, whereas Group ES experienced the lowest comprehension effort levels and the highest engagement/concentration levels. Group EE had the highest mental load levels.

To determine whether any effect obtained between the performance of the groups and their self-reported measures, Spearman Rank Order correlations were performed. Strong positive correlations were found between both comprehension tests and self-reported concentration/engagement for the no subtitles group (Test 1: $r = .80$, $p < .05$; Test 2: $r = .52$, $p < .05$). Strong negative correlations were found for the Sesotho subtitles group between the second comprehension test and self-reported frustration levels ($r = -.68$, $p < .05$) and self-reported comprehension effort ($r = -.75$, $p < .05$). This would seem to suggest that higher perceptions of engagement can be related to higher performance in the unsubtitled condition (Group E), and that lower perceptions of comprehension effort and frustration resulted in higher performance in the presence of Sesotho subtitles (Group ES).

Paas et al. (2003) introduced a useful metric to determine the combined effect of performance and self-reported cognitive load, namely instructional efficiency. Using standardised z-scores it is possible to determine the distance from the performance-effort axis. In Figure 6 it is clear that the subtitles in both Group EE and Group ES resulted in higher instructional efficiency based on the interaction between self-reported comprehension effort and performance directly after watching the video. However, after two weeks, the efficiency for Group EE diminishes substantially while that of Group ES improves. In Figure 6, the position above the diagonal efficiency line indicates positive efficiency, while positions below the line indicate negative efficiency.

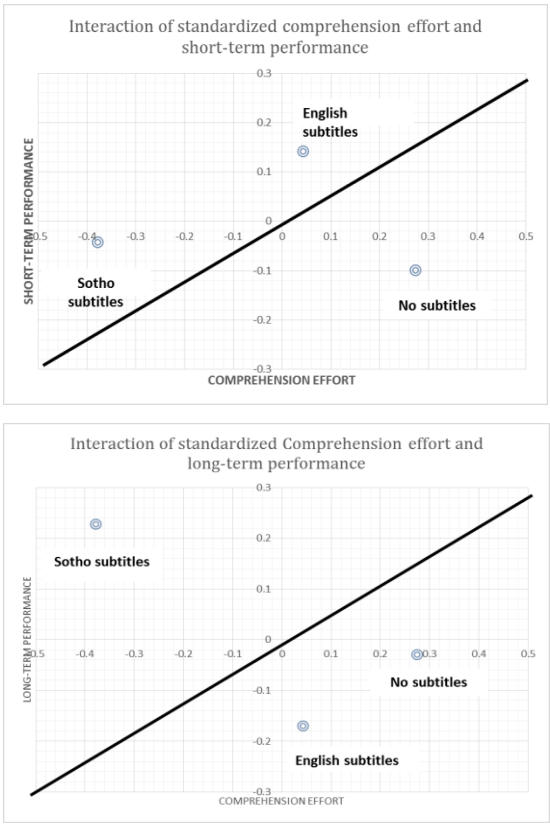


Figure 6. Interaction of standardized z-scores

This effect is perhaps more evident in Figure 7 which depicts the distance from the efficiency axis of each condition. From this it would seem that the instructional efficiency of the Sesotho subtitles is much higher than for either of the two other modes, with the short-term benefits of English subtitles diminishing to a negative position after two weeks. A Kruskal-Wallis ANOVA did not indicate statistical significance, but this preliminary trend deserves more thorough investigation with larger groups.

DOI 10.16910/jemr.7.5.4
This article is licensed under a
Creative Commons Attribution 4.0 International license.
ISSN 1995-8692

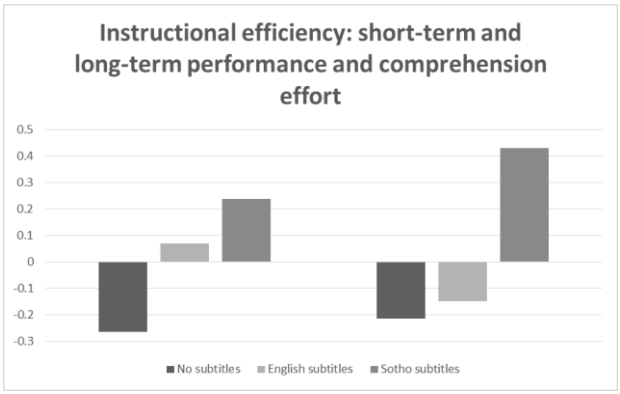
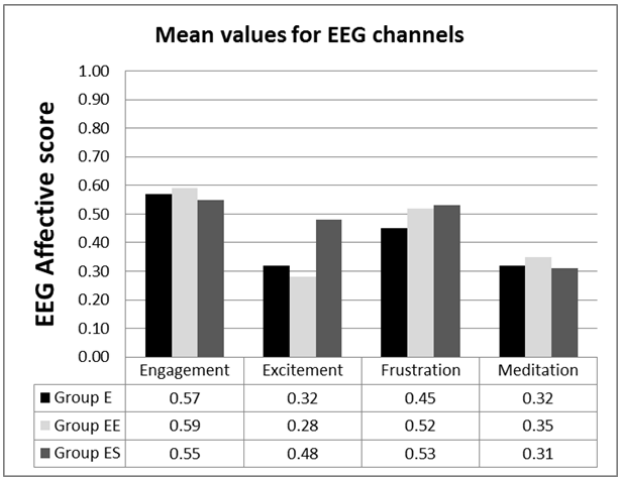


Figure 7. Instructional efficiency for performance and comprehension effort

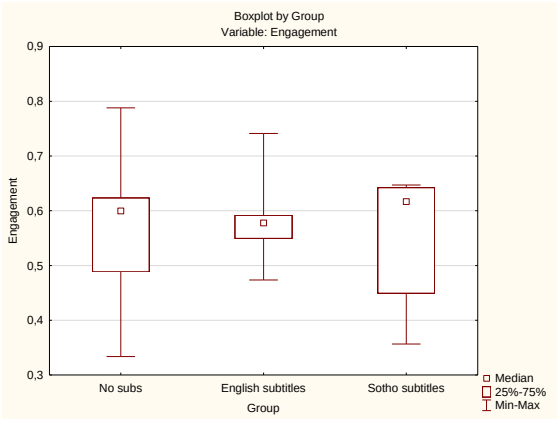
EEG findings. The electroencephalograph (EEG) was used as a direct objective measure of CL. The data indicated little difference between the mean values for engagement, frustration and meditation as measured for the three groups, with no statistical significance in terms of Kruskal-Wallace ANOVAS. The only noticeable difference (although not significant) was found for Group ES, who experienced higher levels in terms of excitement (Figure 8).

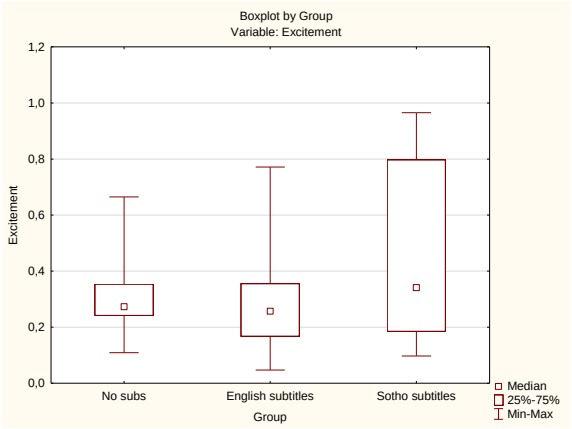


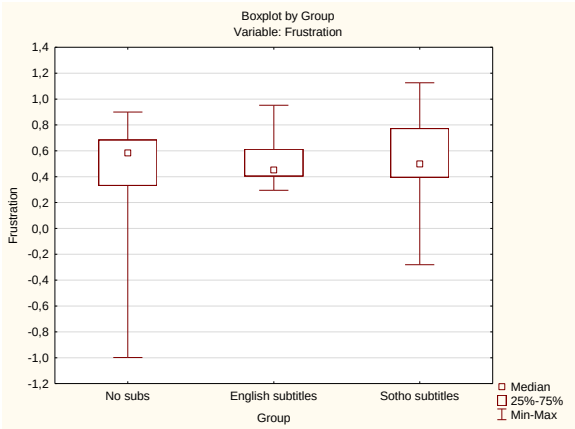
	Engagement	Excitement	Frustration	Meditation
Group E	0.57	0.32	0.45	0.32
Group EE	0.59	0.28	0.52	0.35
Group ES	0.55	0.48	0.53	0.31

Figure 8. Instructional efficiency for performance and comprehension effort

From the boxplots for the four channels (Figure 9) it is evident that the engagement, frustration and meditation levels had a much smaller interquartile range for Group EE which would suggest that this group's emotions were somehow more focused, whereas all four channels had the largest interquartile range for Group ES. This may be ascribed to the difference in reading behaviour in the various groups: participants in Group ES read fewer of the subtitles, whereas the participants in Group EE read the subtitles much more consistently.







DOI 10.16910/jemr.7.5.4

This article is licensed under a
Creative Commons Attribution 4.0 International license.

ISSN 1995-8692

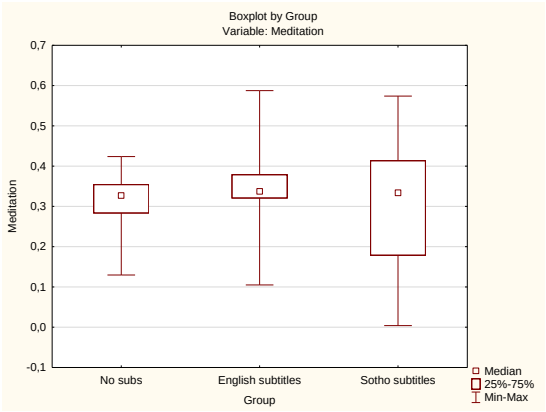


Figure 9. Boxplots representing the data from the EEG channels

As in the case of performance measures as well as subjective measures, the difference between the three groups do not reach significance, once again suggesting that subtitles do not have either a hugely positive effect, nor result in cognitive overload.

RIDT findings. RIDT, as an indirect objective measure of cognitive load directly related to the task of reading, was measured for three specific instances: reading of subtitles in the absence of any other visual textual material; reading subtitles in the presence of other visual textual material (specifically presentation slides); and reading of presentation slides in the absence of subtitles. There was a significant difference between the two subtitled groups in terms of the average RIDT on subtitles in the absence of any other visual textual material ($U=81.00$, $z = -3.70$, $p < 0.05$), with low RIDT recorded for Group ES. This is largely due to the fact that the participants in Group ES avoided the subtitled area; very little reading and therefore also very little processing of the subtitles occurred, with the majority of L1 Sesotho subtitle text being skipped. In Group EE some of the English L1 subtitle text was also skipped, but this was not nearly as significant (Figure 10).

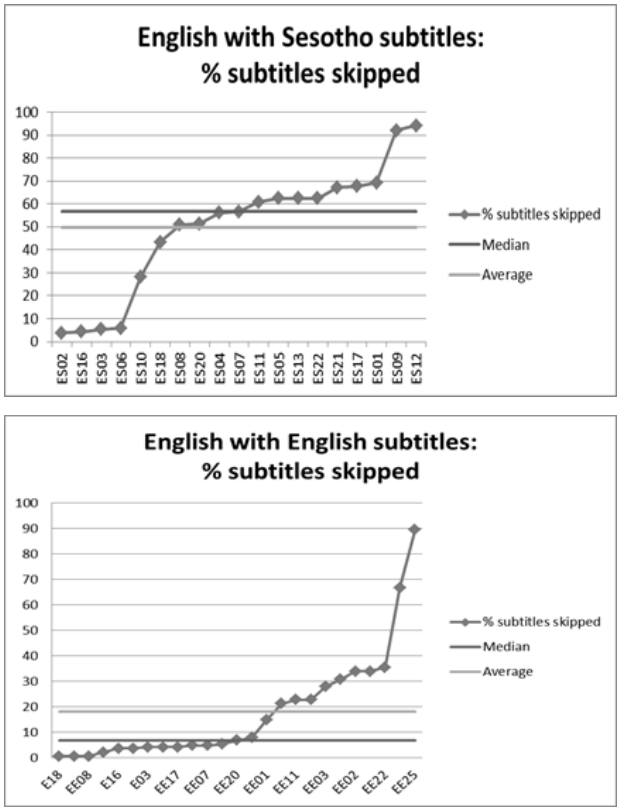


Figure 10. Percentage subtitles skipped for Groups ES and EE

No significant differences were found between the two subtitle groups with regard to the average RIDT on subtitles in the presence of slides ($U=201.00$, $z= -0.85$, $p =0.39$).

The difference in RIDT can be directly attributed to the language of subtitle presentation, and therefore also attention distribution – Groups ES, who spent less time looking at the Sesotho subtitles, spent less time reading and processing the subtitles. Although this cannot be related to cognitive load directly, it could be considered an indication that the participants thought the Sesotho subtitles would be more difficult to read and process and that they would be better off avoiding them altogether. This would also explain why the RIDT scores for Group ES were significantly higher than Group ES for those instances where English presentation slides appear on-screen – they avoid reading in Sesotho, but grasp any opportunity to read in English, the LoTL.

Nevertheless, the fact that Group ES who read the Sesotho subtitles had long-term gains as evidenced by the fact that they retained information better after a period of

DOI 10.16910/jemr.7.5.4
This article is licensed under a
Creative Commons Attribution 4.0 International license.
ISSN 1995-8692

two weeks, would seem to suggest that they made use of the subtitles at strategic times to check terms in their L1, even if they did not read the Sesotho subtitles to the same extent as those participants who read the English subtitles.

Discussion

This study set out to answer two specific questions regarding the use of subtitles in an academic context, namely what is the impact of attention distribution and subtitle language on comprehension, and to what extent subtitles affect cognitive load.

In terms of attention distribution, subtitle language and comprehension, it was found that the language or presence/absence of subtitles did not have any significant impact on performance. Although the three groups distributed their visual attention resources differently, this did not have any serious implications for the extent to which they comprehended the work discussed in the recorded lecture. This was found for Test 1 and Test 2, which can be seen as indicators of short term and long term retention of knowledge respectively. The findings for Group ES may suggest that L1 subtitles result in a higher retention of knowledge in the longer term, which could be due to the benefits of L1 cognitive priming, but this has to be confirmed in follow-up studies.

The language of the subtitles itself had an impact on attention distribution, with Sesotho subtitles being read much less than English subtitles, and even avoided in some instances. This finding contradicts existing literature on eye tracking and subtitle reading which suggests that subtitles are read "effortlessly and almost automatically" regardless of the language used (d'Ydewalle & De Bruycker, 2007:196; d'Ydewalle, Praet, Verfaille & Van Rensburgen, 1991; Van Lommel, Laenen & d'Ydewalle, 2006), and warrants further investigation. Furthermore, Group ES allocated significantly more attention to the presentation slides (written verbal information in L2 English - LOTL) than Group EE. This, along with the fact that participants in Group ES avoided reading the L1 Sesotho subtitles, might indicate a preference for English in an academic context, and an awareness of the fact that, for them, reading in Sesotho would require a higher level of cognitive processing. Due to the redundancy of the information, the additional cognitive effort that would be required to read the L1 Sesotho subtitles while listening to the L2 English audio may have caused this group to use the subtitles for a different purpose, namely to check terms rather than to follow the lecturer.

In terms of subjective measures of cognitive load and subtitle reading, the findings for Group E suggest that the absence of subtitles increase perceptions of frustration, while the presence of Sesotho subtitles resulted in a perception of lower comprehension effort in Group ES. The interaction of self-reported comprehension effort and performance measures to yield an indication of instructional efficiency suggests that the Sesotho subtitles had a higher efficiency both directly following the video and after two weeks, with the video without any subtitles having the lowest efficiency in both cases. These, however, are trends that did not yield statistical significance and that have to be investigated in more detail.

In terms of direct objective measurements, no significant differences could be found between the three groups in terms of the recorded EEG data. The only meaningful result of this data is that the English subtitles resulted in a much smaller variance in the group, with a very narrow interquartile range.

Conclusions

The findings of this study, although limited to a small and specific sample group, do provide some insight on possible ways to improve educational design, keeping in mind long term and short term performance: in terms of subtitle reading (and possibly also the reading of other study material), English is beneficial for short term performance; but given more and more in-depth exposure, L1 subtitles (and possibly also other L1 study materials) could result in better long term performance, a promising result for schema formation in instructional design.

The study also supports the findings of other studies such as D'Ydewalle and De Bruycker (2007) and Perego et al. (2010) on the cognitive effectiveness of subtitles while providing first steps towards more reliable measurement of cognitive load in the presence of subtitles. Although the measurement of cognitive load in the context of dynamic texts still requires careful experimental research, this study provides a starting point. The logical next step in this research would be to do a more comprehensive study on the influence of language history on cognitive load in the presence of subtitles.

DOI 10.16910/jemr.7.5.4
This article is licensed under a
Creative Commons Attribution 4.0 International license.
ISSN 1995-8692

The measurement of pupil dilation promises to offer a direct window into changes in cognitive load, which can only be utilised fully if the reliability of this measurement has been established for dynamic texts with changing levels of luminosity and where the shape of the pupil constantly changes as the viewer explores the different regions of the screen. Simultaneous EEG measurements could also provide valuable data in order to validate claims of changes in cognitive load.

Most importantly, before more concrete claims can be made about the effectiveness of subtitles in educational and other contexts, more research has to be conducted on the contribution of visual vs. auditory information, particularly when redundancy of information occurs.

References

Anderson, C.W., Bratman, J.A. (2008). Translating thoughts into actions by finding patterns in brainwave. In *Proceedings of the Fourteenth Yale Workshop on Adaptive and Learning Systems*, New Haven: CT, 1-6.

Antonenko, P., Paas, F., Grabner, R. & Van Gog, T. (2010). Using electroencephalography to measure cognitive load. *Educational Psychology Review, 22*, 425–438. doi: 10.1007/s10648-010-9130-y

Basar, E. (1999). *Brain function and oscillations: integrative brain functions. Neurophysiology and Cognitive Processes 2*. Berlin: Springer.

Bisson, M.-J., Van Heuven, W., Conklin, K., & Tunney, R.. (2014). Processing of native and foreign language subtitles in films: An eye tracking study. *Applied Psycholinguistics* 35(2), 399-418.

Brünken, R., Plass, J.L., and Leutner, D. (2003). Direct measurement of cognitive load in multimedia learning. *Educational Psychologist, 38*(1), 53–61. doi: 10.1207/S15326985EP3801_7

Carney, R. N. & Levin, J. R. (2002). Pictorial illustrations still improve students' learning from text. *Educational Psychology Review, 14*(1), 5-26. doi: 10.1023/A:1013176309260

De Linde, Z. & Kay, N. (1999). *The semiotics of subtitling*. Manchester: St. Jerome.

Diao, Y., Chandler, P., Sweller, J. (2007). The effect of written text on comprehension of spoken English as a foreign language. *The American Journal of Psychology, 120*(2), 237-261.

d'Ydewalle, G. & De Bruycker, W. (2007). Eye movements of children and adults while reading television subtitles. *European Psychologist, 12*(3), 196-205.

d'Ydewalle, G., Praet, C., Verfaille, K. & Van Rensburgen, J. (1991). Watching subtitled television: automatic reading behaviour. *Communication research, 18*(5), 650-666, Oct. doi: 10.1177/009365091018005005

Gevins, A. & Smith, M. E. (2003). Neurophysiological measures of cognitive workload during human computer interactions. *Theoretical Issues in Ergonomic Science, 4*, 113–131. doi: 10.1080/14639220210159717

Ghia, E. (2012). The impact of translation strategies on subtitle reading. In E. Perego (Ed.) *Eye tracking in audiovisual translation* (pp. 155–182). Rome: Aracne Editrice.

Hefer, E. (2011). *Reading second language subtitles: a case study of South African viewers reading in their native language and L2-English* (Unpublished doctoral dissertation). North-West University, Vanderbijlpark (South Africa).

Hegarty, M., Carpenter, P. A. & Just, M. A. (1991). Diagrams in the comprehension of scientific text. In R. Barr, M. L. Kamil, P. B. Mosenthal & P. D. Pearson (Eds.), *Handbook of reading research* 2, 641–668. New York: Longman.

Janisse, M.P. (1977). *Pupillometry: the psychology of the pupillary response*. Washington: Hemisphere Publishing Corporation.

Kalyuga, S. (2012). Instructional benefits of spoken words: A review of cognitive load factors. *Educational Research Review, 7*(2), 145–159, June. doi: http://dx.doi.org /10.1016/j.edurev.2011.12.002

Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: A review and analysis. *Brain Research Reviews, 29*, 169–195. doi: 10.1016/S0165-0173(98)00056-3

DOI 10.16910/jemr.7.5.4
This article is licensed under a
Creative Commons Attribution 4.0 International license.
ISSN 1995-8692

Klimesch, W., Schack, B. & Sauseng, P. (2005). The functional significance of theta and upper alpha oscillations for working memory: A review. *Experimental Psychology, 52*, 99–108. doi: 10.1027/1618-3169.52.2.99

Krejtz, I., Szarkowska, A., & Krejtz, K. (2013). The effects of shot changes on eye movements in subtitling. *Journal of Eye Movement Research*, 6 (5), 1–12.

Kruger, J.L. (2013). Subtitles in the classroom: balancing the benefits of dual coding with the cost of increased cognitive load. *Journal for Language Teaching, 47*(1), 29-53. doi: http://dx.doi.org/10.4314/jlt.v47i1.2

Kruger, J.L. & Steyn, F. (2014). Subtitles and eye tracking: reading and performance. *Reading Research Quarterly*, *49*(1), 105-120. doi: 10.1002/rrq.59

Mahlasela, J. T. (2013). *Improving comprehension in physical science content subjects through mother tongue subtitling in secondary education* (Unpublished master's thesis). North-West University, Vanderbijlpark (South Africa)

Mampusti, E.T., Ng, J.S., Quinto, J.J.I., Teng, G.L., Suarez, M.T.C., Trogo, R.S. (2011). Measuring academic affective states of students via brainwave signals. In T.B. Ho, R.I. McKay, X.H. Nguyen & T.D. Bui (Eds.), *KSE 2011: Third International Conference on Knowledge and Systems Engineering* (pp. 226–231). Hanoi, Vietnam: Conference Publishing Services.

Matjila, D.S. & Pretorius, E.J. (2004). Bilingual and biliterate? An exploratory study of grade 8 reading skills in Setswana and English. *Per Linguam*, 20(1):1-21.

Mayer, R. E. (2002). Cognitive theory and the design of multimedia instruction: an example of the two-way street between cognition and instruction. *New Directions for Teaching and Learning, 89*, 55-71. doi: 10.1002/tl.47

Mayer, R.E., Heiser, J. & Lohn, S. 2001. Cognitive constraints on multimedia learning: When presenting more material results in less understanding. *Journal of Educational Psychology*, 93(1):187-198.

Nesbit, J.C. & Hadwin, A.F. (2006). Methodological issues in educational psychology. In P.A Alexander and P.H. Winne (Eds.), *Handbook of Educational Psychology 2nd Edition* (pp. 825-848). Mahwah, NJ: Earlbaum.

O'Brien, S. (2008). Processing fuzzy matches in translation memory tools: an eye tracking analysis. In S.

Paas, F., Tuovinen, J., Tabbers, H., and Van Gerven, P.W.M. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist, 38*(1), 63-71. doi: 10.1207/S15326985EP3801_8

Paas, F.G.W.C. and Van Merriënboer, J.G. (1993). The efficiency of instructional conditions: an approach to combine mental effort and performance measures. *Human Factors, 35*(4), 737-741, December.

Paas, F.G.W.C., Van Merriënboer, J.G. and Adam, J.J. (1994). Measurement of cognitive load in instructional research. *Perceptual and Motor Skills, 79*(1), 419-430, August. doi: 10.2466/pms.1994.79.1.419

Pellegrino, J. W., Chudowsky, N. & Glaser, R. (2001). *Knowing what students know: the science and design of educational assessment.* Washington, DC: National Academy Press.

Perego, E., Del Missier, F., Porta , M. & Mosconi, M. (2010). The cognitive effectiveness of subtitle processing. *Media Psychology*, 13(3):243–272.

Pollatsek, A. & Rayner, K. (2006). Eye-movement control in reading. In M.J. Traxler & M.A. Gernsbacher (Eds.), *Handbook of Psycholinguistics, 2nd edition* (pp.613-657) San Diego: Elsevier.

Pretorius, E.J. & Mampuru, D.M. (2007). Playing football without a ball: language, reading and academic performance in a high-poverty school. *Journal of Research in Reading*, 30(1), 33-58. doi: 10.1111/j.1467-9817.2006.00333.x

Rajendran, D.J., Duchowski, A.T., Orero, P., Martínez, J. & Romero-Fresco, P. (2013). Effects of text chunking on subtitling: A quantitative and qualitative examination. Perspectives: Studies in Translatology, 21(1), 5-31.

Szarkowska, A., Krejtz, I., Kłyszejko, 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.

SMI (SensoMotoric Instruments). (2009a). iView X™ system manual: version 2.2. Teltow: SMI.

SMI (SensoMotoric Instruments). (2009b). BeGaze™ manual: version 2.3. Teltow: SMI.

DOI 10.16910/jemr.7.5.4
This article is licensed under a
Creative Commons Attribution 4.0 International license.
ISSN 1995-8692

SouthAfrica.info, (2014). The languages of South Africa, http://www.southafrica.info/about/people/language.htm#sesotho, Accessed on: 25 November 2014.

Van Lommel, S., Laenen, A. & d'Ydewalle, G. (2006). Foreign-grammar acquisition while watching subtitled television programming. *British Journal of Educational Psychology, 76,* 243–258. doi: 10.1348/000709905X38946

Winke, P., Gass, S. & Syderenko, T.. (2013). Factors influencing the use of captions by foreign language learners: an eye tracking study. *The Modern Language Journal*, 97 (1),254–275.

DOI 10.16910/jemr.7.5.4

This article is licensed under a
Creative Commons Attribution 4.0 International license.

ISSN 1995-8692