

Jan-Louis Kruger

North-West University (Vaal Triangle Campus)

Subtitles in the classroom: balancing the benefits of dual coding with the cost of increased cognitive load

Abstract

Language teachers in the twenty-first century cannot ignore the possible benefits of using multimodal texts in the classroom. One such multimodal source that has been used extensively is subtitled videos. Against the background of conflicting theories in the fields of educational psychology and psycholinguistics as well as language acquisition where multimodal texts are concerned, this article presents an experiment aimed at determining the impact of competition between different sources of information on comprehension and attention allocation. The material that is investigated is a recorded and subtitled academic lecture in Economics with PowerPoint slides edited in, as an example of multisource communication. The article in particular engages with the issue of language as it pertains to the use of English as medium of instruction for English Second Language (ESL) students in South Africa. Essentially, the article seeks to shed light on the well documented positive effects of subtitles that are explained by the information delivery hypothesis and Dual Coding Theory, and

the equally well documented negative impact explained by the redundancy effect in Cognitive Load Theory. Some evidence was found in the study that cognitive resources are assigned to more stable information sources like slides and non-verbal visual contextual information when the presentation speed of subtitles increases. This means that when the presentation speed of subtitles increases, learners focus on stable textual information (like slides) and on nonverbal information (like the face of the lecturer). Using the correct presentation speed of subtitles in multisource information in an educational setting is imperative for the activation of the potential benefits of multisource communication (that includes subtitles) for learning. The findings of the study stand to benefit all fields of multimedia educational design, but also have direct relevance to the use of technological support such as subtitles in the classroom.

Key words: cognitive load, dual coding, information delivery, multimedia educational design, subtitles

1. Introduction

In a multilingual South Africa where the language of instruction for the majority of university students is an additional language, a growing need exists to investigate how technology can be used to assist students to overcome “the language” issue (see Heugh, 2002). In this regard, a number of studies have been conducted globally over the past three decades on the use of subtitling in the context of language learning and education. There is, in fact, extensive evidence in the literature that same-language subtitles (SLS) hold significant potential in education. Some of these studies will be discussed below. The benefits of subtitling illustrated in these studies seem to support Paivio’s Dual Coding Theory (1986, 1991, 2007), which is based on the idea that “cognition consists of two separate but interconnected mental subsystems, a verbal system and a nonverbal system” (Sadoski, Paivio & Goetz, 1991: 463). Dual Coding Theory suggests that “a combination of imagery and verbal information improves information processing” (Sydorenko, 2010: 50; see also Paivio, 1986, 1991, 2007), which could have an impact on learning. More specifically, Mayer, Heiser and Lohn (2001: 190) identify the “information delivery hypothesis” as the reason why on-screen text is added to audiovisual educational material, namely the hypothesis that when the same information is delivered by more paths, students learn more. Given these positive findings, it would make sense to make the use of a well-established mode such as subtitling in language teaching. This is made particularly feasible because of the availability of subtitles on a wide range of films in all genres as well as educational videos created for specific courses. These subtitles are most commonly available in South Africa in English for Deaf and hard-of-hearing viewers (a form of subtitling also known as subtitles for the Deaf and hard of hearing or SDH) on DVDs as well as on some satellite channels and increasingly in online video files. In addition to bringing subtitled material into the classroom, learners can be encouraged to make use of the feature consciously when watching television or movies.

However, subtitles, like any other teaching aid, cannot be used indiscriminately, and there is an equally substantial body of research in the field of multimedia learning that takes the dual channel and limited capacity assumptions of cognitive theory (see Mayer, 2002) as starting point. These studies identify a redundancy effect when information is present in more than one channel, and link this to reduced learning (see, for example: Mayer, 2002; Mayer, et al., 2001; Diao, Chandler & Sweller, 2007).

As a first step in a more thorough investigation of the validity of the redundancy effect on the one hand, or the information delivery hypothesis on the other, this article will investigate the impact of subtitles on cognitive load (as evidenced by performance and behavioural measures). The article will attempt to answer the question of whether subtitles result in increased learning as measured by means of a comprehension test. Since the usefulness of subtitles is also determined by whether they are presented at a speed that makes it possible to process them, the article will also try to determine whether the presentation speed of subtitles used in an educational context have an impact on comprehension and eye movements. For this purpose, a comparison will be done of the comprehension scores of participants who were exposed to the same lecture but with subtitles at different presentation speeds.

This investigation of performance measures will be supplemented by a study of attention allocation as evident in eye movements to determine whether a difference can be observed in the attention allocation of the experimental groups when exposed to a subtitled lecture at different subtitle presentation rates, with the added competition of PowerPoint slides. In this part of the experiment, the redundancy effect will be tested when students are exposed (at different subtitle presentation rates) to the same information in three modalities or channels namely the audio, the synchronised subtitles, and the PowerPoint slides that remain on screen for longer, while also being exposed to a fourth channel, namely the non-verbal communication of the lecturer as he presents the lecture. Due to the large number of variables that are involved when so many sources of information compete for attention, this study will serve as an exploratory attempt to arrive at a better understanding of the impact of competing channels through a manipulation of presentation rate of subtitles, which could provide the basis for future refinement of the design of multisource educational materials. A better understanding of issues of cognitive load could provide direction to applied linguists and academic development specialists in terms of the use of subtitles as an aid to minimise the impact of the “language issue” on learning at South African and other universities and schools.

Before the current experiment and hypotheses it aims to test are described in more detail, a brief overview will be given of the two main fields that have investigated the effects of the addition of on-screen text on viewers’ experiences over the past decades.

2. Educational benefits of subtitling and the information processing hypothesis

2.1 Reading and listening comprehension

A number of studies have been conducted over the past three decades on the possible benefits of subtitling in terms of reading and listening comprehension. For example, Garza (1991) conducted a number of studies on subtitling in the context of second language learning. Specifically, he found that the gap between reading and listening comprehension could be bridged with the use of same language subtitles: “the addition of captions to the video material contributes significantly to the memorability of the language of a segment and, consequently, facilitates the student’s ability to use that language in the proper context” (Garza, 1991: 245).

The positive impact of subtitles in terms of word learning and word recognition, with a resulting positive effect on comprehension in a broader educational context, is also confirmed by Bird and Williams (2002). Even from this handful of studies it should be clear that subtitles could hold substantial benefits in language teaching. By introducing subtitled material such as educational videos, particularly in an ESL context, teachers could achieve integration of reading and listening skills, and also impart vast contextual networks of information that may be less effective when taught through one medium only.

Vanderplank (1988, 1990) determined that same language subtitling (SLS) increases comprehension by unlocking accents, dialects and humour, as well as drawing the learners' attention to unfamiliar phrases and words. Danan (2004) similarly found that subtitling can increase listening comprehension skills of learners who learn via an L2 since it leads to additional cognitive benefits including a greater depth of processing (see also Huang & Eskey, 1999). Again, the benefits of subtitles to the language teacher are evident from these findings.

In terms of reading and listening comprehension, as well as word recognition, Markham (1999: 326) found that "university-level ESL students clearly derive substantial listening (specifically word recognition) benefits from viewing second-language captioned video material". In particular, Markham (1999: 327) suggests that "second-language listeners with solid reading abilities can use ... second-language captions to develop their listening abilities". Whether the same gains can be achieved for learners with less than solid reading abilities (something that can be assumed with little risk of error for the majority of ESL students in the population used for this study as confirmed by their results in the reading evaluation conducted at the beginning of their studies) is debatable (see the discussion of Linebarger et al., 2010 below). This finding, is also partially in conflict with the findings of Diao et al. (2007: 237), who found that "listening with the presence of a script and subtitles led to better understanding of the scripted and subtitled passage but poorer performance on a subsequent auditory passage than listening with the auditory material only".

In terms of listening comprehension, it would therefore appear that variables such as the proficiency level of the learners in the additional language and the amount of cognitive load in the different channels play a role in the extent of the benefits.

2.2 Literacy and academic literacy

For the purpose of this article, the benefits of subtitling for academic literacy are particularly interesting and relevant. However, studies done on general literacy also provide useful insights. Linebarger, Piotrowski and Greenwood (2010) draw on two models of learning from television (dual coding theory and travelling lens theory) in discussing the use of subtitles in literacy training. Dual coding theory, which is consistent with the information delivery hypothesis discussed by Mayer et al. (2001), states that the simultaneous presentation of information via auditory and visual modalities can enhance comprehension by increasing the number of cognitive paths that can be followed to retrieve the information. According to Linebarger et al. (2010: 159), travelling lens theory, similarly, basically contends that children "develop cognitive skills to process information in one medium (e.g. television) and are able to use these skills when processing content found in other media forms (e.g. books...)". Although this refers to the transfer of skills, it would suggest that skills developed in a multimodal context such as film could be transferred to unimodal contexts.

Linebarger et al. (2010), like Markham (1999), do point out that the proficiency level of the viewers is extremely important. For example, they suggest that poor reading skills may render on-screen print very challenging, causing viewers to ignore it, while viewers who are fluent readers may also ignore such text because reading has become

habituated. They therefore suggest that the category of readers that would benefit most from subtitles in terms of literacy gains would be emerging readers who have not become fluent yet, because to them the subtitles would still be interesting and challenging “and thus within their ‘travelling lens’” (Linebarger et al., 2010: 150). The study by Linebarger et al. (2010) confirmed the dual coding theory in that both word recognition scores and comprehension scores were higher for caption viewers, even if their study focused mainly on children and not adults as is the case in the current study.

Ayonghe (2010) did an experiment with university-level students in the context of academic literacy. She investigated the use of subtitling as an aid to academic literacy at the University of Buea in Cameroon, and found statistically significant improvement in academic literacy levels in the students who were exposed to subtitled television dramas and documentaries. Although this study provides positive indications, with validity reinforced both by large test and control groups and by the duration of the experiment over a full academic semester with complete texts rather than only isolated texts or parts of texts, the exact reason for the improvement in academic literacy levels could not be isolated.

Lacroix (2012), in a study on the impact of SLS on student comprehension (done in an English as additional language, or EAL context), focused on the use of discipline-specific audiovisual material, namely recorded classes, as opposed to the use of popular television in Ayonghe’s study. Lacroix found improvement of academic performance in an academic course after exposing students to subtitled recordings of lectures over a period of six weeks.

However, these positive findings are qualified by studies in the field of educational psychology and multimedia instruction, which will be discussed next.

3. Cognitive theory of multimedia learning

3.1 Cognitive theory

According to Mayer (2002: 60), cognitive theory (in the context of multimedia design), offers “three theory-based assumptions about how people learn from words and pictures: the dual channel assumption, the limited capacity assumption, and the active processing assumption”. The dual channel assumption is based on a view of the human cognitive system as consisting “of two distinct channels for representing and manipulating knowledge: a visual-pictorial channel and an auditory-verbal channel” (Mayer, 2002: 60). In this theory, the visual-pictorial channel is used for processing pictures that enter the cognitive system through the eyes as pictorial representations. On the other hand, the auditory-verbal channel is used for processing words that enter the cognitive system through the ears as verbal representations (see Mayer, 2002: 60). In the case of on-screen text, according to Mayer et al. (2001: 190), unlike in the case of narration or speech, the words enter “the visual channel via the eyes... and later can be converted to sounds... that are used to build a verbal

representation in the verbal channel". In all of this the limited capacity assumption and the related split-attention hypothesis of cognitive theory become increasingly important.

According to Mayer (2002: 60), the limited capacity assumption holds that "each channel in the human cognitive system has a limited capacity for holding and manipulating knowledge". In the presence of too many spoken words and sounds simultaneously, the auditory channel becomes overloaded, and likewise, "when a lot of pictures (or other visual material) are presented at one time, the visual channel can become overloaded" (Mayer, 2002: 60). It therefore stands to reason that the cognitive load may be further increased by the simultaneous presence of uninterrupted spoken words accompanied, or shadowed, by the written transcription of these words (or an edited version of these words) in the subtitles, as well as the words on slides, and that such a situation, when also combined with the non-verbal visual information in a lecture, may result in cognitive overload. In particular, since the subtitles place a further demand on the visual channel, this may result in the overloading of the visual channel, which could impact negatively on learning.

3.2 Cognitive load theory (CLT) and the redundancy effect

According to Paas, Renkl and Sweller (2004: 1), CLT "is mainly concerned with the learning of complex cognitive tasks, where learners are often overwhelmed by the number of information elements and their interactions that need to be processed simultaneously before meaningful learning can commence". Diao, et al., (2007: 237) define CLT as being "concerned with relationships between working and long-term memory and the effects of those relationships on learning and problem solving". They then proceed to summarise five principles of natural information processing systems that specify such systems as adapted to human cognition. These principles focus on the relationship between long-term and working memory in terms of information storage, schema and borrowing, limits of change in the working memory, and linking between long-term and working memory.

According to CLT, "efficient instructional designs... should be able to manipulate instructional materials and procedures to reduce extraneous or unnecessary cognitive load via the borrowing principle and enhance schema construction and automation" (Diao, et al., 2007: 239). The redundancy effect is one of the instructional procedures generated with the use of CLT.

According to Diao, et al., (2007: 239), "the redundancy effect occurs when the same information is presented to learners in different forms, requiring them to mentally coordinate the multiple forms. In some cases the learner may need to translate one form into the other to confirm that the two forms contain similar or identical information". An example from the material used in the current study would be when an idea unit is simultaneously present in the spoken words of the lecturer, in the subtitles (more or less verbatim and synchronised), and on the PowerPoint slides (in outline form). Students would then have to coordinate the three forms mentally, which may impose an extraneous cognitive load.

The following two figures illustrate the different visual sources of information in the text.

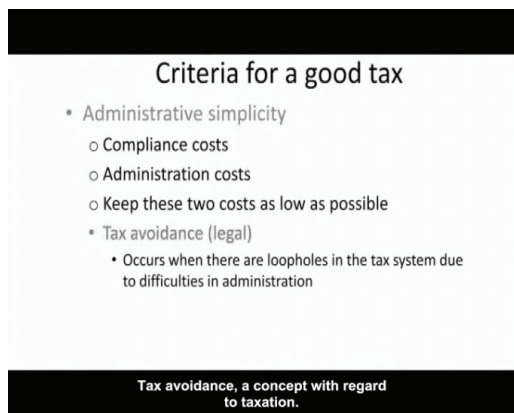


Figure 1: Slides and subtitles

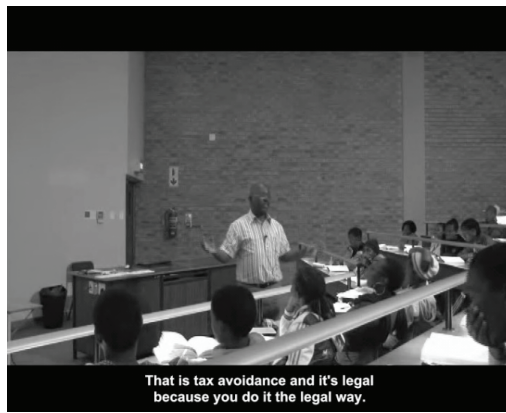


Figure 2: Visuals and subtitles

Importantly, according to Diao et al. (2007: 239), “having to attend to and coordinate spoken and written text results in slower learning than the use of one modality only”. They also mention the study by Moreno and Mayer (2002) that found that the redundancy effect was not present when auditory and written text were presented simultaneously in the absence of animation. This may be because the mental coordination required to match two verbal forms is less taxing than to coordinate verbal content with related visual content like a diagram.

What seems to emerge here is that the introduction of subtitles may very well be beneficial if this is not done together with complicated multimedia material. This would suggest that a subtitled recorded lecture, where the audiovisual material consists mainly of the verbal content of the lecture delivered by the lecturer (which is also available in the subtitles and in PowerPoint slides), may not produce such a cognitive overload. This is the case mainly because students do not have to process all three modes, but may focus on the one most suited to their style of learning. However, this has to be determined empirically.

The modality of input has received some attention in connection with subtitled video. In the context of native speakers of English, for example, Mayer et al. (2001) confirmed the redundancy effect when they found that adding subtitles caused students to split their visual attention between two sources, and to perform worse in terms of retention and transfer than students who had access to the video without subtitles. In particular, they argue that “the redundancy effect is consistent with the cognitive theory of multimedia learning and the split-attention hypothesis, which can be derived from it... The locus of the effect seems to be at the point of visual attentional scanning, as posited by the split-attention hypothesis” (Mayer, et al., 2001: 195). Nevertheless, Mayer et al. (2001: 196) caution that their findings confirming the redundancy principle “should not be taken as evidence that printed text and spoken text must never be presented together”.

This is confirmed by Diao et al (2007: 251) who established that, although “the redundancy effect does play a role in multimedia EFL instruction when students are learning to listen”, the simultaneous presentation of written and spoken material does facilitate comprehension and recall of information. Even in the field of multimedia learning and cognitive load theory, it would therefore seem that there is evidence that the double exposure to verbal content in auditory and written format may escape the redundancy effect under certain conditions.

Other studies have also established that subtitles do not necessarily result in cognitive overload, but that viewers can process information from various channels simultaneously (cf. for example, Perego, Del Missier, Porta and Mosconi, 2010). Specifically in the context of reading, it has been argued by various authors quoted in Perego et al., (2010: 245-246) that reading is largely an automatic activity for most adults because of learning processes, also in the case of subtitling (see d’Ydewalle et al., 1991; d’Ydewalle & Gielen, 1992; Koolstra, Van der Voort & d’Ydewalle, 1999). Furthermore, according to Perego et al., (2010: 243), “research on subtitles can ... be helpful to understand ... forms of multi-source communication in which the individual has to pay attention to information delivered through various sensorial channels and information sources”.

Mayer et al. (2001) illustrate the fact that speech and text could escape the redundancy effect with the example of PowerPoint presentations “in which a presenter both speaks and presents printed words on screen” which “can be effective even though words are presented in two modalities”. As reason for this, they cite the slower presentation rate of words in a PowerPoint presentation. As the current study will look at the use of subtitles added to a lecture in which PowerPoint slides are also used, it will have to be determined whether the redundancy effect does occur under these conditions.

3.3 Pedagogical implications and the South African context

From the studies discussed in this section it should be clear that the use of subtitles in an education context (such as language teaching or in support of other academic disciplines) is complicated by cognitive issues, but also by contextual and other issues such as medium of instruction, low literacy skills, complexity of the stimulus, and other factors. The fact that various studies have reported on the positive impact of subtitles (both SLS and interlingual) means that it is worth investigating as a pedagogical tool in the language classroom and other classrooms. However, the dangers of a redundancy effect and cognitive overload has to be managed carefully, something that can only be done once we know more about the reception of subtitled material in a particular context, such as an ESL tertiary environment in South Africa. Eye tracking provides one way of gaining a better understanding of cognitive load and attention allocation as will be explained in the next section.

4. Eye tracking as behavioural measure

Eye tracking has been used quite extensively in reading research. According to Irwin (2004: 94), “eye position may seem to be an ideal dependent variable because eye

movements are a natural and frequently occurring human behaviour; people typically look at something when they want to acquire information from it". They continue to state that, probably, "fixation location corresponds to the spatial locus of cognitive processing and that fixation or gaze duration corresponds to the duration of cognitive processing of the material located at fixation". This assumption underlies much eye tracking research and is sometimes called the eye-mind assumption (see Just and Carpenter, 1980, in Irwin, 2004: 94).

Holmqvist et al. (2011: 382) state that studies on reading, scene perception and usability all indicate "functional links between what is fixated and cognitive processing of that item – the longer the fixation, the 'deeper' the processing". In reading, lower frequency words attract longer fixations, as do complicated texts and grammatical structures. Important for this study from studies on scene perception, is the fact that objects that are out of context attract longer fixations. It would therefore seem that a higher cognitive load results when something that is unfamiliar or unexpected appears on screen. In this study the unfamiliar terminology of an Economics lecture to students not majoring in Economics may therefore be expected to attract longer fixations, which should be visible in relation to the slides and subtitles.

Holmqvist et al. (2011) also note that fixation counts in a specific area of interest (or AOI) have, among others, been shown to increase under conditions of higher semantic importance, under conditions where the search task is complicated (presumably therefore also when there is high competition between sources of information), and under conditions where a fixated object appears repeatedly over a period of time resulting in the build-up of memory with fewer fixations.

Due to the fact that there is quite a strong competition effect and potential redundancy effect in the investigation of the eye movements of participants looking at a subtitled recorded academic lecture with slides, the focus in this study will be on attention allocation according to specific eye tracking measures. In view of the above and the findings by Perego et al. (2010) that subtitles attracted a higher number of fixations, but that the mean fixation duration in the subtitled area was significantly shorter than in the visuals, the mean fixation duration as well as the fixation count of a particular AOI will be investigated. Dwell time, according to the BeGaze 2.4 manual (SensoMotoric Instruments, 2010: 161), "starts at the moment the AOI is fixated, and ends the moment the last fixation on the AOI ends", or it can be defined as the "sum of durations from all fixations and saccades that hit the AOI". This attentional measure therefore includes both visual intake and no-visual-intake measurements because of the evidence that some cognitive processing also takes place during saccades as no-visual intake events.

These three measures will all be weighted. Dwell time will be investigated as percentage of the visible time of each AOI (*weighted dwell time* = dwell time for AOI in ms divided by visible time of AOI in ms times 100). Fixation count will be investigated as number of fixations per second of the visible time of each AOI (*weighted fixation count* = fixation count for AOI divided by the visible time of the AOI in ms x 1000). *Mean fixation duration* is already standardised to the mean duration in ms of fixations

in each AOI (sum of all fixation durations in AOI divided by number of fixations in AOI). The reason for these weighted measures is that other studies that investigate subtitle processing through eye tracking typically compare the attention distribution between subtitles and visuals only; both AOIs being on screen for the full duration of the clip. In this case, there was written text in the subtitles as well as the PowerPoint slides, and visual attention was divided between subtitles, slides and visuals of the lecturer when there were no slides.

5. Experimental design

5.1 Aims and hypotheses

Perego et al. (2010: 246) point out that “studies on subtitle processing usually do not provide a comprehensive analysis of processing strategies and of their effectiveness, because they do not take into account at the same time eye movements, visual scene processing performance, and subtitle processing performance”. The current experiment was therefore designed in such a way that a combination of comprehension measures or processing performance and eye movements could be studied. There were two sets of dependent variables, namely the behavioural measures of eye movements, and the cognitive measure of comprehension.

The first aim of the article is to determine whether the attention allocation of participants is related to their comprehension (whether participants who performed better on the comprehension test looked at the different information sources differently).

- *Hypothesis 1* is that a positive correlation will obtain between the eye tracking measures on the subtitles and slides and comprehension scores, and a negative correlation will obtain between the eye tracking measures on the visuals and comprehension scores.

This aim can be refined by determining whether the presentation rate of subtitles has an impact on the attention distribution of participants.

- *Hypothesis 2* is that there will be a statistically significant difference in the attention distribution between groups that are exposed to subtitles at different presentation rates, with the difference in attention allocation to subtitles and slides as well as to subtitles and visuals (favouring subtitles for dwell time and fixation count, and favouring slides and visuals for mean fixation duration) being statistically significantly greater when the subtitle presentation rate is higher.

Should Hypothesis 1 hold true, it would mean that subtitles and other visual support like text on slides result in increased cognitive processing, making them useful tools in ESL support at all educational levels. Should Hypothesis 2 hold true, it would mean that the South African ESL context may require lower presentation rates to prevent cognitive overload.

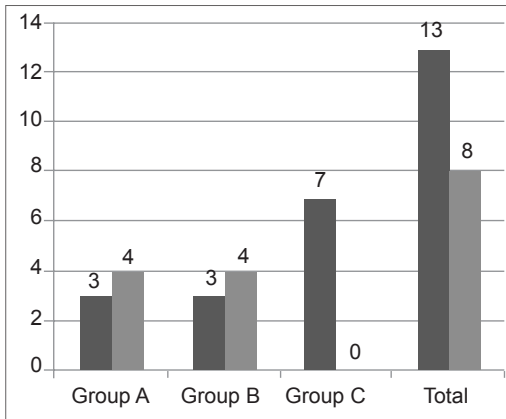
5.2 Material and sampling

The material for the study was a recorded lecture from the first semester of Economics I on the Vaal Triangle Campus of North-West University (NWU), dealing with fiscal policy and taxation. The class was edited down to the main content with all introductory greetings and non-related information edited out. The final video was approximately 30 minutes in duration.

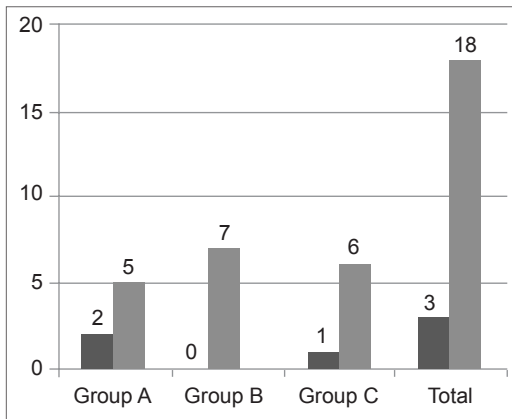
Since this was an introductory lecture during the first semester of 2011, it is assumed that an audience without a background in Economics should be able to understand the basic concepts, but would not be able to reach the same levels of comprehension as students who take Economics. To control for the influence of prior knowledge, the participants for the study were recruited from the first-year English I class (consisting of 160 students) who do not typically take Economics. After an open invitation to the full group, a total of 21 students agreed to participate and they were assigned randomly to one of the three treatment groups until there were 7 in each group. This was done by assigning the first participant to Group A, the second to Group B, the third to Group C, and then starting at Group A again with the fourth participant, and so on.

All three groups saw the video (delivered in English) with English subtitles. In order to investigate the impact of presentation rate, it was decided to subtitle the video as *near verbatim*, *somewhat reduced* and *substantially reduced* renditions of the speech of the lecturer. In all three conditions the subtitles remained in full sentences. The respective presentation rates were calculated afterwards as an average of 132wpm (Group A), 110wpm (Group B) and 74wpm (Group C). This speed was calculated by dividing the sum of the display time of all the subtitles, by the total number of words for each version of the video. A minimum gap of 2 frames were left between consecutive subtitles, and each subtitle consisted of a maximum of two lines displayed centred at the bottom of the screen (see Figure 1 and 2).

The average age of participants was 21 with the majority of participants ranging from the minimum age of 18 to 22 years of age, with only one participant being older than 22, at 45 years of age. The home language and gender distribution of the three groups can be seen in the following graphs:



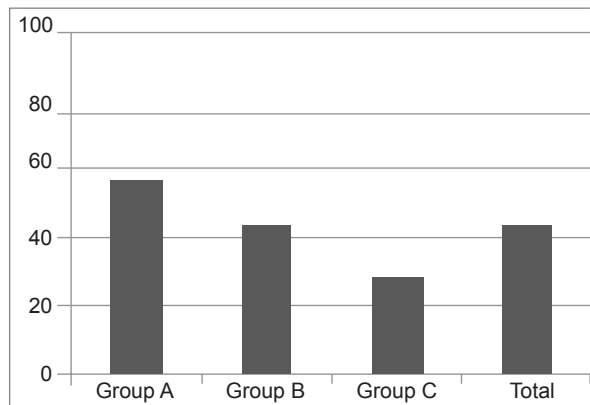
Graph 1: Language distribution



Graph 2: Gender distribution

Since language background could very well be a confounding variable in this context, the academic literacy levels of participants, as evidenced by the results of participants in the “Test for Academic Literacy Levels” (TALL)¹ were also compared in order to ensure the comparability of the groups.

All first-year students at the NWU write this test at the beginning of their studies. The average TALL scores for the three groups can be seen in Graph 3.



Graph 3: Academic literacy distribution

Since the TALL test provides an indication of the academic literacy levels of students, and has been found to be a good indicator of academic success or students’ receptive abilities (ability to understand academic material) in an academic context, this might very well also prove to be a confounding variable.

¹ See Van Dyk & Weideman (2004), Weideman (2006) and Butler (2007) for a description of this test.

A score of above 50% on the TALL is typically used as an indication that a student has reached a sufficient level of academic literacy not to require intervention in the form of the basic academic literacy course at the institutions using the evaluation. On the frequencies alone it would therefore seem that groups A and B lie just above and just below the threshold, with Group C substantially below. To determine whether the observed differences in means between the three groups are significant, a one-way ANOVA was performed on the TALL scores. A Shapiro-Wilk W test for normality confirmed that the TALL scores were normally distributed, with $p > 0.05$, making an ANOVA possible. The ANOVA indicated that there is a statistically significant difference in main effect for the variable group (stimulus), with $F(2, 18) = 6.4$, $p < 0.05$.

To determine where this difference is located, independent-sample t-tests by group were performed, indicating a statistically significant difference between the TALL scores of Group A (132wpm) ($M = 57.14$, $SD = 14.29$) and Group C (74wpm) ($M = 28.86$, $SD = 6.96$); $t(12) = 4.71$, $p < 0.05$. The effect size for this difference is medium ($r = 0.78$, $d = 2.5$). Based on the fact that Group C consisted only of African language respondents, as well as the difference between Group A and Group C in terms of the TALL scores, it was decided to discard Group C for the purposes of the data analysis. Although this means that the data of only 14 participants could be used, this number is in line with the number of participants used in similar eye tracking studies involving subtitles (see Perego et al., 2010 who used 16 participants, and references therein). Since the number of participants was already low (necessitated by the fact that collecting eye tracking data is particularly time consuming), this was considered a necessary step in order to ensure at least some validity in terms of the findings with Group C potentially skewing the results because of the confounding variables of academic literacy levels and home language. Since low academic literacy levels in themselves are an important consideration in designing interventions, further studies are under way to investigate the impact of subtitles on cognitive load for such groups in particular although it falls outside the scope of the current study.

5.3 Experiment

Participants reported to the eye tracking laboratory one by one. When they arrived, the experiment was explained to them and they had the opportunity to read and sign the informed consent form, which formed part of the ethics clearance for the project with the NWU. In order not to influence their behaviour, they were simply told that they would be watching a recorded lecture on economics while their eyes are recorded, and that they would answer a few questions on the content afterwards. They were told that the aim of the experiment was to determine how students process educational material (without calling any attention to the subtitles in the instructions).

Each participant was then asked to sit in a comfortable position in front of the stimulus monitor, at a distance of approximately 700mm from the stimulus screen. Their eye movements were monitored and recorded using the SMI iViewX RED eye-tracking system. This eye-tracking system works with a built-in camera with a sampling rate of 50Hz. The detection of events in the RED system uses fixations as the primary events, with the algorithm that is used to detect fixations being based on dispersion. Fixation

detection parameters are set to detect a fixation when the eyes remain in the same area (smaller than 100px, the maximum dispersion) for a period of more than 80ms (SensoMotoric Instruments, 2010). The screen resolution was 1280px by 1024px. This eye tracker allows more freedom and resembles the natural computer viewing situation more closely than high speed eye trackers with chin and forehead rests. Each participant's eyes were then calibrated in iViewX using a 9-point calibration and validated to ensure the accuracy of the data. Participants were also instructed to sit as still as possible to ensure accuracy, although the system does allow for some movement. The average calibration deviation for participants on the X-axis was 0.4 degrees, and on the Y-axis, 0.6 degrees. The average tracking ratio or quality² for participants across the trial was 91%. After having watched the video, subjects were given a 10-mark questionnaire with four items to complete.

5.4 Data

In order to determine the attention allocation of participants when watching the subtitled video recording, areas of interest (or AOIs) were marked in SMI's BeGaze 2.5 (SensoMotoric Instruments, 2010), and the behaviour of the eyes of subjects that had previously been recorded using the SMI iViewX RED eye-tracking system was then analysed with respect to these AOIs. The first AOI was the full subtitle area for the entire video (referred to simply as "subtitles" in the rest of the article), with the second AOI being the full PowerPoint slides ("slides"). The final AOI that was defined is that part of the screen not taken up by either the subtitles or the slides, and therefore containing visuals of the lecturer presenting the lecture ("visuals"). The slides and the visuals were therefore toggled on and off, while the subtitles remained marked for the entire duration. This was done in order to compare the attention distribution between subtitles, the slides, and the visuals.

The data obtained from the comprehension test and the TALL test, as well as the data obtained by means of the eye tracking, and converted to event and AOI statistics in BeGaze 2.5, were then subjected to statistical analyses using Statsoft's Statistica 10 (StatSoft Inc., 2012). In the case of t-tests that yielded statistical significance, effect size was calculated with Cohen's d, with a small effect obtaining when $0.2 \leq r < 0.5$, a medium effect when $0.5 \leq r < 0.8$, and a large effect when $r \geq 0.8$. The difference in standard deviation is pooled³.

The questionnaire consisted of four questions. The first question is included in this article in order to illustrate the use of subtitles as well as the different sources of information. Table 1 presents the first question, the two different sets of subtitles, the transcription of the lecturer's words in the auditory channel, the information presented on the slides, and the answers to the question used to grade the replies.

2 Tracking quality is the number of non-zero gaze positions divided by the sampling frequency, multiplied by the run duration, expressed in percent (see SensoMotoric Instruments, 2010).

3 The pooled standard deviation is used here (see Hedges and Olkin, 1985) and not just one of the two standard deviations as in Cohen's original formula (see Cohen, 1988).

Table 1: Question 1

Question1:	Name two kinds (instruments) of fiscal policy (2)
Information in subtitles (110wpm)	00:07:16:06-00:07:16:09 So this fiscal policy... 00:07:18:10-00:07:18:18 consists mainly of 2 elements. 00:07:21:07-00:07:21:10 You've got contractionary fiscal/policy... 00:07:25:15-00:07:25:19 And then you have expansionary fiscal/policy.
Information in subtitles (132wpm)	00:07:16:06-00:07:16:09 So this fiscal policy... 00:07:18:10-00:07:18:18 consists mainly of 2 elements. 00:07:20:24-00:07:21:10 You've got what is referred to as/ contractionary fiscal policy... 00:07:25:13-00:07:25:19 And then you also have expansionary/ fiscal policy.
Information in soundtrack	So this fiscal policy consists mainly of 2 elements. You've got what is referred to as contractionary fiscal policy, and then you also have expansionary fiscal policy.
Information in PowerPoint	FISCAL POLICY (Instruments) CONTRACTIONARY FISCAL POLICY EXPANSIONARY FISCAL POLICY
Answer:	a) Contractionary (fiscal policy) b) Expansionary (fiscal policy)

The remaining questions and answers are presented in Table 2 below.

Table 2: Questions and memorandum (2 to 4)

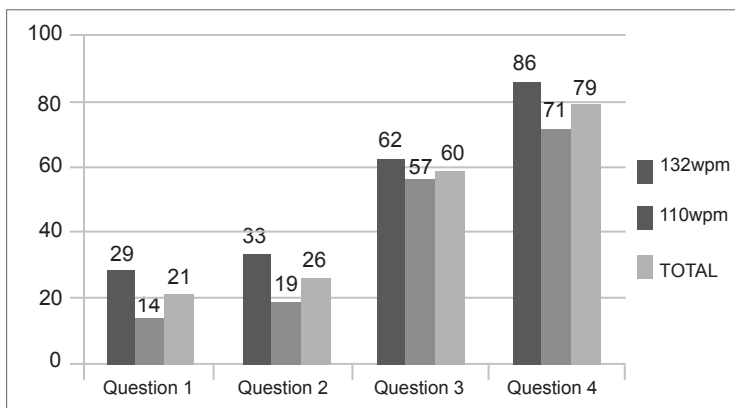
Question	Answer
2. List 3 criteria for good tax (3)	Any 3 of a) Neutrality b) Administrative simplicity c) Fairness/equitability d) Should not be a disincentive
3. True/false (3): 3a. Tax avoidance is legal 3b. Tax evasion is illegal 3c. Indirect taxes are levied on persons, individuals or companies	3a) True 3b) True 3c) False
4: Give examples of direct and indirect taxes (2)	Any one each of the following: a) Direct: income tax/company tax/estate duty b) Indirect: VAT/ customs duties/excise duties

The questions were compiled by the post-graduate academic assistant for the course, and checked by the lecturer for accuracy and to ensure a good distribution of difficulty level. Question 1 involved straightforward retention of facts (although the terminology would be unfamiliar to participants). Question 2 required more than retention since participants had to understand the concept of what constitutes good tax from the examples used by the lecturer to explain the criteria. Question three consisted of three true or false statements. This question is likewise not dependent only on retention, but would involve some insight. However, there was a chance that participants may have been able to provide the correct answer by chance or based on logic. Question 4 required the retention of some examples discussed by the lecturer. The questions are also posed in the same order in which the information pertaining to the questions was provided in the lecture, with the result that the information for the first two questions may not have been as readily available to the short-term memory of students as the information for the last two questions.

6. Results

6.1 Presentation rate and comprehension

The results for the four questions (converted to percentages) are provided below in Graph 4.



Graph 4: Average comprehension scores by question

From the descriptive statistics it seems that there is indeed a difference in comprehension between the two groups, with the mean comprehension score for Group A (132wpm) being 51.43%, and that for Group B (110wpm) being 40%. However, a t-test for independent samples on the comprehension scores indicated that this difference is not statistically significant between the two groups. This would seem to indicate that presentation rate of subtitles in this experiment did not have a significant impact on comprehension, and that

the difference in scores may simply have been the result of chance or inter-individual differences in academic literacy.

6.2 Correlations between comprehension and attention allocation (Hypothesis 1)

Following the lack of statistical significance between the two groups in terms of comprehension, the behavioural eye tracking measures were correlated with the comprehension scores, in order to determine whether a relationship obtains between comprehension and attention allocation. Pearson product-moment correlation coefficients were therefore calculated to assess these relationships.

The only statistically significant correlation that could be found for the full group was between comprehension and the weighted fixation count on the slides, with $r = 0.58$, $n = 14$ and $p < 0.05$. In other words, more fixations on the slides generally resulted in increased comprehension. This confirms previous findings where deeper processing was linked to more fixations (see Holmqvist et al., 2011). The only other measure that approaches (but fails to reach) a statistically significant correlation with comprehension for the full group is the weighted dwell time in the slides, with $r = 0.49$, $n = 14$ and $p = 0.07$, which again seems to confirm expectations from the literature. For the group as a whole, Hypothesis 1 is therefore partially supported, but only in the case of the slides.

However, there were no statistically significant correlations for the group that saw the subtitles at 132wpm between comprehension and eye tracking measures. For this group, Hypothesis 1 is therefore not supported.

In the case of the 110wpm stimulus, no statistically significant correlations could be found between comprehension and the eye tracking measures on the subtitles, although there were weak negative correlations (with r between 0.2 and 0.3). This would almost suggest a trend that the more the subtitles were attended to by the 110wpm group, the lower their comprehension was. For Group B, Hypothesis 1 is therefore not supported with respect to the subtitles.

In terms of the two other AOIs, there were positive correlations between comprehension and all three eye tracking measures in terms of the visuals, and two of the three measures in the case of the slides, although only the mean fixation duration in the case of the visuals yielded statistical significance. The correlations between comprehension and these measures were:

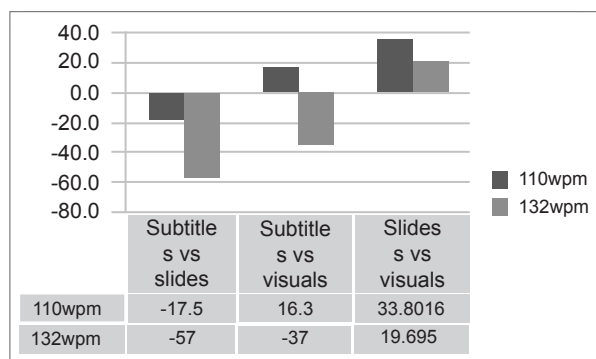
- Slides: weighted fixation count ($r = 0.66$, $n = 7$, $p = 0.11$)
- Slides: weighted dwell time ($r = 0.67$, $n = 7$, $p = 0.10$)
- **Visuals: mean fixation duration ($r = 0.79$, $n = 7$, $p < 0.05$)**

- Visuals: weighted fixation count ($r=0.4$, $n=7$, $p=0.38$)
- Visuals: weighted dwell time ($r=0.59$, $n=7$, $p=0.17$)

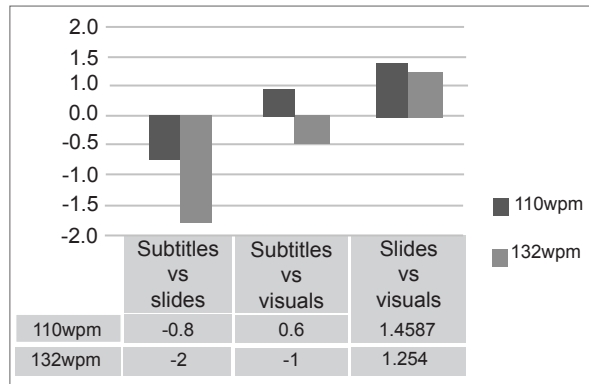
These findings seem to suggest that, for the group that saw the subtitles at 110wpm, comprehension or retention increased when participants attended to both the slides and the visuals (containing contextual non-verbal clues) whereas the added cognitive load of subtitles had a negative impact on comprehension or retention. In particular, at the slower rate, the longer mean fixation durations in the visuals resulted in higher comprehension. For Group B, Hypothesis 1 is therefore only supported in the case of the slides (as for the group as a whole), but not for the visuals or subtitles, where more attention rather than less attention resulted in higher comprehension.

6.3 Presentation rate and attention distribution (Hypothesis 2)

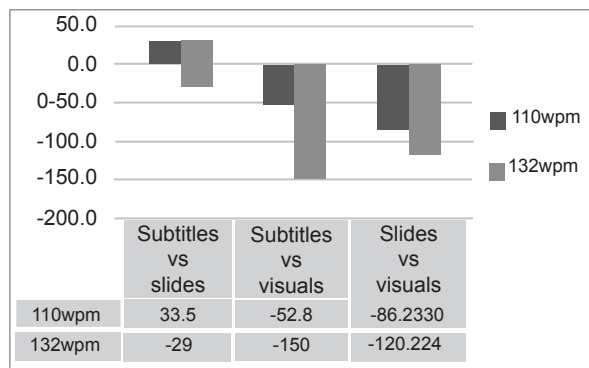
In order to determine the impact of presentation rate on attention allocation, the descriptive statistics reveals the following distribution of means in terms of the difference in attention distribution between subtitles and slides, subtitles and visuals, and slides and visuals for the three measures:



Graph 5: Difference in weighted dwell time between AOIs



Graph 6: Difference in weighted fixation count between AOs



Graph 7: Difference in mean fixation duration between AOs

To determine the significance of these differences in attention allocation, independent-sample t-tests were run, yielding statistical significance on the difference in measures in the following cases:

- **Weighted dwell time on subtitles vs. slides** between Group A (M=-57%, SD=33.1), and Group B (M=-17.5%, SD=33.9) with $t(12)=-2.2$, and $p<0.05$. This difference has a medium effect size (Cohen's $d = -1.2$, $r=-0.5$).
- **Weighted dwell time on subtitles vs. visuals** between Group A (M=-37%, SD=47.6) and Group B (M=16.3%, SD=24.3), with $t(12)=-2.6$, and $p<0.05$ (medium effect size with Cohen's $d = -1.4$, $r=-0.6$).

For both the attention distribution between subtitles and slides, and subtitles and visuals, a higher presentation rate therefore resulted in a statistically significantly lower weighted dwell time on the subtitles than on slides or visuals. Hypothesis 2 is therefore not supported in terms of the weighted dwell time.

- **Weighted fixation count on subtitles vs. visuals** between Group A (M=-0.5, SD=1.2) and Group B (M=0.6, SD=0.6) with $t(12)=-2.3$ and $p<0.05$, with a medium effect size (Cohen's $d = -1.2$, $r=-0.5$).

This seems to indicate that a higher presentation rate resulted in a statistically significantly lower weighted fixation count on the subtitles than on the visuals, which is contrary to the findings of Perego et al. (2010) who found significantly more fixations on subtitles. Hypothesis 2 is therefore not supported.

- **Mean fixation duration on subtitles vs. slides between Group A (M=-29, SD=37.8), Group B (M=33.5, SD=40.4), $t(12)=-3$, $p<.05$ (with medium effect size; Cohen's $d =-1.6$, $r=-0.6$).**
- **Mean fixation duration on subtitles vs. visuals** with Group A (M=-150, SD=91.7), Group B (M=-52.8, SD=62.5), $t(12)=-2.3$, $p<0.05$ (with medium effect size; Cohen's $d = -1.2$, $r=-0.5$).

The faster the presentation rate of the subtitles, the shorter the fixations seem to be on the subtitles, and the longer they tend to be on the slides. The same applies to subtitles vs. visuals. The difference in presentation rate therefore seems to make a significant difference in this distribution between the groups. Hypothesis 2 is therefore supported for the attention distribution between subtitles and slides, and subtitles and visuals, but not for slides and visuals.

7. Conclusions

This study set out to test whether the presentation rate of subtitles has an impact on cognitive load, or whether the presentation rate of subtitles used in an educational context has an impact on comprehension. It would appear that in this exploratory study that the presentation rate does not impact on comprehension. However, further studies are required with larger samples and with a control group who watch the recording without subtitles to determine whether these preliminary findings can be generalised.

Although there were no statistically significant correlations between comprehension and any of the three eye tracking measures on any of the three sources of information in the case of the group that saw the subtitles at 132wpm, in the case of the 110wpm group there were some correlations. Weak negative correlations were found between comprehension and eye tracking measures on subtitles, which would seem to support the redundancy effect hypothesis of CLT. More interestingly, for this group, strong positive correlations were found between comprehension and attention allocation to the slides and the visuals. This suggests two things: first, it would seem that a lower presentation rate results in a stronger relationship between comprehension and attendance to salient visual codes like slides and a talking person who provides contextual support, with subtitles seeming to result in cognitive overload. When the presentation rate is higher, this relationship seems to be absent, in other words, it is not possible to state with certainty whether participants got information from listening to the lecturer or reading the

slides or the subtitles. Secondly, for the group as a whole it would appear that students tend to get more information from the slides than from the subtitles. This makes sense if the slides provide a succinct summary that is not prone to the wandering of unscripted speech in the subtitles and the spoken words of the lecturer.

The investigation of eye tracking measures related to the three visual AOIs between the two different stimuli (132 wpm and 110wpm) was done to test the hypothesis that the presentation rate would impact on the eye behaviour. From the analysis of the data it appears that a higher presentation rate caused participants to shun the subtitles in favour of the slides and the visuals, rather attempting to get information from the contextual signs in the visuals or the more succinct verbal text in the slides. It would therefore appear that the presentation rate of subtitles did have an impact on eye behaviour, with the faster presentation rate causing participants to favour the visuals and to some extent the slides over the subtitles.

The final part of this study concerned the impact of presentation rate on attention distribution. There was a significant difference in attention distribution (between slides and subtitles) between the two groups, with the direction of the difference favouring the slides for Group A. Group A also spent statistically significantly more time in the visuals than the subtitles (with Group B spending more time in subtitles than visuals), suggesting that a higher presentation rate impacts negatively on the processing of subtitles.

In the case of mean fixation duration, there was statistical significance in the difference between the two groups relating to the competition between subtitles and slides as well as subtitles and visuals. In the case of the competition between subtitles and slides, Group A had shorter fixations in the subtitles, whereas Group B had shorter fixations in the slides. It would therefore seem that the higher presentation rate had a negative impact on the processing of the subtitles in the comparison of these two textual sources of information. In the case of the competition between subtitles and visuals both groups had shorter fixations on the subtitles than on the visuals, in line with earlier studies, but the difference in mean fixation duration for Group A was significantly bigger than for Group B, favouring the visuals. This once again seems to indicate that presentation rate has a negative impact on the processing of subtitles.

Although presentation rate therefore seems to have an impact on the processing of difference sources of information in a subtitled lecture that also makes use of PowerPoint slides, and this presentation rate seems to impact negatively on the processing of subtitles, more studies are required on the competition between these sources of information. If the attention distribution is to be correlated with comprehension in a more specific manner that will allow researchers to pinpoint the nexus of the various cognitive processes, more attention will have to be devoted to the testing of comprehension and retention, perhaps also involving more qualitative research measures.

In real terms the study provides preliminary indications that subtitles, and particular faster subtitles that allow less time for reading, result in higher cognitive load, supporting the redundancy hypothesis. Nevertheless, the potential benefits of subtitles, particularly in

terms of the benefits of dual coding, means that his mode has to be explored in more detail. For teachers in an ESL context in South Africa it would be advisable to limit redundancy in using subtitled material by using material in which the subtitles are presented at a lower rate. This type of material is available commercially on some DVDs, and can also be created with the assistance of fairly low-cost software and even freeware that is increasingly becoming available online. It would be important to conduct specific research in order to arrive at optimal presentation speeds and information density in subtitles.

Finally, the manipulation of the sources of information (with more control groups in which the competition between slides and visuals and between subtitles and visuals is also investigated separately, and with a specific focus on the benefits to low-literacy learners) would bring us closer to answers concerning the impact of each of the sources of information on the processing of academic lectures, and the possible effect of language on learning in university classrooms.

References

- Ayonghe, L.S. 2010. *Subtitling as an aid in academic literacy programmes: The University of Buea*. Unpublished PhD thesis. Vanderbijlpark: NWU.
- Bird, S.A. & Williams, J.N. 2002. The effect of bimodal input on implicit and explicit memory: An investigation into the benefits of within-language subtitling. *Applied Psycholinguistics* 23(4): 509-533.
- Butler, H.G. 2007. *A framework for course design in academic writing for tertiary education*. Unpublished PhD thesis. Pretoria: University of Pretoria.
- Cohen, J. 1988. *Statistical power analysis for the behavioral sciences*. (2nd Edition). Hillsdale, NJ: Lawrence Earlbaum Associates.
- Danan, M. 2004. Captioning and subtitling: Undervalued language learning strategies. *Meta: Translators' Journal* 49(1): 67-77.
- 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., Praet, C., Verfaillie K. & Van Rensbergen, J. 1991. Watching subtitled television. Automatic reading behaviour. *Communication Research* 18: 650-666.
- D'Ydewalle, G. & Gielen, I. 1992. Attention allocation with overlapping sound, image, and text. In: Rayner, K. (Ed.) 1992. *Eye movements and visual cognition: Scene perception and reading*. New York: Springer-Verlag. pp. 415-427.

- Garza, T.J. 1991. Evaluating the use of captioned video materials in advanced foreign language learning. *Foreign Language Annals* 24(3): 239-258.
- Hedges, L.V. & Olkin, I. 1985. *Statistical methods for meta-analysis*. Orlando: Academic Press.
- Heugh, K. 2002. Recovering multilingualism: Recent language-policy developments. In: Mesthrie, R. (Ed.) 2000. *Language in South Africa*. Cambridge: Cambridge University Press. pp. 449-475.
- Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H. & Van de Weijer, J. 2011. *Eye tracking: A comprehensive guide to methods and measures*. Oxford: Oxford University Press.
- Huang, H. and Eskey, D. 1999. The effects of closed-captioned television on the listening comprehension of intermediate English as a Second Language students. *Education Technology Systems* 28: 75-96.
- Irwin, D.E. 2004. Fixation location and fixation duration as indices of cognitive processing. In: Henderson, J.M. & Ferreira, F. (Eds.) 2004. *The interface of language, vision, and action: Eye movements and the visual world*. New York, NY: Psychology Press. pp. 105-133.
- Koolstra, C.M., Van der Voort, T.H. & d'Ydewalle, G. 1999. Lengthening the presentation time of subtitles on television: Effects on children's reading time and recognition. *Communications* 24: 407-422.
- Lacroix, F. 2012. *The impact of same-language subtitling on student comprehension in an English as an Additional Language (EAL) context*. Unpublished MA dissertation. Vanderbijlpark: North-West University.
- Linebarger, D., Piotrowski, J.T. & Greenwood, C.R. 2010. On-screen print: the role of captions as a supplemental literacy tool. *Journal of Research in Reading* 33(2): 148-167.
- Markham P.L. 1999. Captioned video-tapes and second language listening word recognition. *Foreign Language Annals* 32-3: 321-328.
- 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.
- 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.

- Moreno, R. & Mayer, R.E. 2002. Verbal redundancy in multimedia learning: When reading helps listening. *Journal of Educational Psychology* 94(1): 156-163.
- Paas, F., Renkl, A. & Sweller, J. 2004. Cognitive Load Theory: Instructional implications of the interaction between information structures and cognitive architecture. *Instructional Science* 32: 1-8.
- Paivio, A. 1986. *Mental representations: A dual coding approach*. Oxford: Oxford University Press.
- Paivio, A. 1991. Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology* 45: 255–287.
- Paivio, A. 2007. *Mind and its evolution: A dual coding theoretical approach*. Mahwah, NJ: Erlbaum.
- Perego, E., Del Missier, F., Porta, M. & Mosconi, M. 2010. The cognitive effectiveness of subtitle processing. *Media Psychology* 13(3): 243-272.
- Sadoski, M., Paivio, A. & Goetz, E.T. 1991. A critique of schema theory in reading and a dual coding alternative. *Reading Research Quarterly* 26(4): 463-484.
- SensoMotoric Instruments. 2010. BeGaze™ manual: version 2.4. Teltow: SMI.
- StatSoft Inc. 2012. Statistica (data analysis software system), version 10.0. www.statsoft.com.
- Sydorenko, T. 2010. Modality of input and vocabulary acquisition. *Language Learning & Technology* 14(2): 50-73.
- Van Dyk, T. & Weideman, A. 2004. Finding the right measure: From blueprint to specification to item type. *Journal for Language Teaching* 38(1): 15-24.
- Vanderplank, R. 1988. The value of teletext sub-titles in language learning. *English Language Teaching (ELT) Journal* 42(4): 272-281.
- Vanderplank, R. 1990. Paying attention to the words: Practical and theoretical problems in watching television programmes with uni-lingual (CEEFAX) sub-titles. *System* 18(2): 221-234.
- Weideman, A. 2006. Assessing academic literacy in a task-based approach. *Language Matters* 37(1): 81-101.
-

ABOUT THE AUTHOR

Jan-Louis Kruger

North-West University (Vaal Triangle Campus)
PO Box 1174, Vanderbijlpark, 1900, South Africa
JanLouis.Kruger@nwu.ac.za

Jan-Louis Kruger is associate professor in, and Director of the School of Languages on the Vaal Triangle Campus of North-West University. His research focuses on reception studies in AVT including the use of behavioural measures (eye tracking and EEG) combined with performance measures (comprehension, immersion) in fiction film as well as educational material. He is a co-editor of *Perspectives: Studies in Translatology*.