

Post-event Information and False Memories

Thea Rolskov Sloth (201706833), Kiri Koppelgaard (201708294), and Signe Kirk Brødbæk (201707519)
Cognition and Communication, Cognitive Science, Arts, University of Aarhus
Jens Chr. Skous Vej 2, 8000 Aarhus, Denmark
Supervisor: Kristian Tylén

Abstract (TRS, KK)¹

This paper investigates the effect of post-event information and priming regarding false memories. The basis of this experiment is investigations done by Loftus and Palmer (1974), who tested the effects of leading verbs and Loftus and Zanni (1975), who tested the effects of leading articles. Additionally, this experiment tests the interaction effect of these. We hypothesise that when being misled by leading verbs, participants will be influenced in how they estimate speed. Furthermore, when presented with definite articles, participants will be more prone to remember elements, such as a broken windscreen, that was not present. Lastly, we hypothesise that the leading verb *smash* and a definite article combined will have the strongest effect on falsely remembering a broken windscreen. Two experiments in Danish were conducted to test these hypotheses. Participants watched a short video of a car crash and received post-event information in the shape of differently phrased questions. The data was analysed with a t-test, simple logistic regression and a 2x2 factorial logistic regression model in RStudio. The results of this study suggest no significant impact of post-event information neither on speed estimation nor falsely remembering a broken windscreen.

Keywords: False memories; post-event information; priming; fuzzy-trace theory

(KK, TRS) Memories may seem unchangeable and solid. Yet there is evidence, that memories are flexible and affectable (e.g. Loftus & Palmer, 1974; Loftus & Zanni, 1975). Through misleading suggestions, they can be altered and make us remember things differently than how an event truly unfolded; the creation of so called *false memories*. A false memory is the disturbance of a memory leading to a misinterpretation of the original memory.

Challenges of false memories

(KK) The phenomenon, false memories, has shown to be a challenge under various

circumstances. During the 1990s, a vast debate on the accuracy of recovered memories of childhood sexual abuse raged (Schacter, 1999, p.182-183). Patients started recovering repressed memories of childhood abuse during therapy sessions. Some of the recovered memories appeared to be accurate, but there is reason to believe that others were products of suggestive treatment, which ended in devastating consequences for the relatives (Lindsay & Read, 1994).

Likewise, false memories have given rise to challenges in the legal system. People being wrongfully convicted of crimes due to false eyewitness testimonies is a reality today.

¹ The primary author is listed first, the secondary author(s) is listed second.

Amongst others this was the case for Steve Gary Titus (Henderson, 1981). Resembling the perpetrator and being near the crime scene at the time of the rape, he was treated as a suspect and put in a line-up. The rape victim expressed prior to the trial that Titus resembled the perpetrator “the most”. However, when the victim took the stand, the utterance went from strongly suggestive to a direct determination of Titus as the perpetrator. The victim’s memory of the perpetrator had been distorted resulting in a false memory.

(TRS, KK) The topic of leading questions in court is an ongoing issue in legal systems. In Denmark, lawyers are especially attentive of the uncertainties related to eye-witnesses and the way they are questioned. E.g. one experienced Danish criminal lawyer, Bjarne Frøberg, announced that he ascribes no confidence in testimonies whatsoever due to their high rate of inaccuracies (Hauerslev, 2014). The approach to eyewitness testimonies is under re-evaluation, which is expressed in an issue of the Danish magazine, *Advokaten* (Hauerslev, 2014).

Likewise, this is relevant for the legal system in the United States, where Titus was convicted. Especially, since a jury must be convinced in court (Simpson & Selden, 1998). Leading questions are under discussion, and awareness is raised around this issue as they can have fatal consequences for wrongfully convicted individuals (Simpson & Selden, 1998).

The debate of false memories has been followed by extensive research on both memory and false memories in which people confidently claim to recall events that did not happen or recall them differently than how they occurred (e.g. Loftus & Palmer, 1974).

Bartlett’s *War of ghosts*

(KK) One of the scientists that has investigated the topic of memory distortion is Bartlett (1932). He hypothesised that memory is reconstructive and that humans encode and retrieve information according to their expectations which are formed by cultural schemas (Bartlett, 1932, pp. 60-120).

Bartlett presented the story of *The War of Ghosts*, a native American legend, to British participants and requested them to either reproduce the story regularly to themselves or to recall the story and repeat it to another person.

Bartlett concluded that there was no significant difference in the way the two groups recalled the story. However, both groups distorted the story, when trying to remember it. The participants in general remembered the main themes of the story but altered the features they found unfamiliar to match their own cultural expectations. According to Bartlett (1932), remembering is an active process, where information is retrieved and altered to accommodate the cultural schemas and make sense of the perceived information. The study founded the base for Bartlett’s later theory of reconstructive memory, which states that memories should be treated as reconstructions of reality rather than copies of experiences.

Memory is far from a recording device, that can be turned on and off. In studies by e.g. Loftus and Palmer, elaborated on later in this report, it has become clear that it is possible to influence the answer to a question by post-event information and by the wording of a question. According to Loftus and Palmer, the wording affects the cognition by influencing people’s memory of the incident. Therefore, the nature of language is

crucial when discussing false memories.

Linguistic relativity and determinism

(KK, SKB) In 1929, Sapir presented the idea that cognition is affected by language. He stated that it is an illusion that our sense of reality is independent of our use of language, and that language solely is a tool for solving problems of communication. The real world is in fact to a large extent constructed by the language habits of groups we belong to (Sapir, 1929, p. 209). According to Sapir, the linguistic categories and structure of language influences our non-linguistic cognition.

From this idea, Whorf made a new and stronger thesis. Whorf argued that our native languages are the ground stone of our dissection of nature. We are to a large extent part of an agreement that organizes nature in a particular way, i.e., an agreement made from our language dependent communities which is systemised in our language (Whorf, Carroll, & Chase, 1956, p. 213). According to Whorf, language does not only influence our thoughts. It is the defining framework. The linguistic categories of language set limits and determine our cognitive categories and, thereby, how we perceive the world.

The theses of Sapir and Whorf have been gathered in what is popularly called the *Sapir-Whorf hypothesis*. This hypothesis is often defined to include two versions. The weaker version, *linguistic relativity*, states that language affects cognition as opposed to the more radical version, *linguistic determinism*, which states that language determines our thoughts and that we are only able to think what our language has words for (Silverstein, 2015, p. 96).

According to the Sapir-Whorf hypothesis, language affects thought. Thought

is an important element in attention, and attention is needed in order to store memories (Anderson, 2015, pp. 55-56, p. 179). Therefore, language is an important aspect of memory.

Memory and false memories in the brain

(TRS) Nearly all of the brain is involved in memory. However, the frontal structures of the brain are most critical when encoding new memories and when retrieving memories from long-term memory. Though, when storing memories permanently, the temporal structures play the dependent role. Especially, the hippocampus contributes majorly in this process (Anderson, 2015, pp. 124-181).

Hippocampus also plays the critical role in the phenomenon false memories. To examine the neural grounds of false memories, the *Deese-Roediger-McDermott paradigm* was developed (Anderson, 2015, pp. 167-69). In this research, participants studied lists consisting of closely associated words. Afterwards, participants studied another word list consisting of either the studied words, associated words or unrelated words. The results showed that participants succeeded in recognising the studied words and rejecting the unrelated words. However, almost 80% of the associated words were falsely accepted as studied words (Cabeza, Rao, Wagner, Mayer, and Schacter, 2001, as cited in Anderson, 2015, p. 168).

Cabeza et al. (2001, as cited in Anderson, 2015, p. 168) also investigated the activation in the cortex when participants produced the words of the different categories. They found that the fMRI responses produced by the studied words and the associated words were nearly indistinguishable in the hippocampus and much more

forceful than the response produced by the unrelated words. These hemodynamic responses correspond to a large degree with the behaviour of the participants in the experiment examining the Deese-Roediger-McDermott paradigm, where people had difficulties in discriminating memories of studied words from associated words. As the hippocampus responds equally high to actual memories and false memories, it is unsuccessful in distinguishing between what is genuine experiences and what is not (Anderson, 2015, p. 168).

False memories and their implications

(TRS) A theory that explains the false memory phenomenon is the *fuzzy-trace theory* (FTT) (Brainerd & Reyna, 2002). The bottom line of this theory is that memory mechanisms work opposingly (Reyna & Brainerd, 2005). The theory extends on the *dual-process theory* (Evans, 2007, Stanovich, 2011, as cited in Anderson, 2015), which concerns the processing of mental representations of past events through two pathways; either an automatic, unconscious process or a deliberate, controlled process. The dual-process theory states that these systems work independently of each other (Anderson, 2015, pp. 257-258). The FTT, on the other hand, suggests that the two pathways, when storing and retrieving memories, work through gist and verbatim traces parallelly. It is built on five principles of which four of them are especially relevant to this study. These are formulated by Brainerd and Reyna (2002):

The first principle is *Parallel Storage of Verbatim and Gist traces*, which suggests that the elements we remember leaves traces on surface levels and meaning levels in our memory. *Verbatim traces* capture memories in their exact and specific form.

These representations are memorised accurately and detailed based on their surface forms. On the contrary, *gist traces* represent the general meaning of elements and elaborate on their relations to the world. These representations are intuitive and rely on the comprehension of the event. Unconsciously and automatically, this connects meaning across topics. When something is stored in memory, it is processed in both verbatim and gist traces parallelly, i.e., both the surface and context are included.

The second principle formulated by Brainerd and Reyna (2002) states that our memory performance is built on how verbatim and gist traces are retrieved. Directional distortions of a retrieval cue, like interviews or instructions on how to retrieve something, can mislead the retrieval. The exact combination of verbatim and gist traces establishes the degree of false memories. It is worth noticing that in regard to forgetting, verbatim traces become unreachable faster than gist traces.

The third principle proposes that in combination, verbatim and gist traces can increase the accuracy of memories when their content correspond. Despite this, they affect false memories opposingly. Whereas verbatim traces restrain false memories by overcoming false generalisations, the retrieval of gist traces can enhance false memories. This can occur if the memory is retrieved in the wrong context or biased by earlier experience, and hence attributes to incorrect generalisations. This could explain the findings of Cabeza et al. (2001) by the participants relying on gist traces when retrieving their memories of the studied word lists.

The fifth principle covers the fact that a vivid memory exists when relying on

verbatim traces and usually also when relying on gist traces. However, when relying only on biased and strong gist traces, a vague memory is retrieved. This can cause so called phantom memories of things that were not actually experienced.

A developmental reversal exists in which adults tend to have more inaccurate memories than children, i.e. adults rely more heavily on gist traces and with this reliance make more errors. Thus, false memories arise partly from the reliance on gist traces.

In everyday life, it is practical to rely on gist traces and so called *top-down processing*, where context or general knowledge of the world guides our perception. This saves us from having to remember every single detail we run upon (Anderson, 2015, p. 47).

Flashbulb memories

(KK) Many of the experiments involving false memories are manipulations or distortions of memories that the participants were more or less indifferent to (e.g. Loftus & Palmer, 1974). In this regard, an interesting phenomenon is the concept, *flashbulb memories*. The term, flashbulb memories, was first introduced by Brown and Kulik (1977).

Flashbulb memories are characterised as memories of events of high importance to the individual experiencing them, which causes the memory to be stored permanently (Anderson, 2015, p. 145-148). In the original study by Brown and Kulik (1977), they investigated American participants' memories of the Kennedy assassination. They found that while this incident created an exceptionally powerful memory, it is not only the nature of the event that has created the strong memory. According to

Brown and Kulik (1977) two factors seem to determine how the memory is encoded: a high level of surprise and a high level of emotional arousal. If these variables are not sufficiently high under the encoding, no flashbulb memory is encoded. If the variables do attain sufficiently high levels, they seem to influence how frequent the memory is rehearsed both covert and overt, which then in turn affects the ability to recall the memory. Thus, Kulik and Brown suggest these factors will create stronger memories.

Opposed to Brown and Kulik (1977), a recent study examined whether the memories of personally relevant and highly stressful events could be distorted by misinformation (Morgan, Southwick, Steffian, Hazlett, & Loftus, 2013). The study was performed on 800 military personnel undergoing Survival School training. This training phase consists of a mock-captivity in a prisoner war camp. During this training, the participants underwent a 30-minute-long interrogation with types of stressors experienced by actual war prisoners such as facial slaps, abdominal punches, etc. After the stressful interrogation, the participants were asked to identify their perpetrator. When they were fed misleading information about the perpetrator, e.g. exposure to a photograph of a different person, more than half of the participants falsely identified their perpetrator (Morgan et al., 2013).

The fact that memories of stressful events are manipulatable is suggested in these findings. Moreover, this modification is even possible in military personnel whose training and experience are thought to result in relative immunity to such impacts.

Priming and spreading activation

(SKB) To examine how people's memories and decisions are manipulatable, experimenters have used *priming*. Priming is the enhancement of a signal due to previous stimuli (Anderson, 2015, p. 370).

Several theories have been suggested to account for the mechanisms of priming. One of these is known as the *spreading activation theory*.

Anderson (2015, p. 135) presents spreading activation as a process which makes associated memories to currently attended items more accessible. McKoon and Ratcliff (1992) presents it as a theory which suggests words to be interconnected with different strengths. For instance, if one hears the word *pool*, one might think of a swimming pool. However, if one hears the words *pool* and *game*, one might think of a pool table. This shows how words are connected and how different information results in different associations. Yet, these word connections may be different amongst individuals (Anderson, 2015, pp. 136-137). According to the spreading activation theory, it is possible to prime someone because the information provided in the priming assists the retrieval of a specific reply (McKoon & Ratcliff, 1992).

Loftus and Palmer

(SKB) Loftus and Palmer (1974) investigated the accuracy of memories. The motivation for the study was the belief that there must be some factors which are particularly powerful in influencing memories of an event.

The report contained two experiments. The first experiment showed seven films of car accidents after which participants were asked to give a description of the

accident and then answer a series of questions. The critical question existed in several forms by using different verbs: "About how fast were the cars going when they *smashed* into each other?". The verbs used were *contacted*, *hit*, *bumped*, *collided*, and *smashed*, respectively. Hereby, they strived to prime the subsequent answer from the participant in a particular way. The mean speed estimates for the different verb conditions showed systematic increase from *contacted* till *smashed*. These results suggest that the choice of a single word (in this case a verb) can change estimates of speed in a systematic way.

Loftus and Palmer proposed two possible explanations for their results. One explanation was *response-bias factors*, which influence one's estimate in a particular way; If one is uncertain whether the car drove 40 or 50 km/h, the phrasing of the subsequent question may bias one's response to choose the higher estimate if primed with the verb *smashed*.

The other possible explanation proposed was that the phrasing of the question causes actual changes in the participant's representation of the incident. If this is true, the verb *smashed* may result in the subject remembering the accident as more dramatic. Thereby, it could result in the subject remembering the incident differently than it truly unfolded.

In the second experiment, students were shown one film in which a multiple car accident took place. First, students described the accident with their own words. Afterwards, a series of questions about the accident was asked in which the critical one concerned the speed of the car. Three conditions were present. One using the verb *smashed*, another using the verb *hit*, and in the third condition, no question was asked.

After one week, the participants were asked a series of questions about the accident without seeing the video again. The critical question was “*Did you see any broken glass?*”. There was no broken glass in the accident. Since broken glass is more associated with accidents occurring at high speed, the experimenters expected participants in the *smash* condition to answer “Yes” more often to this question.

Results showed a significant difference in estimates of speed in the two conditions with the leading question. Moreover, Loftus and Palmer’s results showed that the probability of answering “yes” to the question concerning broken glass was approximately twice as high in the *smash* condition as in the *hit* condition. Thus, the verb *smashed* led to higher speed estimates as well as more “yes” responses.

Loftus and Palmer suggested that the phrasing of a question can influence the answer given considerably. They proposed that two kinds of information are encoded in our memory for the incidents we experience; the first kind is the original incident itself, and the second kind is external information provided after the incident. This post-event information interferes with our memory of the original experience in a way that results in us being unable to disassemble these two kinds of information. According to Loftus and Palmer, we experience having one inseparable memory.

Moreover, they proposed that their results suggest that the participants first form a representation of the incident, which is later influenced by post-event information (questions phrased with either *hit* or *smashed*). This interference with the original representation influences participants to be more biased in their speed estimates and

more prone to answering “Yes” to seeing broken glass at the scene.

The experimenters concluded that verbal labels, such as *smash*, altered the way participants represented memories of the accidents. This alteration pointed towards the representation proposed by the verbal label.

Loftus and Zanni

(KK) In 1975, Loftus and Zanni set out to study how language has the ability to alter memories, closely related to Loftus and Palmer (1974).

The main subject investigated was whether human perception, recollection and verbalisation of previous events really are true copies of reality, or if phrasings of questions post-event are able to distort memories.

In line with Loftus and Palmer (1974), a video of a multiple car accident was shown to the participants. Different from Loftus and Palmer, they examined whether questions with a definite article would be more likely to be affirmed e.g. “Did you see *the* broken headlight?” versus “Did you see *a* broken headlight?”.

Loftus and Zanni found that when using an indefinite article, it resulted in false “Yes” responses 7% of the time, but with a definite article the false “Yes” occurred 15% of the time. This suggests that when posing a question with a definite article, it is more than twice as likely that a false memory will occur. Furthermore, they found that a definite article would be less likely to generate the answer “I don’t know” (Loftus & Zanni, 1975).

Human associative memory (HAM)

(KK) A model relevant to Loftus and Zanni’s results is the *human associative*

theory (HAM). HAM is a computational model, that makes predictions about linguistic assumptions in human cognition, i.e. sentence memory and other verbal learning phenomena (Trayhurn et al., 2005). When a speaker is already familiar with a particular item a definite article is used. HAM simulates this and looks for the referent of that term in its memory. Though, when a noun with an indefinite article is presented the model assumes that it is a novel term that has been introduced to its memory (Loftus & Zanni, 1975) in line with what a human would do.

This study

(SKB, KK) By now, different studies and arguments have been accounted for which propose memories as being manipulatable. However, many have tried to replicate the results of Loftus and Palmer and have not succeeded (e.g. Read, Barnsley, Ankers & Whishaw, 1978, as cited in Goldschmied, Sheptock, Kim, Galily, 2017). The motivation of our study is the ambiguity of results on manipulatable memories.

Our study is based on the experiments done by Loftus and Palmer (1974) and by Loftus and Zanni (1975). However, this experiment focuses on the interaction effect between the phrasing of a post-event question and the use of an indefinite or a definite article. We conduct this study in Danish to make it equivalent to earlier research by investigating the native tongue of the participants.

We hypothesise the following:

H1: When misled by post-event information, people will tend to estimate speed as higher when an extreme verb (*smash*) is used than when a neutral verb (*hit*) is used.

H2: When misled by the use of definite articles (*the*) and indefinite articles (*a*), people will tend to remember something falsely more frequently when primed with a definite article.

H3: When misled by the extreme verb (*smash*) and being misled by a definite article (*the*), people will be more prone to falsely remember a broken windscreen.

Materials and methods

In the following section, there will be accounted for the used materials and methods.

Participants

(TRS) Mainly, participants were collected randomly from Aarhus University and from our acquaintances.

The survey was run in Danish to make sure the only variable tested was memory, and not the use of second languages, and to get a larger number of participants. Thus, only Danes were included in the experiment.

People that to a large degree had prior knowledge to the study by Loftus and Palmer (1974) were excluded.

The total number of participants was 63 distributed as 39 females and 24 males. The large proportion of university students affects the mean age of participants, that was 24.8 years ranging from 20 to 56 years (SD = 9.26).

People under 18 years were excluded from the experiment due to lack of experience with driving a car, since we suspect this could be a source of errors in speed estimates.

Materials and stimuli

(SKB) The study was divided into two experiments both conducted as surveys on *Kwiksurveys.com*.

The first part of the experiment included two conditions. Here participants were shown a video of a car crash. The duration of the video was 12 seconds and there was no sound in the video. It was edited in *Windows Movie Maker*, where a red circle was added around the car that participants should pay attention to.

The video shows a light-coloured car driving towards the viewer on a city expressway. The driver of a black car makes a left-hand-turn from an adjacent road and fails to see the oncoming light-coloured car. This results in the light-coloured car colliding frontally into the black car. (link to the video, see appendix (a))
After watching the video, participants were asked one of two questions:

Hit condition:

“Hvor hurtigt kørte den lyse bil i den røde ring, da den kørte ind i den sorte (km/t)?” (Eng: *How fast did the light-coloured car in the red circle drive when it hit the black car (km/h)?*).

Smash condition:

“Hvor hurtigt kørte den lyse bil i den røde ring, da den smadrede ind i den sorte (km/t)?” (Eng: *How fast did the light-coloured car in the red circle drive when it smashed into the black car (km/h)?*).

In the first part of the experiment, the measurement was the participants' speed estimates. Thus, the independent variable was

the verb used (*hit* or *smashed*), and the dependent variable was the continuous estimate of speed.

In the second part of the experiment, a follow-up survey was sent to the participants containing four questions including one critical question (see appendix (b)). One of the two versions of the critical question was asked:

Definite article condition:

“Så du den smadrede forrude?” (Eng: *Did you see the broken windscreen?*)

Indefinite article condition:

“Så du en smadret forrude?” (Eng: *Did you see a broken windscreen?*)

Thus, the only difference between these conditions is either a definite article or an indefinite article.

A “yes” or “no” response was measured. Thus, the independent variable was the article (definite or indefinite), and the dependent variable was the categorical answer “yes” or “no”. It is important to notice that no broken windscreen was present in the video.

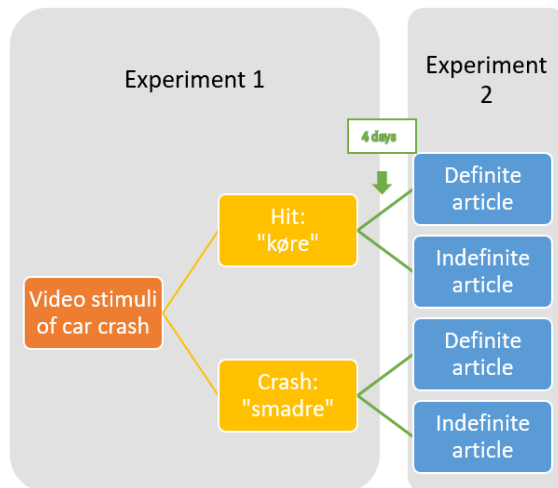


Figure 1: The design of the two experiments conducted in this study

Procedure

(SKB) Firstly, all participants were sent a link to their own computer with the first part of the experiment containing the video. Secondly, they were shown a short introduction text in which they were told that they were about to see a video of a car crash and were asked to pay close attention. Thirdly, they were shown the car crash video. Lastly, participants were asked either the question from the *hit* condition or the question from the *smash* condition. The participants were only allowed to watch the video once and, thus, only had one trial.

After four days, the same participants received a link to the follow-up survey. The two conditions from experiment 1, *hit* and *smash*, respectively, were divided into two subgroups of similar sizes receiving either the critical question with a definite or indefinite article.

Participants were randomly assigned to conditions in both experiment 1 and experiment 2.

Analysis

(KK, TRS/SKB) The two experiments in this study investigate each their hypotheses. Experiment 1 investigates *H1*, seeking to determine the difference in the speed estimates under the influence of leading questions. Experiment 2 investigates *H2* examining the effect of definite and indefinite articles in questions, and together the experiments investigate *H3* examining interaction effects.

To test the hypotheses, the data gathered from *Kwiksurveys.com* was imported into *Rstudio* (RStudio team, 2016) and pre-processed to enable further analysis. Incomplete surveys were excluded from the experiment.

In order to investigate *H1*, the mean score of speed estimates for each condition was compared using a robust t-test (package *WRS2*, Mair & Wilcox, 2017), since the assumption of normality was violated.

To investigate *H2*, simple logistic regression with a binary outcome was performed. Assumptions were examined and met.

$P(\text{falsely remembering a broken wind-screen}) \sim \text{condition}(\text{exp2})$

To investigate *H3*, a 2x2 factorial design of a logistic regression model was performed. Assumptions were examined and met. Two models were constructed to facilitate a comparison of these in order to find the best explanatory model: The first model examined main effects with the conditions of both experiments as predictors. The second model examined interaction effects.

Model 1:

$P(\text{falsely remembering a broken windscreen}) \sim \text{condition}(\text{exp1}) + \text{condition}(\text{exp2})$

Model 2:

$P(\text{falsely remembering a broken windscreen}) \sim \text{condition}(\text{exp1}) \times \text{condition}(\text{exp2})$

By multiplying the predictors, as in *Model 2*, interaction effects were examined as shown in *Table 1*.

smash (“smadre”) + indefinite article	Hit (“køre”) + indefinite article
smash (“smadre”) + definite article	Hit (“køre”) + definite article

Table 1: 2x2 factorial design to examine interaction effects

Both models fitted the data poorly. A further investigation was performed in order to determine, if the models leaved more variance unexplained than the null model. However, the null model was not significantly better than the informed model.

Henceforth, this study will report the best explanatory model, i.e. model 2, which answers *H3* most accurately of the two models.

Results

(KK, SKB/TRS) The distribution of the speed estimates is shown in *Figure 2*.

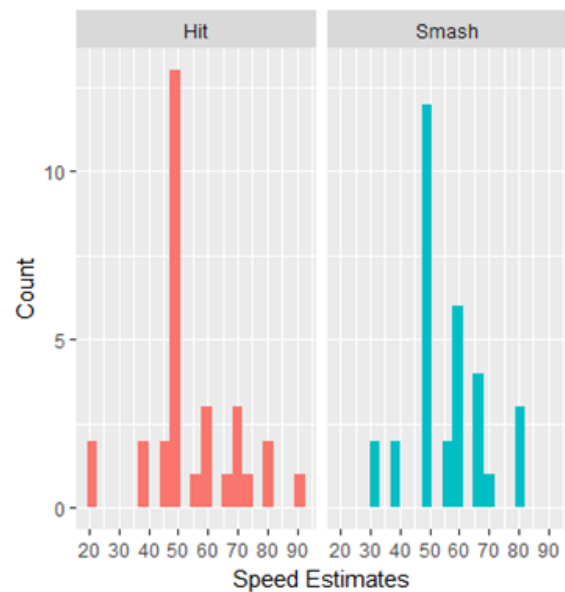


Figure 2: Distribution of speed estimates in experiment 1

Investigating *H1*, a robust t-test showed no significant results, $t(33.52) = 0.42$, $p = .678$, $r = .07$. The results are plotted in *Figure 3*.

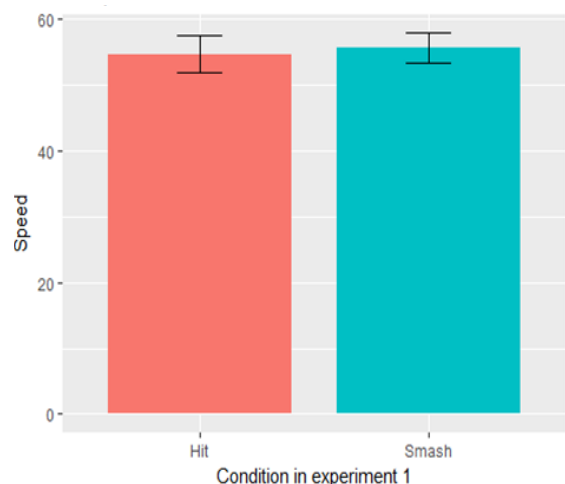


Figure 3: Independent robust t-test conducted on speed estimates in experiment 1

The results of the logistic regression investigating *H2* showed that the use of definite and indefinite articles had no significant impact on falsely remembering a broken windscreen, $b = 0.03$ ($SE = 1.44$), $z = 0.02$, $p > .98$. The results are visualised in *Figure 4*.

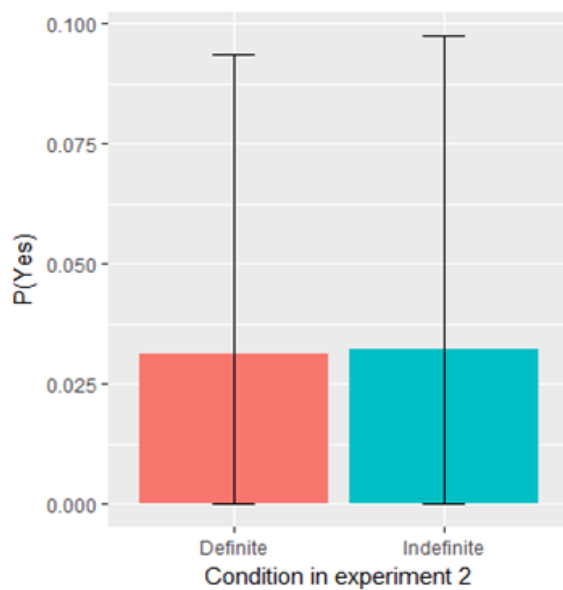


Figure 4: The results of the logistic regression investigating *H2*

The results of the 2x2 factorial design of the logistic regression model investigating *H3* showed no significant interaction effect between the verb in the first experiment and the use of definite or indefinite articles in the second experiment, $b = -0.13$ ($SE = 6372.25$), $z = 0.00$, $p > .99$. Results are shown in Figure 5.

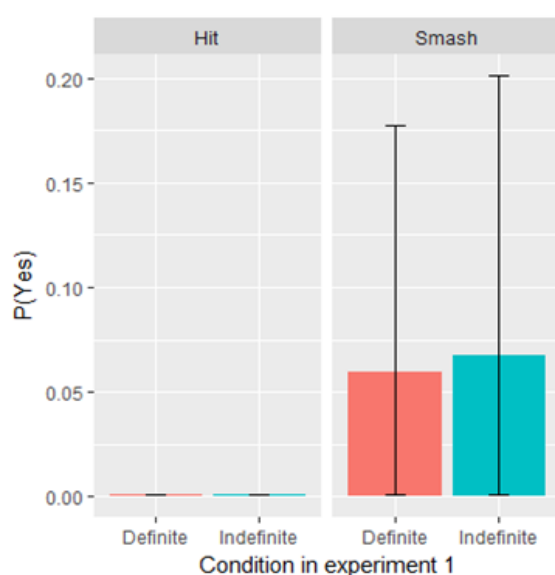


Figure 5: Interaction effect between conditions from both experiments

Discussion

(KK) The results from the t-test investigating *H1* suggest that there is no significant difference between the use of the verbs *smash* and *hit* when participants estimated speed. This is visualised in Figure 3 by overlapping error bars. The effect size equally indicates no effect, however, relatively close to a small effect. According to these results, participants will not be significantly misled by this particular kind of post-event information. On average, participants in the *smash* condition estimated speed to be higher than participants in the *hit* condition, though not significantly higher. Thus, the null-hypothesis that relates to *H1* cannot be rejected.

The results regarding *H2* suggest that the different phrasing with definite and indefinite articles, respectively, does not significantly mislead the participants to more frequently remembering an object which was not present (in this study, a broken windscreen). Therefore, it is not possible to reject the null-hypothesis that relates to *H2*. The insignificant results are visualised in Figure 4 by overlapping error bars.

Likewise, in regard to *H3*, the null-hypothesis cannot be rejected. The results of the 2x2 factorial design of a logistic regression model suggest that there is no interaction effect between conditions in experiment 1 and experiment 2. The insignificant results are visualised in Figure 5 by overlapping error bars. Even though the results are insignificant, a falsely remembered windscreen is only observed in the *smash* condition, which seems as though the condition in experiment 1 could have an impact while the condition in experiment 2 has not. However, further investigation is

needed in order to examine if this is an actual generalisable effect.

Potentially, the insignificant results could have many sources some of which is discussed in the following sections.

Limitations of this study

(TRS, SKB) The number of participants might be critical to the results. If having collected more participants, the data might have approached normality in experiment 1.

Optimally, a higher rate of randomness in the collection of participants would have been preferable. Since the participants to a great extent were collected from acquaintances, there is a higher risk of the participants having discussed the video in the waiting period and, hence, influenced each other's memory of the car accident. If participants had been explicitly told not to discuss the experiment with one another during the waiting period, the accuracy of their memories might have been different.

Furthermore, the representativeness of the population investigated can be questioned, due to the fact that participants were mainly university students.

Since the participants of the study were highly dominated by university students, whom we to a large extent assume to have driver's licenses, we infer that they have some driving experience. However, we did not ask participants to clarify whether they had a driver's license, which would have been preferable to know.

The distribution of speed estimates (Figure 2) displays that a remarkably large portion of the participants estimated the speed of the car to be 50 km/h. Since the speed limit in Danish cities is 50 km/h, this could suggest that participants have been

biased in their estimation due to the fact that the car crash took place in the city. Since the length of the video was relatively short, the participants might draw on their general knowledge of speed limits in cities.

(TRS, KK) The duration of the video was 12 seconds, which is a relatively short time to adjust one's mind to a different setting. Even though the participants were prepared for the video of the car accident with a short introduction text in the survey, the effect of surprise might not have been eliminated entirely. A longer video with audio could have increased the ecological validity of the study, because the participants to a higher degree would have adjusted to the setting of the video. By including audio more senses would be included. 12 seconds might not be enough to create the grounds for somewhat natural stimuli (Goldschmied et al., 2017). Thus, the length of the video could result in lack of context.

This could have affected the results investigating *H1*, since the participants had a very short amount of time to estimate the speed. The duration might also have affected the results investigating *H2*. Participants had a short amount of time to observe the car crash and memorise the details.

The red circle around the car was added in order to minimise the effect of the short time to orientate in the video. This, however, confused some of the participants, who thought it was a "trick" in line with the change blindness experiments (e.g. McConkie & Currie, 1996, as cited in Anderson, 2015).

(SKB, TRS) In the following paragraph the phrasing of the survey questions is revised.

In experiment 1, the *hit* condition contained the Danish word "køre" twice.

Firstly, in the meaning of *drove* and, secondly, in the meaning of *hit*. In the meaning of *hit*, “køre” is a predominantly neutral word. In retrospect, it would have been interesting to investigate another leading verb priming the participants to think of the accident as being less severe, e.g. “støde” or “ramme” (which both translates as *hit*, but are more tranquil in Danish). This would create a larger spectrum of experimental manipulation.

In addition, the phrase “*hvor hurtigt ...*” (Eng: *how fast ...*) was used in both conditions of experiment 1 in line with Loftus and Palmer (1974). This wording has an implication of the car driving fast as opposed to slow. Instead, a wording similar to “*What do you estimate the speed of the car to be ...*” could have been used. Nevertheless, the same wording was used in both the *hit* and *smash* condition which, in theory, should result in the same effect on both groups. However, a neutral wording probably would have been more ideal.

Other research (e.g. Lipscomb, Bregman & McCallister, 1985, as cited in Goldschmied et al., 2017) has examined the effect of using the different phrases “*how fast was the car going?*” or “*how slow was the car going?*” across ages. Their results showed higher speed estimates when the first phrase was used. Moreover, they investigated the same question phrasings as in the original Loftus and Palmer study (1974) across five age groups. They found that only for seventh- to eighth-graders, *smashed* resulted in higher speed estimates than *hit*. Due to the effect only being observed in this age group, the experimenters questioned the generality of this effect.

Since the mean age of this study was approximately 25 years and no effect of the manipulation was found, the general effect

Loftus and Palmer found could be questioned in line with what Lipscomb et al. (1985, as cited in Goldschmied et al., 2017)) suggested.

In experiment 2, *H2* was investigated and no significant results were found. Four questions were asked in order to hide the critical question. Ideally, more questions could have been added in order to conceal the critical question better. Furthermore, the questions were shown in a fixed order but could ideally have been randomised.

The critical question in experiment 2 was a yes-no task. Consequently, it was not possible to answer, “I do not know”. This results in a forced-choice paradigm in which participants must commit to an answer. An advantage of this is that more responses can be analysed. However, by forcing participants to commit to either seeing or not seeing a broken windscreen, the participants may have answered one of the two options without being certain of what they remember. As a consequence, the data from experiment 2 may be inaccurate.

To further investigate the effect of indefinite and definite articles, repeated measures could be conducted. In such a scenario, each participant would receive several questions containing definite and indefinite articles. This procedure could lead to larger accuracy and eliminate some individual baseline variance by informing the model of these mixed effects.

Due to the unclear phrasing, we cannot be absolutely sure about what specific words had an effect on the participants’ answers. A reconsideration of the wording in the survey could have ensured which factors were actually measured.

(TRS, KK/SKB) In hindsight, a pilot study could have prevented the study from a number of inaccuracies. For instance, in experiment 2, a ceiling effect was observed. That is, nearly all participants answered correctly to the critical question. A broken windscreen is more easily noticed than broken glass, which was the element Loftus and Palmer (1974) investigated. The choice of a broken windscreen as element may have contributed to the observed ceiling effect. With a pilot study, we could have discovered this and attempted to implement an adjustment.

Many of the inaccuracies of this study could also have been factors to improve in the original studies by Loftus and Palmer (1974) and Loftus and Zanni (1975).

Challenges concerning Loftus and Zanni (KK) In the study by Loftus and Zanni (1975), they argued that HAM supports their hypothesis, that participants primed with a definite article will be more prone to falsely remembering a broken headlight. According to Loftus and Zanni, this is due to the fact that participants are partly given the answer when a question is asked with a definite article.

When asking: “Did you see a broken headlight?”, two questions are implicitly asked: “Was there any broken headlight?” and “Did you happen to see the broken headlight?”. Thus, when asking a question with an indefinite article, the participant has to answer both whether there was in fact any broken headlight present and if the participant had happened to see it. When having answered the first part of the question, he should be fairly certain of the second part of the question.

When asking a question with a definite article, the first question is implicitly answered by the narrator. Thereby, the participant only has to answer whether he had happened to see the broken headlight, and, as in the case with the question with the indefinite article, the participant can be fairly certain of his response (Loftus & Zanni, 1975).

However, the results of experiment 2 suggest that it does not happen to be as Loftus and Zanni hypothesised and could suggest a different approach to HAM. No significant results were found when investigating *H2*. According to HAM, when people are presented with a definite article, they search for a reference to a previous introduced term. However, in this study a memory of a broken windscreen could not have been encoded. Thus, when participants unsuccessfully tried to recall it from memory, they falsified the question and answered *no*, as if they were presented with an indefinite article. This suggests that there is no particular difference between *a* and *the*, as earlier found in studies by Burt (1931, as cited in Loftus & Zanni, 1975).

A weakness Loftus and Zanni pointed out in their study was that “a broken headlight” could refer to any headlight in the video. Since it is impossible to attend to all headlights in the video, the participants had no chance of being sure about whether there in fact was a broken headlight. In their study, this would predict a large number of the answer: “I do not know”. In this study, this could give rise to a large inaccuracy in the results using indefinite articles compared to using definite articles. When using definite articles, the question more strongly implies the presence of a broken headlight. Thereby, participants only have to identify one broken headlight and do not have to

deal with the uncertainty of if it even existed. However, the results of this study do not imply this difference.

Loftus and Zanni found that the use of definite articles leads the answer given to a particular question. This is not found to be significant in our study. An explanatory factor to the difference between the studies could be repeated measures in Loftus and Zanni's follow-up survey. The participants were asked multiple questions with half of the objects present and the other half absent. Potentially, this results in a larger accuracy in the measurements, showing the extent of the effect of indefinite and definite articles. This method could be interesting to implement in further investigations.

Challenges concerning Loftus and Palmer

(SKB) In Loftus and Palmer's study (1974), four of the seven videos were staged. They were segments from longer driver's education films. For this reason, the experimenters knew how fast the cars were actually going. By comparing the actual speeds with the participants' speed estimate, they concluded that, in general, participants were bad at estimating speed. This notion is relevant to *H1*, as *Figure 2* shows that participants' estimates, overall, varied from 20-90 km/h.

Read and Bruce (1984, as cited in Goldschmied et al., 2017) found that when participants were more familiar with the environmental framework, the effects of leading questions decreased. Since one can get a driver's license from the age of 18 in Denmark and the mean age in this study was approximately 25 years, their results suggest that these perhaps less experienced drivers in this study should be easier to manipulate. However, no significant difference between

conditions were found suggesting no effect of the manipulation even in less experienced drivers.

Since our study and others (e.g. Read, Barnsley, Ankers & Whishaw, 1978, as cited in Goldschmied et al., 2017) have not been able to replicate the findings of Loftus and Palmer (1974), it is relevant to look at their study with scepticism. A critique by Goldschmied et al. (2017) is the ecological validity of the study. In the first part of the experiment, participants watched seven videos of car crashes. This is not equivalent to a one-time exposure to an event in real life and, therefore, is not directly applicable to real-life eyewitness testimonies. In our study and in the second part of Loftus and Palmer's study, only one video was shown which makes it more related to a real-life experience even though the earlier mentioned limitations of one short video are still relevant.

Another relevant factor involving real-life experiences, is the fact that this study is in line with Loftus and Palmer's asked participants to recall an event they had witnessed. This is similar to how eyewitness testimonies are put to use by the police and the legal system, making it applicable outside laboratory settings.

A study by Yuille and Cutshall (1986) suggests that misleading information might have a greater effect in laboratory settings than when one witnesses a true traumatic incident. They interviewed witnesses to a real armed robbery and found that they could not alter the witnesses' memories through misleading information. A possible explanation of this finding could be that the witnesses encoded flashbulb memories of the incident, which according to Brown &

Kulik (1977, as cited in Anderson, 2015, p. 145-146) are stored permanently.

For further research, one could in line with Yuille and Cutshall (1986) conduct an experiment trying to influence actual car crash witnesses with leading questions.

Brown and Kulik (1977) suggested that factors such as a high level of surprise and a high level of emotional arousal are important if one is to encode a flashbulb memory. However, watching a video does not make as large an impact as witnessing a car crash or another traumatic incident in real life. Any level of surprise by watching the video was further diminished by informing the participants that they were about to see a car crash, leaving minimal chances for the participants to encode a flashbulb memory. Therefore, it would be interesting to interview witnesses of actual car crashes to examine if the effects found by Loftus and Palmer (1974) and Loftus and Zanni (1975) would be observed or, like in this study, no effect would be observed.

Priming and spreading activation

(SKB) In our study, we attempted to prime participants in both experiments. However, the priming in both experiments turned out insignificant, suggesting the amount of priming in this study was inadequate. Nevertheless, the effect of priming has been found in other experiments in similar contexts and in other contexts (e.g. Loftus & Palmer (1974), and Kaplan (1989, as cited in Anderson, 2015, p. 136).

The spreading activation theory becomes relevant to our study given the fact that this theory states that currently attended items makes associated memories

become more available. Even though the results of this study did not show a significant difference in speed estimates between the two conditions, only participants in the *smash* condition answered “Yes” to seeing a broken windscreen: One in the *indefinite article* condition and one in the *definite article* condition (see *Figure 5*). In relation to the spreading activation theory, the verb *smash* may be more strongly linked with severe accidents than the verb *hit*. By hearing the word *smashed* participants might associate the accident more closely with a severe accident and, thereby, a broken windscreen. Maybe more so, than when the verb *hit* was used. However, since the results of this study were not significant further research must be conducted to clarify this implication.

(TRS, KK/SKB) Since there were no measurements of the brain activity in the participants, this study has minimal grounds for discussing the activation in the brain during the study. In the Deese-Roediger-McDermott paradigm, participants had to recognize studied words (Anderson, 2015, pp. 167-169). In our experiment, participants had to rely on their recollection of the car accident. The paradigm cannot directly account for the results in experiment 2, since it is recollection and not recognition that was being tested. However, the paradigm proposes that people rely on associations and generalisations of their memories. The results of our study indicate that this is not the case, since participants were able to answer accurately.

It would be interesting to examine recognition by showing the participants a video again four days later, the same video with minor changes, to see if they noticed the differences.

However, in this study participants were tested solely on the retrieval of the memory of the car crash. The FTT can explain some of the critical factors in this retrieval process.

(TRS) The first principle in FTT (Brainerd & Reyna, 2002) accounts for the parallel storage of verbatim and gist traces; what can be retrieved depends on the specificity of the gist traces. Many factors could potentially influence the specificity of a gist trace, since participants could be individually affected by the stimuli. In relation to our experiment, this would mean that if the video of the car accident is stored with a high specificity of the particular context, this will be recalled differently than if it was stored with less specificity.

The second principle concerns forgetting and argues that verbatim traces decay faster than gist traces. This is important in relation to the waiting period from completing experiment 1 until participants had to recall it four days later in experiment 2. Thus, an explanatory factor in the high accuracy of the critical question in experiment 2 might be the length of the waiting period. In the original study by Loftus & Palmer (1974), the waiting period was seven days. Four days might not have been enough time for participants' verbatim traces to become unavailable. If the waiting period had been longer e.g. one or two weeks, the participants might have relied on gist traces to a greater extent. If the memory of the car accident had become a more generalised memory, there might have been a larger effect of the used articles. The short waiting period might have resulted in the memory still being somewhat vivid and contributing to the earlier mentioned ceiling effect.

According to the third principle, the high accuracy could also be explained by the combination of gist and verbatim traces. This combination makes for a high accuracy of a memory, if their content corresponds. This seems to apply for the results of our study.

According to FTT, the results of the investigation of *H1* and *H2* suggest that the verbatim traces constrained false generalisations of the car accident; the participants still had a rather vivid memory of the car accident four days later. It appears that participants have generally managed to rely highly on their verbatim traces in this study. As suggested in the fifth principle this will create a stronger memory. This could be what the results of experiment 2 indicates.

Linguistic challenges

(KK, SKB) According to the Sapir-Whorf hypothesis, language is a leading factor in cognition including memory. While most linguists agree on the weaker version of the Sapir-Whorf hypothesis, the stronger version has received critique (e.g. Harshey, Czerniakowski, Robinson, & Steimel, n.d.). The results of this study suggest that the conducted experimental manipulation does not determine thought. Thus, some critique of the Sapir-Whorf hypothesis might be relevant to consider.

Especially, Whorf has received critique on the idea of language as the defining framework of thoughts. The idea of language and thought as being inseparable is criticised by Pinker (2003). According to Pinker, the work of Whorf and his founding base for linguistic determinism should be questioned to a large degree. Equally, the results of this study questions linguistic determinism, since the wording of the ques-

tions asked does not seem to define the participants' memories. However, one could hypothesise, that if this study did not have a ceiling effect, one might have seen linguistic relativity supported by results suggesting that wording of a question influences memory. The results of experiment 1 showed an indication of a minor differentiation between conditions, however, due to insignificant results further investigation is needed to support this.

In this study, we were unable to replicate the results of Loftus and Palmer (1974) and Loftus and Zanni (1975). A consideration for further investigation could be the fact that we did the experiment in another native tongue than the original study. Therefore, it could be interesting to investigate how distinct Danish and English are to see if "*the categories we dissect nature by*" (Whorf, Carroll, & Chase, 1956, p. 213) indeed are different. Thereby, we would be able to investigate if these differences could clarify the different results when the experiment was conducted in Danish and in English.

Bartlett concluded from his *War of Ghosts* (1932) that memory is reconstructive, and we encode and retrieve information according to our expectations. However, in this study we are not able to display this distortion as hypothesised. Some of the critique Bartlett has received is the fact that he did not have a control group and neither standardised instructions nor standardised time measures, which could question his results (Crane, n.d.). In order to compare our study closely with that of Bartlett, a further investigation is necessary. Equally, it could be interesting to investigate which cultural schemes the participants in this study had in

order to compare it with the possible subsequent distortion.

False memories

(KK/TRS) Relating false memories to reality, it is relevant to investigate the degree in which we can rely on memory in relation to eyewitnesses. To rightfully judge which role eyewitnesses should play in our legal system, it is essential to know the effect of leading questions. Loftus and Palmer (1974) and Loftus and Zanni (1975) suggested that wording affects memory. However, the results of our study suggest no effect of leading questions. Therefore, further investigation is needed in order to clarify the effect. Ambiguity in the research on the matter forces alertness when operating with testimonies. Even though attention is raised to this topic and rules regarding leading questions has been tightened, it is still a potential issue. It is possible to object in court against leading questions, but it is critical to notice that when a leading question has been asked, it has affected the witness even though it is sustained and rephrased (Simpson & Selden, 1998). With larger knowledge on the subject, we can navigate around this issue more deliberately in order to approach justice.

Conclusion

(SKB) In conclusion, the results of this study suggest that manipulative post-event information in the shape of leading verbs and leading articles does not have a significant impact on neither speed estimates nor false memories (in this study falsely remembering a broken windscreen). However, further investigation is needed to determine whether these results are generalisable.

Acknowledgments

Participants, students of CogSci for solidarity, family and friends for support, Kristian T for guidance, Diana for understanding of

legal matter, and Jakob Lillelund for the understanding of logistic regression and help in Rstudio.

References

- Anderson, J. R. (2015). *Cognitive psychology and its implications* (8th ed.). Macmillan.
- Bartlett, F. C. (1932). *Remembering: An experimental and social study*. Cambridge: Cambridge University.
- Brainerd, C. J., & Reyna, V. F. (2002). Fuzzy-trace theory and false memory. *Current Directions in Psychological Science*, 11(5), 164-169. doi:10.1111/1467-8721.00192
- Brown, R., & Kulik, J. (1977). Flashbulb memories. *Cognition*, 5(1), 73-99.
- Crane, J. (n.d.). *Bartlett (1932)*. Retrieved 15 December 2017, from <https://www.thinkib.net/psychology/page/8195/bartlett-1932>
- Goldschmied, N., Sheptock, M., Kim, K., & Galily, Y. (2017). Appraising Loftus and Palmer (1974) post-event information versus concurrent commentary in the context of sport. *The Quarterly Journal of Experimental Psychology*, 70(11), 2347-2356, DOI: 10.1080/17470218.2016.1237980
- Harshey, C., Czerniakowski, M., Robinson, S. & Steimel, K (n.d.). Ask A linguist FAQ. Retrieved 19 December 2017, from <http://linguistlist.org/ask-ling/sapir.cfm>
- Hauerslev, H. (2014, 04). Vidnebeviset er fyldt med fejl. Advokaten, pp. 8-20. Retrieved 15 December 2017, from <https://bibliotek.dk/da/moreinfo/netarchive/870971-tsart%253A36087358>
- Henderson, P. (1981). Looking back at Titus case. Retrieved 13 December 2017 from <http://old.seattletimes.com/news/local/tituscas/lookingback.html>
- Lindsay, D. S., & Read, J. D. (1994). Psychotherapy and memories of childhood sexual abuse: A cognitive perspective. *Applied Cognitive Psychology*, 8(4), 281-338.
- Loftus, E. F., & Palmer, J. C. (1974). Reconstruction of automobile destruction: An example of the interaction between language and memory. *Journal of Verbal Learning and Verbal Behavior*, 13(5), 585-589.
- Loftus, E. F., & Zanni, G. (1975). Eyewitness testimony: The influence of the wording of a question. *Bulletin of the Psychonomic Society*, 5(1), 86-88.
- Mair, P., & Wilcox, R. (2017). *Package 'WRS2'*

- McKoon, G., & Ratcliff, R. (1992). Spreading activation versus compound cue accounts of priming: Mediated Priming Revisited. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18(6), 1155-1172. doi:10.1037/0278-7393.18.6.1155
- Morgan, C. A., Southwick, S., Steffian, G., Hazlett, G. A., & Loftus, E. F. (2013). Misinformation can influence memory for recently experienced, highly stressful events. *International Journal of Law and Psychiatry*, 36(1), 11-17.
- Pinker, S. (2003). *The language instinct: How the mind creates language*. Penguin UK.
- Reyna, V. F., & Brainerd, C. J. (2005). *The science of false memory*. New York: Oxford University Press. doi:oso/9780195154054.001.0001
- RStudio Team (2016). RStudio: Integrated Development for R. RStudio, (Version 1.0.153) [Computer Software]. Boston, MA. Retrieved September 5, 2017 Available from <http://www.rstudio.com/>
- Sapir, E. (1929). The status of linguistics as a science. *Language*, 5(4), 207-214. doi:10.2307/409588
- Schacter, D. L. (1999). The seven sins of memory. *American Psychologist*, 54(3), 182-203. doi:10.1037/0003-066X.54.3.182
- Silverstein, A. (Ed.). (2015). *Human communication: theoretical explorations* (Vol. 12). Routledge.
- Simpson, F. A., & Selden, D. J. (1998). Objection: Leading Question!. *Texas Bar Journal*, 61, 1123-1135.
- Trayhurn, P., Simons, P. C. M., Versteegh, H. A. J., Jongbloed, A. W., Kemme, P. A., Slump, P., Verschoor, G. J. (2005). Citation classic. *British Journal of Nutrition*, 93(1), 136-152. doi:10.1079/BJN20041301
- Whorf, B. L., Carroll, J. B., & Chase, S. (1956). Language, thought and reality. *Les Etudes Philosophiques*, 11(4), 695.
- Yuille, J. C., & Cutshall, J. L. (1986). A case study of eyewitness memory of a crime. *Journal of Applied Psychology*, 71(2), 291-301. doi:10.1037/0021-9010.71.2.291

Appendix

Appendix a)

Koppelgaard, K. [Kiri Kop]. (2017-12-04). Video Stimuli. Retrieved 8 December 2017, from https://www.youtube.com/watch?time_continue=5&v=EU94NsVZdWI

Appendix b)

Questions asked in experiment 2

- Hvilken farve havde den bil, der blev påkørt? (What colour was the car that was run into?)
 - Hvid (White)
 - Sort (Black)

(The critical question that differed in the two conditions):

- Indefinite article condition: Så du en smadret forrude? (Did you see a broken windscreen?)
- Definite article condition: Så du den smadrede forrude? (Did you see the broken windscreen?)
 - Ja (Yes)
 - Nej (No)
- Hvor skete ulykken? (Where did the accident occur?)
 - I byen (In the city)
 - På landet (In the countryside)
- Kørte nogen over for rødt? (Did anybody run a red light?)
 - Ja (Yes)
 - Nej (No)