Den mørke side af semantisk kontekst [The dark side of semantic context]: Semantic context boosts people's confidence in their ability to comprehend Danish.

Kayla Jordan[1], Niels Peter Nielsen[2], Daniel Bernstein[3], Eryn Newman[4], Rachel Zajac[5], Maryanne Garry[1]

[1] School of Psychology, The University of Waikato, 1 Knighton Road, Hamilton, 3240, New Zealand.

[2] Department of Psychology and Behavioural Sciences – CON AMORE, Aarhus University, Bartholins Allé 9 building 1351, 423, 8000 Aarhus C, Denmark.

[3] Department of Psychology, Kwantlen Polytechnic University, 12666 72 Ave, Surrey, British Columbia, V3W2M8, Canada.

[4] School of Medicine and Psychology, The Australian National University Canberra, ACT, 2600, Australia.

[5] School of Psychology, University of Otago, 362 Leith Street, Dunedin, 9016, New Zealand.

Author Note

Abstract

Semantic context helps people comprehend information, but it can also create illusions of truth and understanding. To what extent does increased semantic context boost people's confidence in their ability to perform highly technical skills, such as understanding a foreign language? We addressed this question across five experiments by showing subjects a video clip of people speaking Danish. Some subjects saw the subtitled video while others saw the unsubtitled version. Then we asked subjects to rate how well they thought they would be able to understand Danish in new situations. We found that people who saw the subtitled videos thought they would understand more of that language in new situations compared to those who saw the unsubtitled videos, even though their actual understanding didn't improve. These findings suggest that relative to situations of lesser semantic context, greater semantic context creates illusions not simply of understanding, but also of greater skill.

General audience summary

When people encounter new information, it can be made easier to understand by an accompanying cocktail of words, gestures, and behaviors. The problem is this same cocktail —called semantic context—can also create the illusion of understanding. Take, for example, foreign films and television series. Subtitles help viewers understand what characters are saying, and what is happening. What's interesting is that viewers attend to subtitles effortlessly and may even lose awareness of the subtitles despite still relying on the subtitled information. But could subtitles create a semantic context that encourages viewers to be more confident they had learned the foreign language even when they had not? To answer this question, we conducted five experiments in which we showed participants a video clip of people speaking Danish—either with or without subtitles—and asked everyone to rate their ability to understand Danish in new situations. Then we asked people to translate Danish audio clips to see if they had learned any Danish. We found those who saw the subtitled video were more confident in their ability to understand Danish in new situations compared to those who saw the unsubtitled clips, even though they were not able to translate any more of the Danish audio clips. These findings suggest that relative to situations of lesser semantic context, greater semantic context can create illusions of one's ability to do something implausible.

Introduction

Our research team, like much of the world, spent most of 2020-2023 at home watching Netflix. And we noticed a curious phenomenon. When we watched foreign television shows or films, we seemed to lose awareness of the presence of subtitles and even came to think we were learning the language. And then, within seconds of turning off the subtitles, we realised we were not learning much at all. How are we to make sense of this illusion? Here we suggest subtitles provide "semantic context"—a kind of conceptual scaffold that facilitates fluent processing of otherwise very difficult to process content. While semantic context can work to facilitate incoming information by connecting it to existing knowledge structures, this increase in comprehension should be limited to the specific content encountered (Bransford & Johnson, 1972). But the psychological literature provides us with several reasons to suspect semantic context might encourage people to misattribute that ease of processing beyond the content they have encountered, to something more grand—that they better understand a foreign language more generally. Across five experiments, we report evidence supporting the idea that subtitles provide semantic context that can rapidly boost English speakers' confidence in their ability to comprehend Danish in novel situations.

Of course, it is widely demonstrated that subtitles can be helpful. When people watch videos, subtitles can improve attention, vocabulary, and written and verbal comprehension (see, for a review, Gernsbacher, 2015). Moreover, evidence from eye-tracking studies suggests people attend to subtitles automatically, with no tradeoff between the processing of the subtitles and the images in the film (Bisson et al., 2014; Perego et al., 2009). One way to think about subtitles is that they provide semantic context. For example, most of us who were glued to Squid Game followed the story because the subtitles provided meaning we could not get from the Korean audio track.

Beyond subtitles, the wider literature suggests semantic context—such as titles, advance organizers, illustrations, or photos—eases comprehension of information in the moment (Ausubel, 1960; Bransford & Johnson, 1972; Mayer & Gallini, 1990). And so, people who know upfront that an

obscure passage is titled "washing clothes" understand and later recall that passage better than people who don't know the title (Bransford & Johnson, 1972). Effects like these are thought to occur because semantic context helps people to fit new information into existing representations or hierarchical networks (Bransford & Johnson, 1972). We see similar patterns in work on narrative structure. For instance, when semantic context clarifies the goals of characters in a story, people better comprehend the story and recall more of it later. In one study, when people read about a farmer and his donkey, learning early on that the farmer was trying to get the donkey into its shed helped reveal the structure of all the subgoals and associated actions in the story and improved recall (Bower, 1978). Those who read the same story without knowing the farmer's goal had worse understanding and recall. This example suggests that semantic context helps people partition incoming information accordingly, to apprehend or construct a structure, and to integrate information together into a coherent narrative. Finally, we see similar effects in the literature on event perception. When people view unfolding events, knowing about an actor's goal helps those viewers to predict what will happen next (see, for a review, Zacks, 2020). For instance, when people know an actor's goal is to reach a certain object, they predict or anticipate the reaching—looking at the target object before the actor's hand even moves (Eisenberg et al. 2018). Taken together, the wider literature suggests semantic context eases comprehension and confers a host of benefits.

Collectively, these literatures all suggest subtitles have a bright side. But there could also be a dark side. Although subtitles ease comprehension of information in the moment, there is evidence to suggest people may draw on these feelings of ease to make inferences about a range of unrelated decisions. After all, we know that people draw on feelings in the moment to inform all kinds of judgments—including liking, familiarity and trust (Alter & Oppenheimer, 2009; Oppenheimer, 2008; Schwarz et al., 2021). In fact, feelings of ease can lead to a range of illusions. For example, when subjects read sentences constructed to help them predict a terminal word (*the stormy seas tossed the…*), they were more likely to falsely remember having seen the word "boat" in an earlier study phase than

when a sentence was not predictive (*he saved up his money and bought a …*; Whittlesea, 1993). There is also evidence that semantic context creates feelings of learning; both learning and the illusion of learning evoke positive feelings thought to influence subsequent judgements (Lakshmanan & Krishnan, 2011). For example, in one study, people evaluated wines more positively when the wine label contained context that taught them a new concept—such as a wine called "Yellow Rick" accompanied by a picture of a yellow haystack (Cardwell et al., 2017). Of course, learning that a rick is a haystack has no bearing on the quality of the wine, but in this experiment, the more confident people were that they understood new words such as "rick," the higher they rated wine quality. Considered together, this work raises the possibility that subtitles could provide semantic context that encourages feelings of familiarity, prior exposure, and even illusions of learning a foreign language.

To the extent these possibilities are true, we might also see that Netflix's subtitled foreign shows push people towards being more confident they understand that foreign language, and might even be able to use it in the future. Several lines of research from the cognitive literature support this possibility. First, increased semantic context encourages people to predict an optimistic future. For example, in a series of experiments, people judged claims that certain commodities would increase in price as truer when they appeared with a semantically-related photo (Newman et al., 2018). Relatedly, providing semantic context about a dietary supplement (say, an image of a heart that conveys its function) can increase people's evaluations about the benefits they will receive from consuming that supplement (Delivett et al, 2020). Second, semantic context can even create overconfidence about one's abilities. People have come to report distorted confidence they could accurately throw a dart, execute a complex magic trick, and even land a plane in an emergency, after watching short videos of experts performing those skills (Jordan et al., 2022; Kardas and O'Brien, 2018). These examples provide evidence that semantic context boosts self-evaluations of future performance.

In sum, these findings suggest that watching a subtitled, foreign-language video clip should enhance people's comprehension of information in the moment. But these findings also suggest that

people might also mistake this ease of comprehension as evidence they could better understand the foreign language in other contexts. So, when people who speak only English watch a short video clip from the Danish TV series "Rita," English subtitles should make it easier to comprehend the story—but people might also mistakenly interpret this ease as evidence more broadly and infer they would be better able to understand Danish in new situations.

To examine these hypotheses, we asked this general question: To what extent would people who watched a subtitled Danish video clip become more confident in their ability to understand Danish in general, compared with those who watched the same video clip without subtitles? In Experiments 1a and 1b, we show that people who watched a brief, subtitled video from a Danish TV show thought they would understand a greater percentage of Danish dialogue if they were to see a new clip, without subtitles, from the same episode of that show. In Experiment 2, we show a similar illusion, such that people who watched the subtitled video thought they would be better able to understand Danish in novel social situations. In Experiment 3, we show this illusion is not dependent on a particular video, nor is it caused by people actually learning Danish, nor by people believing they could use English instead of Danish in the novel situations. In Experiment 4, we show the illusion persisted even after we asked people to listen only to the audio track from a video, thereby reducing social information and other contextual cues available in images. When taken as a whole, our results support the hypothesis that [a] semantic context boosts people's confidence by making it easier to understand the situation at hand, but [b] people can confuse that ease of processing as evidence of their ability to comprehend a foreign language in future novel situations.

Experiment 1a

Our primary aim in Experiment 1a was to develop materials and a method for future research and to assess the efficacy of both.

Method

We preregistered all experiments. These preregistrations, as well as materials and data, are available at: https://researchbox.org/853&PEER_REVIEW_passcode=BMTCAC.

Subjects

We recruited subjects from Amazon's Mechanical Turk (MTurk), an online source of a diverse population (Mason & Suri, 2012). In the absence of data about the size of an effect, we aimed for 100 subjects for each condition, anticipating exclusions. But because MTurk and Qualtrics interact in such a way that it is possible to overcollect a target sample size, we collected data from 451 Mechanical Turk workers. We analysed the data from a total of 366[1] subjects after exclusions ($M_{age} = 40.60$, $SD_{age} =$ 13.02; 33% identified as men, 66% identified as women, and 1% identified as gender diverse; 94% of subjects indicated that English was their first language).

Design

We used a 2(subtitles: subtitles, no subtitles) x 2(video origin: same episode, unspecified origin) between-subjects design.

Procedure

First, subjects read about the task they would complete:

---

[1] We discovered during a small pilot launch that the branching logic wasn't working in the survey design. We fixed this problem before collecting the full sample. Then we analyzed the data both including and excluding the 20 pilot participants and observed the same pattern (a significant difference between the subtitles and no subtitles conditions for the "same episode" question and no significant difference between conditions for the "general" question). Therefore we included these data to maximize the n.

With the rising popularity of online streaming services such as Netflix, people now have

easy access to a wide range of media, including foreign films. Today you are going to

watch a short clip from the Danish television series "Rita."

All subjects then watched a 1-minute video clip from "Rita." In the scene, a teacher shows a new

hire around a school, and everyone is speaking Danish. Half of the subjects were randomly assigned to

see this clip with English subtitles; half saw it without.

Next, all subjects were told "Later on in this survey we are going to show you another short video

clip, without subtitles, of people speaking Danish. This new video is 30-seconds long." Then, half of all

subjects were told "This new video clip is from the same episode of this TV show" while the other half

did not receive additional information, thereby leaving the origin of the video unspecified. Then

everyone was asked to predict" How much of the spoken dialogue do you think you will understand in

this video?" (0 = *None of the dialogue*, 100 = *All of the dialogue*). Next, everyone was asked "How

much expertise do you think is involved in understanding a foreign language" (1 = *No expertise*, 5 = *A*

great deal of expertise).

We then asked subjects a series of questions comprising checks for attention and compliance with

the instructions specifying that "As long as you complete the survey, we are going to pay you no matter

what you tell us now, so please be honest. We need your honest answer so we know how to analyze the

data you have provided us" (see supplemental for more information).

Results

Recall our primary research question: To what extent would people who watched a subtitled

Danish video clip become more confident in their ability to understand Danish, compared to those who

watched the same video clip without subtitles? Before answering this question, we excluded from

analysis the data from 85 (18.85%) subjects who [a] failed our attention check/compliance measures;

[b] provided a nonsensical description of the situation shown in the video; [c] had watched the

television show "Rita" before; or [d] reported speaking Danish. Across all exclusion criteria, 60.47% were excluded for failing one criterion, 3.49% for two, and 36.04% for more than two.

Next, we checked that people knew that understanding a foreign language requires a great deal of skill. We asked everyone" How much expertise do you think is involved in understanding another language?" (1 = *No expertise*, 5= *A great deal of expertise*). We found that regardless of whether they watched the video clip with or without subtitles, subjects reported that understanding a foreign language requires much expertise ($M_{subtitles}$=4.23, $M_{nosubtitles}$ = 4.09, M_{diff} = 0.14, 95%CI [-.07, .35], Welch's $t(319) = 1.29$ $p = .20$; Wilcoxon $Z = .93$, $p = .35$). That is, subjects knew that foreign language ability is a highly specialized skill.

Figure 1.

Subjects' predicted Danish understanding by condition (subtitles, no subtitles) in Experiments 1a and
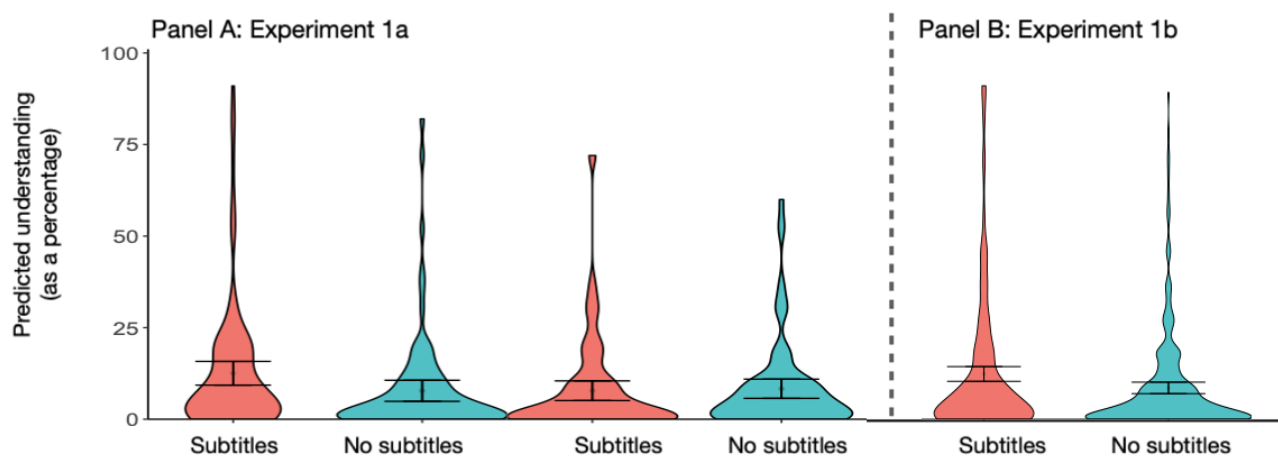
1b.

Note. Panel A displays subjects' predicted understanding of Danish in Experiment 1a. Panel B displays subjects' predicted Danish understanding ratings in Experiment 1b. Error bars represent 95% confidence intervals of cell means.

We now return to our primary question. To answer this question, we first classified subjects' ratings of understanding according to whether they saw a video with subtitles or no subtitles, and then again by whether or not they were told the "new video clip is from the same episode of this TV show" before making their rating. As Panel A shows, there was some evidence that subtitles boosted subjects' predicted understanding of Danish, relative to the no subtitles condition, but only when they were told the new clip was from the same episode of the TV show. That is, in a non-preregistered analysis, there was a marginal interaction, $F(1, 365)=3.92$, $p = .048$.

Recall that as we preregistered, we compared the effect of subtitles in two different ways. First, for people who were asked to predict their understanding of a subsequent scene in the same episode, we preregistered a t-test comparing those who saw subtitles with those who did not. Second, for people who were asked to predict their understanding of another unspecified Danish video, we also preregistered a similar t-test here.

We observed a marked skew in subjects' ratings of subjective understanding. The lower quartile was 1 and the modal response was 0; clearly, no transformation could restore normality. But t-tests are robust to most deviations in normality, especially Welch's—a parametric test that does not assume equal variance (Fagerland & Sandvik, 2009; Delacre et al., 2017). Therefore, we conducted a Welch's t-test. We also conducted a non-parametric test—a Wilcoxon signed-rank test—to buttress our main analysis. Both analyses revealed the same pattern. We display the results in Figure 1, Panel A.

As the left side of Panel A shows (and compared with their counterparts who watched the clip without subtitles) subjects who watched the video clip with subtitles predicted they would understand more Danish dialogue if they were to see a new clip from the same episode ($M_{\text{diff}} = 5.34$, 95%CI [0.99, 9.69]; Welch's t(194) = 2.55, $p = .02$; Wilcoxon $Z = 3.01$, $p = 0.01$). These findings fit with the idea that subtitles might boost people's comprehension of information in the moment and potentially create illusions of skill.

But as the right side of Panel A shows (and compared with their counterparts who watched the clip without subtitles) there was no evidence subjects who watched the clip with subtitles thought they would understand more Danish dialogue in a new clip of unspecified origin ($M_{\text{diff}} = -0.04$, 95%CI [-3.96, 3.13], Welch's $t(164) = -.23$, $p = .41$; Wilcoxon $Z = .73$, $p = .47$. These findings do not support the idea that comprehension of information necessarily creates illusions of skill in unspecified situations.

In an unplanned analysis, we ran a 2(subtitles: subtitles, no subtitles) x 2(clip origin: same episode, unspecified) between-subjects ANOVA on the data. We found an interaction, such that subtitles only inflated confidence when people were asked to consider watching another clip from the same episode, $F(1, 365)=3.92$, $p = .048$. But because this analysis is exploratory, we treat this conclusion as speculative.

Recall the primary purpose of Experiment 1a was to develop materials and a method for future research and to assess the efficacy of both. Our data provide preliminary evidence of a relative effect of subtitles—watching a foreign-language video with subtitles inflates people's predictions of their ability to understand that language in similar situations (the same episode), but not in unspecified situations. We might think of these two types of situations as providing "near" and "far" transfer, respectively (Royer, 1986; Schunk, 2012). These conclusions are speculative because we did not design Experiment 1a to compare across the "same episode" and "general" ratings. But because the data provided better evidence that our materials and method produced a "same episode" effect, Experiment 1b focused on replicating only that effect with a larger sample that would afford better estimation of the true size of the effect. We then return to the issue of far transfer in Experiment 2.

Experiment 1b

Method

Subjects

We collected data from 887 MTurk workers. After applying preregistered exclusions, 662 subjects remained in the dataset (M_{age} = 42.77, SD_{age} = 13.89; 33% identified as men, 65% identified as women, and 2% identified as gender diverse; 96% of subjects indicated that English was their first language).

Design

We used a two-group design manipulating subtitles (subtitles, no subtitles) between-subjects.

Procedure

The procedure was identical to that of Experiment 1a, except with only the "same episode" instructions. That is, when subjects were asked to predict how much of the spoken dialogue they would understand in the next video, they were all told the new video would be from the same episode of the TV show.

Results

We first excluded from analysis data from 225 subjects who [a] failed our attention checks (mostly people who failed to comply with instructions); [b] provided a nonsensical description of the situation shown in the video clip; [c] had watched the television show "Rita" before; or [d] reported speaking Danish. Across all exclusion criteria, 41% were excluded for failing one criterion, 12% for two, and 47% for more than two. We therefore retained a total sample of 662 subjects.

Recall our primary research question was: To what extent would people who watched a subtitled Danish video clip become more confident in their ability to understand Danish, compared with those who watched the same video clip without subtitles? To answer this question, we first classified subjects' understanding ratings according to whether they saw the video clip with or without subtitles, and display the results in Figure 1, Panel B. As Panel B shows, we again found that subjects who

watched the subtitled video of people speaking Danish were more confident in their own ability to understand the dialogue in a new clip from the same episode, compared with people who watched the video without subtitles (M_{diff} = 4.17, 95%CI [1.38, 6.97], Welch's $t(639)$ = 2.93, p = .02; Wilcoxon Z = -1.96, p = .05). These findings provide further support for the idea that subtitles inflate people's predictions of their ability to understand that language in similar situations.

But again—as in Experiment 1a—we found that subjects reported that understanding another language requires a great deal of expertise, regardless of whether they watched the video with subtitles or without subtitles ($M_{subtitles}$ = 4.08, $M_{nosubtitles}$ = 4.00, M_{diff} = 0.08, 95% CI [-0.09, 0.24], Welch's $t(649)$= .97, p = .33; Wilcoxon Z = 1.26, p = .21).

Experiment 2

In Experiment 2, we addressed the possibility that the reason we saw no evidence people thought they could understand the dialogue in the unspecified video is not that it was a far transfer task, but that it was too abstract, and hard to imagine. For instance, asking people to consider seeing another video of people speaking Danish sets up an unconstrained future, in which the video could be almost anything—ice hockey, standup comedy, a Ministry of Finance press conference, or two people having coffee. We attempted to address this concern in two different ways, returning to "far transfer" tasks that were more constrained. Rather than having people consider understanding another video of people speaking Danish, we asked people to imagine interacting with the people in the same Danish video and then to rate their ability to [a] follow directions, and [b] make friends in this context.

Method

Subjects

In the absence of relevant data about the size of an effect for the new "follow directions" and "make friends" dependent measures, we aimed for 150 subjects for each condition, anticipating exclusions. We collected data from 317 MTurk workers. After exclusions, we analysed the data from 295 subjects (M_{age} = 44.48, SD_{age} = 13.67; 33% identified as men, 67% identified as women; 95% of subjects indicated that English was their first language).

Design

We used a two-group design manipulating subtitles (subtitles, no subtitles) between-subjects.

Procedure

The procedure was same as that of Experiment 1a and 1b, with the exception of the key dependent measures. After subjects watched the clip, we asked them to "now insert yourself into this scenario and imagine interacting with these people." Then everyone answered two questions in a random order: " If one of these teachers were to give you directions to a classroom, how well would you be able to follow those directions?" and "If you were a new student at this school, how well would you be able to make friends?" To increase fluent responding, and because we were no longer asking

subjects to estimate a percentage, we changed the scale from 1-100 to 1-5 and changed the anchors so people considered the fluent response first (1 = *Very well*, 5= *Not at all well*; Schwarz, 1999). But for ease of interpretation in this experiment, and in all others, we recoded and report the data to adhere to the convention in which higher numbers represent more of the construct being measured—here, better understanding of Danish.

Results

We first excluded from analysis data from 22 people who [a] failed our compliance measures; [b] provided a nonsensical description of the situation shown in the video; [c] had watched the television show "Rita" before; or [d] reported speaking Danish. Across all exclusion criteria, 77% were excluded for failing one criterion, 18% for 2 two, and 5% for more than two. We therefore analysed the data from 295 subjects.

Recall our primary research question: To what extent does watching a subtitled video of people speaking a foreign language inflate people's judgments about how well they can understand that language in new situations? We followed the same approach as in Experiment 1a and 1b to classify and calculate the data. We display these results in Figure 2. As the figure shows, subjects who watched the video of people speaking Danish with subtitles thought they would be better able to follow directions and to make friends than people who watched the video without subtitles (Follow Directions: $M_{\text{diff}} = 0.44$, 95%CI [0.14, 0.74], Welch's $t(291) = 2.90$, $p = .01$; Wilcoxon $Z = -2.86$, $p = .01$; Make Friends: $M_{\text{diff}} = .27$, 95%CI [0.01, 0.53], Welch's $t(293) = 2.05$, $p = .04$; Wilcoxon $Z = 2.15$, $p = .03$). Put another way, if we consider the no subtitles group as people's baseline ability to understand Danish, watching the video with subtitles increased subjective understanding by 16.3%[2]. These findings provide further support for the idea that when subtitles ease comprehension of information, people use that ease as evidence they could understand a foreign language in new situations.

---

[2] We calculated this relative difference score via the following equation: ($M_{\text{subtitles}}$ - $M_{\text{nosubtitles}}$ / $M_{\text{nosubtitles}}$). Relative to the no subtitles group, the subtitles group rated their subjective Danish understanding 16.3% higher than the no-subtitles group.

Figure 2.
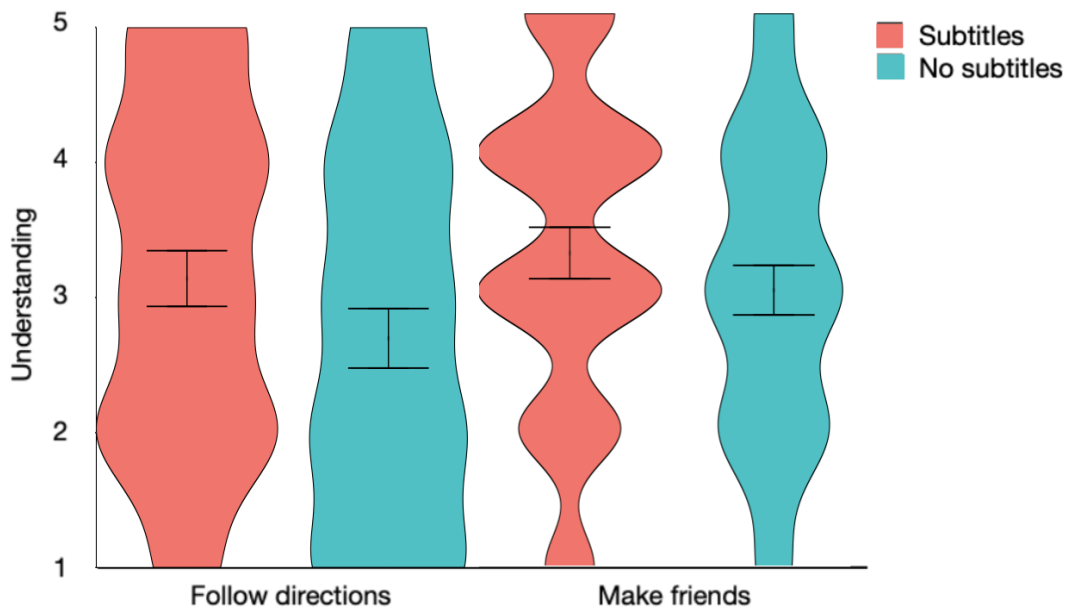
Subjects' mean ratings about how well they could "make friends" and "follow directions" in Danish.

Note. Error bars represent 95% confidence intervals of cell means.

As in Experiment 1a and 1b, regardless of whether they watched the video clip with or without subtitles, subjects reported that understanding a foreign language requires much expertise, $M_{subtitles}$ = 4.18, $M_{nosubtitles}$ = 4.01, M_{diff} = 0.17, 95%CI [-0.05, 0.40], Welch's $t(293)$ = 1.53, p = .13; Wilcoxon Z = 1.89, p = .06.

Considered together, these results suggest that relative to the no subtitle clip, watching a subtitled Danish video clip can inflate people's ratings of how well they will be able to understand Danish in future scenarios. Having said that, several obvious counterexplanations are worth addressing. First, it is possible people who see subtitles actually learn more Danish than those who don't, and so their inflated ratings of Danish understanding are justified. Therefore, to address this counter-explanation, in Experiment 3 we gave everyone a short test of Danish vocabulary, at the end of the session.

Second, it could be that this effect is tied to the specific video. For instance, it might be that something about the friendly banter in the scene might encourage people to think they would thrive in Danish social settings. Therefore, to address this alternative explanation in Experiment 3 we counterbalanced whether people saw the original Rita video, or another Danish video depicting a serious, chaotic political debate.

Third, it is possible subjects thought they could communicate in English in the transfer situations. After all, the framing of the situation did not make it clear that everyone in the transfer situations would be speaking Danish. Of course, such a criticism does not explain the confidence-boosting effect of subtitles, but it might be introducing noise. Therefore, to address this counter-explanation, in Experiment 3 we made it clear that the people would be speaking Danish only in the transfer situations.

Experiment 3

Method

Subjects

We used the Shiny Web app "power for two independent groups *t*-test" (https://designingexperiments.com/shiny-r-web-apps/) to calculate sample size based on the data from Experiment 2. Using those data and a desired power of .90, the target *n* per condition was 157, or a total of 314. We collected data from 357 workers before exclusions. After exclusions, we analysed the data from 308 subjects (M_{age} = 40.34, SD_{age} = 14.38; 36% identified as men, 63% identified as women, and 1% identified as gender diverse; 90% of subjects indicated that English was their first language).

Design

We used a between-subjects design with two conditions (subtitles: subtitles, no subtitles).

Procedure

First, subjects were randomly assigned to watch one of two videos of people speaking Danish, either with subtitles or without subtitles. One video was the "Rita" school tour used in Experiments 1a, 1b, and 2. The other video was from the Danish political drama "Borgen" and depicted several politicians engaged in a live, chaotic, televised debate. We counterbalanced the video that subjects saw. Subjects were then asked the following 4 questions, in a randomized order: "How well would you be able to follow Danish instructions in an emergency?", "If you were a new student at a Danish school and your peers spoke only Danish, how well would you be able to make friends?", "If you were listening to the Danish TV news, how well would you be able to understand the first story covered?", and "How well would you be able to understand the weather forecast on Danish TV?". All items were rated on 5-point likert scales (1 = *Quite well*, 5 = *Not at all well*). Next, everyone was asked "How much expertise do you think is involved in understanding a foreign language?" (1 = *No expertise*, 5= *A great deal of expertise*). Subjects then listened to a series of Danish words (17 in total; some that they

had heard in the video and some others that are commonly used[3]) and were asked to translate these words into English.

Results

Planned analyses

First, we excluded from analysis data from 49 subjects who [a] failed our compliance measures; [b] provided a nonsensical description of the situation shown in the video; [c] had watched the television show "Rita" or "Borgen" before; or [d] reported speaking Danish. Across all exclusion criteria, 72% were excluded for failing one criterion, 14% for two, and 14% for more than two.

We calculated the mean of all four dependent measures for each subject (these measures were correlated; see supplemental material for more information). We then classified that mean according to whether or not subjects watched the video with subtitles. We display these findings in Figure 3. As the figure shows, subjects who saw subtitles thought they would be better able to apply their newly-acquired comprehension of Danish to novel situations than those who didn't see subtitles, $M_{subtitles}=1.87$, $M_{nosubtitles}=1.59$, $M_{diff}=.29$, 95%CI [0.11, 0.46], Welch's $t(288)=3.28$, $p=.01$; Wilcoxon $Z=3.13$, $p=.01$.
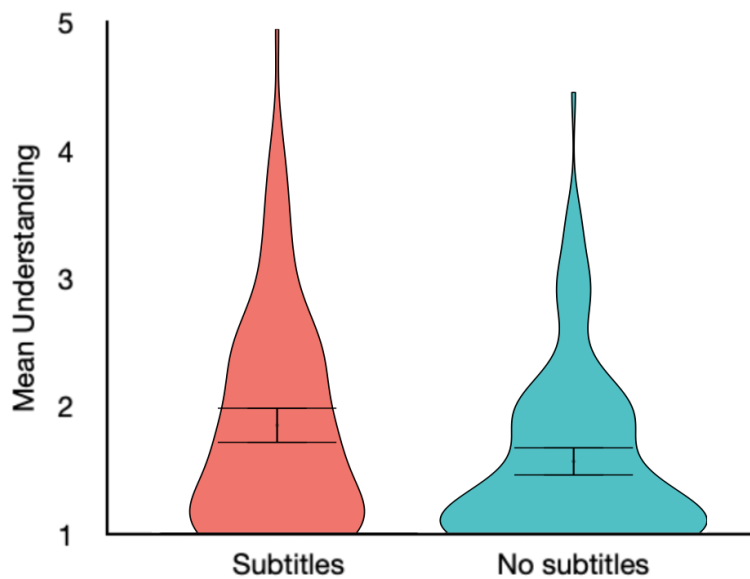
Figure 3.

---

[3] We included some common Danish words that were not in the videos but were related to the content of the scenarios we were asking people about. If people did have any knowledge of a few simple Danish words, it's possible they would do better in the transfer scenarios.

Subject's mean ratings about how well they would be able to understand Danish in new situations by condition (subtitles, no subtitles).

Note. The figure displays subjects' mean ratings across the four dependent variables measuring subjects' subjective ability to understand Danish in new situations. Error bars represent 95% confidence intervals of cell means.

In a follow-up preregistered analysis, we repeated the analyses for each dependent measure considered individually. Subjects who saw subtitles thought they would be better able to follow Danish instructions in an emergency, understand the first story on the Danish TV news, and understand the weather forecast than subjects who didn't see subtitles (all reverse coded so higher bars indicate better understanding: Follow Instructions: $M_{diff} = 0.34$, 95%CI [0.15, 0.53], Welch's $t(275) = 3.46$, $p =$ <.001; Wilcoxon $Z = 3.39$, $p =$ <.001; News: $M_{diff} = .35$, 95%CI [0.13, 0.57], Welch's $t(290) = 3.14$, $p = .01$; Wilcoxon $Z = 3.44$, $p =$ <.001; Weather: $M_{diff} = .34$, 95%CI [0.11, 0.57], Welch's $t(297) = 2.85$, $p = .01$; Wilcoxon $Z = 2.84$, $p = .01$)[4]. There was no evidence subjects who saw subtitles thought they

---

[4] The supplemental material contains figures showing the pattern for each of the individual items comprising that mean: news, weather, instruction, and make friends.

would be better able to make friends than those who didn't (M_{diff} = 0.11, 95%CI [-0.13, 0.35], Welch's $t(303)$ = .93, p = .35 ; Wilcoxon Z = .86, p = .39).

We then calculated subjects' mean percentage score on the translation test. Subjects performed poorly, regardless of subtitle condition ($M_{subtitles}$= 0.31%, $M_{nosubtitles}$ =0.30%, M_{diff} = 0.01%, 95%CI [-0.28, 0.30], Welch's t(306) = .08, p = .94 ; Wilcoxon Z = .07, p = .94)[5]. These results suggest that relative to the no-subtitle clip, watching the subtitled clip did not result in better learning. In other words, we found no evidence that people who watched the subtitled video were learning any Danish from the short video.

Finally, both those who saw subtitles and those who did not still rated the skill of understanding a foreign language as requiring considerable expertise, ($M_{subtitles}$ = 4.09, $M_{nosubtitles}$ = 4.19, Mdiff = -0.10, 95%CI [-0.32, 0.11], Welch's $t(303)$ = .94, p = .35; Wilcoxon Z = .88, p = .38), a finding consistent with our prior experiments.

Considered together, the results from Experiment 3 provide evidence that subtitles produce a consistent illusion, whereby people think they could better understand the foreign language more generally, even though they didn't learn anything. But there is still no direct evidence of subtitles increasing experienced comprehension in the moment relative to the no subtitles condition. The following analyses address this question.

Unplanned analyses

As part of routine analysis to assess new materials, we received feedback from colleagues that they found the new Borgen debate video harder to follow than the Rita school tour video. On the basis of this feedback, we examined our data to determine if the two different videos produced different mean understanding ratings. We carried out an unplanned 2(subtitles: subtitles, no subtitles) x 2(clip type: Rita, Borgen) analysis on subjects' mean rating of all four dependent measures (news, weather,

---

[5] There was no difference in translation test accuracy between words heard in the video and common Danish words.

instruction, and make friends). We found that it mattered which video subjects were randomly assigned to watch. That is, people who watched Rita thought they would understand more Danish than people who watched Borgen, $F(1, 304) = 4.61$, $p = .03$. But those saw the subtitled videos thought they would understand more Danish than those who did not, $F(1, 304) = 10.59$, $p = .001$. There was no evidence of an interaction. Because this analysis was unplanned, we treat these results as tentative.

We then addressed these same issues a priori by collecting new data (see supplemental materials for more information). More specifically, we wanted to know whether the lower understanding ratings for the new Borgen video were produced because that video was hard to follow. We used a 2(subtitles: subtitles, no subtitles) x 2(clip type: Rita, Borgen) between-subjects design. But instead of measuring subjects' ability to apply their newly-acquired, subjective comprehension of Danish to novel situations, we asked 7 questions about how easy it was for subjects to understand and follow the storyline. These questions included "How well do you believe you understood what happened in the story?", and "As you watched the film, how easy was it for you to predict what would happen next?" When we took the mean across these 7 items, the data converge on the idea that the clip mattered; people who watched the Rita clip found it easier to follow the story than those who watched the Borgen clip, $F(1,136)=39.45$, $p <.001$. We also found subjects who watched the clip with subtitles found it easier to follow the story than those who watched it without subtitles, $F(1,136)=26.89$, $p <.001$. (see supplemental information for full data on individual items).

If we step back now, our exploratory analysis and our a priori analysis support the idea that subtitles enhance people's comprehension of information in the moment, and people mistake this ease of comprehension as evidence they could understand the foreign language more generally.

Experiment 4

Although our general research question concerns the extent to which adding semantic context boosts illusions of foreign language ability, in Experiments 1-3 we provide visual context— in addition

to our subtitles manipulation —that we can't control. We know, for example, that nonverbal information such as body movement, facial expressions, and gesturing plays an important role in communication (Kellerman, 1992). The characters in our video clips gestured frequently, and we know that sociogestural information can convey meaning that observers reliably interpret (Goldin-Meadow & Alibali, 2013). In fact, second language learners' comprehension improves when they watch videos that show facial expressions and gesturing in addition to hearing foreign language audio (Sueyoshi & Hardison, 2005).

In Experiment 4 we aimed to develop new materials that would minimize social information and eliminate sociogestural information. Therefore, we removed all visual information from the video clips so subjects could hear only the foreign audio, isolating, and perhaps even emphasising, the effect of verbal context. We also adapted work by Bransford and Johnson (1972) and Wiley and Rayner (2000). We asked subjects to listen to a native Danish speaker read a passage, entirely in Danish, that was difficult to comprehend—such as the well-known passage about washing clothes (Bransford & Johnson, 1972; refer to supplemental material). Half of the subjects first saw the title of the passage (for example, "washing clothes") and then listened to the spoken Danish while English subtitles appeared on the screen against a black background. This manipulation was equivalent to the "easier-to-comprehend" condition in Bransford and Johnson, and made it easier for people to comprehend the passage. Remaining subjects simply listened to the passage spoken in Danish and saw nothing else except the black background. This manipulation should have produced even less comprehension than the "difficult-to-comprehend" condition in Bransford and Johnson because the passage was in Danish. Subjects' experience, then, would be akin to turning on the radio to listen to the Danish news. Therefore, in Experiment 4 we addressed this question: to what extent would people who heard the Danish passage, knowing the topic and seeing English subtitles, mistake their ease of comprehension as evidence they could understand Danish in new situations?

Before answering this question, we first normed the materials, and then collected data to get a better idea of the size of the effect. The findings from these two preliminary experiments are reported in the supplemental material, but by way of overview, we found these data support the hypothesis that only the "title + subtitles" condition was easy to comprehend. We then used these findings to finalise the design and calculate power for the experiment proper.

Method

Subjects

We used the Shiny Web app, "power for two independent groups *t*-test" (https://designingexperiments.com/shiny-r-web-apps/) to calculate sample size based on pilot data. We powered for what we thought would be the smallest effect size based on those data and a desired power of 0.90; the target *n* per condition was 144, or a total of 288. We collected data from 317 workers before exclusions. After exclusions, we analysed the data from 305 subjects ($M_{age} = 40.00$, $SD_{age} = 13.74$; 34% identified as men, 65% identified as women, 1% identified as gender diverse; 94% of subjects indicated that English was their first language).

Design

The experiment used a between-subjects design with two conditions (subtitles: title + subtitles, no title or subtitles).

Procedure

There were three phases in this experiment. In the first phase, subjects were randomly assigned to listen to one of three audioclips of people speaking a difficult to comprehend passage in Danish. Half of the subjects first saw the title of the passage (for example, "washing clothes", "a trip to space", or "flying a kite") and then listened to the spoken Danish while English subtitles appeared on the screen against a black background. The other half of subjects did not see the title, and only listened to the Danish audio. Each video depicted a black screen, but the subtitles group saw English subtitles in time with the audioclip.

In the second phase, subjects were asked the following three questions in a randomised order: "How well would you be able to follow Danish instructions in an emergency?" (*1 = Quite well, 5 = Not at all well*), "If you were listening to the Danish TV news, how well would you be able to understand the first story covered?", and "How well would you be able to understand the weather forecast on Danish TV?" Then we asked subjects to rate "If you were a new student at a Danish school and your peers spoke only Danish, how well would you be able to make friends?" (*1 = Quite well, 5 = Not at all well*). Because this "make friends" dependent variable did not replicate from Experiment 2 to Experiment 3, we speculated that asking this question in a random order might contaminate the other dependent variables, so it was always the fourth question[6]. Next, everyone was asked "How much expertise do you think is involved in understanding a foreign language?" (1 = *No expertise*, 5= *A great deal of expertise*) and "How difficult was it to comprehend the video?" (1 = *Very easy*, 5= *Very difficult*).

In the third phase, subjects listened to a series of Danish words (some that they had heard in the video and some that are commonly used) and were asked to translate these words into English.

Results

Recall our primary research question was: To what extent does watching a subtitled video of people speaking a foreign language inflate people's judgments about how well they can understand that language in new situations?

Before answering this question, we excluded from analysis data from 12 subjects who [a] failed our compliance measures; or [b] reported speaking Danish. Across all exclusion criteria, 75% were excluded for failing one criterion and 25% for two.

[6] In addition, a comment from an Mturk worker suggests that this item might be interpreted differently to the other dependent variables. "I have a comment: There have been several occasions in my life when I have made friends with people who did not speak English, and whose language I did not speak; we enjoyed each other's company, were able to communicate on a basic level, and taught other words in our native languages."

We now return to our primary research question. To answer this question, we first classified subjects' understanding ratings according to whether they saw the video with the title and subtitles or without the title and subtitles. We then calculated the mean across our three key understanding ratings and classified that mean according to whether subjects watched the subtitled video of people speaking Danish, or the video without subtitles, and display the findings in Figure 4. As the figure shows, subjects who saw the title and subtitles thought they would be better able to apply their newly-acquired comprehension of Danish to novel situations than those who saw neither the title nor subtitles, $M_{diff} = 0.51$, 95%CI [0.31, 0.70], Welch's $t(259) = 5.19$, $p < .001$; Wilcoxon $Z = 5.60$, $p < .001$.

Figure 4

Subjects' mean ratings about how well they would be able to understand Danish in new situations by condition (title + subtitles, no title or subtitles).
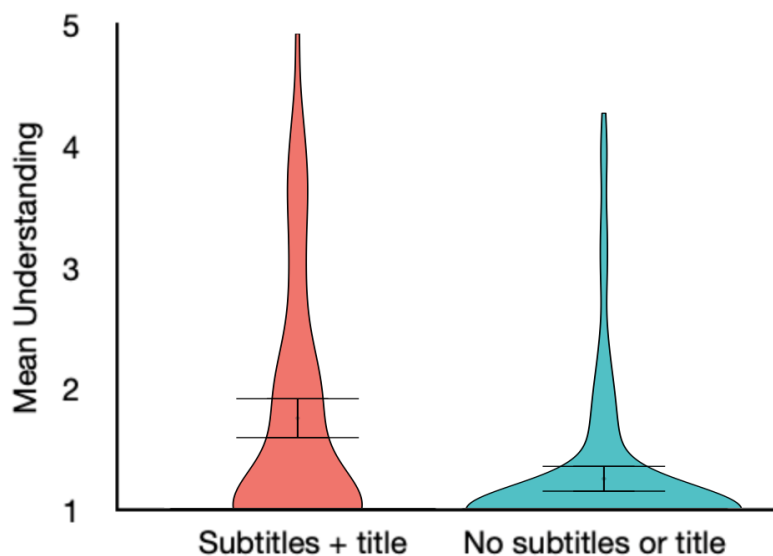
Note. The figure displays subjects' mean ratings across the three key dependent variables measuring subjects' ability to understand Danish in new situations (News, weather, instructions). Error bars represent 95% confidence intervals of cell means.

We found that subjects who watched the video of people speaking Danish with a title and subtitles thought they would be better able to follow Danish instructions in an emergency, understand the first story on the Danish TV news, and understand the weather forecast on Danish TV than those who watched the video without a title or subtitles (all reverse coded: Follow Instructions: $M_{diff} = 0.49$, 95%CI [0.27, 0.70], Welch's $t(259) = 4.49$, $p < .001$; Wilcoxon $Z = 4.78$, $p < .001$]; News: $M_{diff} = .56$, 95%CI [0.35, 0.77], Welch's $t(254) = 5.17$, $p < .001$ Wilcoxon $Z = 5.59$, $p < .001$; Weather: $M_{diff} = .47$, 95%CI [0.26, 0.69], Welch's $t(274) = 4.37$, $p < .001$; Wilcoxon $Z = 4.62$, $p < .001$)[7]. Subjects who saw a title and subtitles also thought they would be better able to make friends than subjects who saw neither ($M_{diff} = 0.26$, 95%CI [0.04, 0.48], Welch's $t(293) = 2.37$, $p = .02$; Wilcoxon $Z = 2.27$, $p = .02$). These findings provide further support for the idea that semantic context boosts people's confidence in their ability to apply their newly-acquired comprehension of Danish to novel situations.

In line with our prior data, we found support for the idea that the subtitles and title boosted people's comprehension of the video relative to people who didn't watch the video with subtitles and a title, (reverse coded: $M_{subtitles} = 3.54$, $M_{nosubtitles} = 1.32$, $M_{diff} = 2.22$, 95%CI [1.88, 2.56], Welch's $t(224) = 12.83$, $p < .001$; Wilcoxon $Z = 11.58$, $p < .001$). We also found people's ease of comprehension and mean subjective confidence ratings across the three key dependent variables were moderately associated ($r_{305} = -0.40$, 95% CI [-0.49, -0.30], $p < .001$).

But "subtitles + title" subjects did not perform better on the Danish translation test than their "no subtitles no title" counterparts. In other words, we found no evidence that people who watched the

---

[7] The supplemental material contains figures showing the pattern for each of the individual items comprising that mean and the "Friends" dependent variable.

video with subtitles were learning any Danish from the short video. $M_{subtitles} = 4.34\%$, $M_{nosubtitles} = 5.42\%$, $M_{diff} = -1.08\%$, 95%CI [-2.42, 0.26], Welch's $t(302) = 1.59$, $p = .11$; Wilcoxon $Z = 1.79$, $p = .07$.

Finally, regardless of whether or not subjects watched the video with a title and subtitles, subjects still rated the skill of understanding a foreign language as requiring considerable expertise, ($M_{subtitles} = 4.12$, $M_{nosubtitles} = 3.89$, $M_{diff} = 0.22$, 95%CI [-0.04, 0.49], Welch's $t(294) = 1.64$, $p = .10$; Wilcoxon $Z = 1.30$, $p = .19$), a finding consistent with our prior experiments.
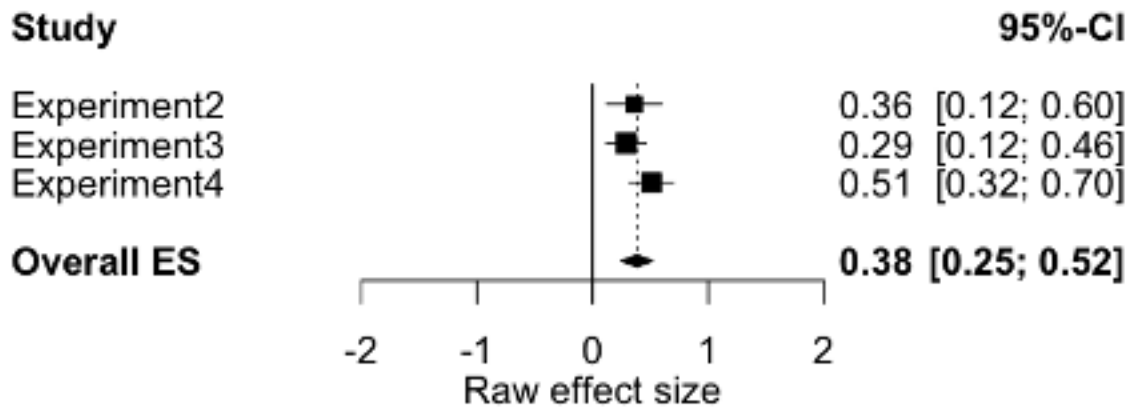
Considered together, the results from Experiment 4 suggest that the context effects we have examined this far replicate with verbal-only materials and provide further support for the hypothesis that semantic context (operationalised here as a title and subtitles) enhances people's subjective comprehension of foreign language information in the moment. But people mistake this ease of comprehension of current content as evidence they could better understand and utilise the foreign language more generally.

Mini meta-analysis

To obtain a more precise estimate of the effect of watching a foreign video with subtitles on people's confidence in their ability to apply their newly-acquired comprehension of Danish to novel situations, we conducted a mini meta-analysis of the data from Experiments 2, 3, and 4 (the experiments that measured subjective Danish understanding on a 5-point likert scale). We report the results in Figure 5.

Figure 5.

Mini meta-analysis of subject's mean understanding ratings by condition (subtitles, no subtitles) in

experiments 2,3 and 4 (R code; Carter & McCullough, 2014).

Study   95%-CI

Experiment2		0.36 [0.12; 0.60]
Experiment3		0.29 [0.12; 0.46]
Experiment4		0.51 [0.32; 0.70]
Overall ES		**0.38 [0.25; 0.52]**

```
    -2      -1       0       1       2
           Raw effect size
```

The right side of the vertical line in Figure 5 shows an effect in which subtitles (or in Experiment

4, subtitles + title) made subjects more confident in their ability to understand Danish in novel

situations. More specifically, we found a weighted raw effect size of $M_{diff} = .38$ [.25, .52], $p < 0.01$, or a

9.5% shift towards overconfidence. These findings fit with the idea that increased semantic context

creates overconfidence in one's ability to understand a foreign language.

We can put this effect size in context by turning to the related literature. If we look to the

cognitive literature, our effect size resembles the increased false recall of words presented in a

predictive semantic context (Whittlesea, 1993). Our effect sizes also resemble people's illusion of

better understanding complicated processes when they are presented with a tangentially related picture

(Cardwell, 2017). In the educational literature, our findings are consistent with work showing student's

overconfidence in their understanding of scientific concepts accompanied by uninformative

photographs (Wiley, 2019). Finally, our data are also consistent with work showing people inflate their

skill in dart-throwing after watching 20 demonstrations of an expert throwing a dart at a bullseye

(Kardas & O'Brien, 2018). Considered together, our findings align with the large body of work

showing that context can mislead people about how much they know or can do. And beyond the psychology literature, our effect size expressed as a correlation ($r = 0.214$[8]) is similar to the relationship between conscientiousness scores and job proficiency, the effect of ibuprofen on pain reduction, and the likelihood of developing lung cancer if you've ever smoked (see, for a review, Meyer et al., 2001).

General Discussion

Across five primary experiments comprising 1936 subjects, we found that situations with greater semantic context can produce rapid illusions of foreign language ability, relative to situations of lesser semantic context. More specifically, when English speakers watched a Danish video with subtitles, they thought they would be better able to understand Danish in new situations, compared to the counterparts who watched the same video without subtitles.

We also addressed several counterexplanations. For example, data from Experiment 3 suggest that this illusion is not produced because people actually learn Danish, nor because people thought they could default to English (rather than speak Danish) in the transfer situations. We also found the illusion did not hinge on a specific video. In fact, in Experiment 4, we showed the illusion does not even hinge on the video per se: we found the same pattern when we minimized social information and eliminated sociogestural information. Considered together, our findings support the idea that—relative to situations of lesser semantic context—greater semantic context produces rapid illusions of foreign language ability by making it easier for people to comprehend the situation in the moment. People then confuse that ease for comprehending the foreign language more generally.

These findings contribute to several literatures. First, they extend what we know about the "dark side" of semantic context—how increased semantic context creates illusions of understanding (Laukkonen et al., 2020). In prior work, people were asked how well they understand a specific—often

---

[8] We calculated Cohen's d and converted it to r using the following equation: $r = \{d/sqrt\{d^2 + 4\}\}$

complex—process, such as how rainbows form (Cardwell et al., 2017). When an uninformative photo (say, a rainbow) accompanied the question, people reported a better understanding of how that process works—even though subsequent analysis showed no such improved understanding. Here, we extend this work by showing that relative to situations of lesser semantic context, greater semantic context creates illusions not simply of understanding, but also of greater skill—which is not backed up by performance on an objective test of vocabulary— coupled with an ability to apply that skill to novel situations.

Second, our findings provide evidence of a potential mechanism for these sorts of illusions, in which semantic context produces feelings of easier comprehension which are then misattributed as evidence of understanding. More specifically, in Experiment 3, people who watched a video clip with accompanying subtitles, or a clearer event structure, reported the storyline was easier to follow and understand. Along the same lines, in Experiment 4, people who watched subtitled audio (plus a title) of a spoken Danish passage reported easier comprehension. Collectively, our findings support the idea that people are influenced by increased semantic context—here, mistaking easier comprehension as evidence they would be better able to understand Danish in new situations, compared to others who did not have access to this semantic context.

Third, these findings fit with work suggesting that processing fluency can be produced by a number of techniques, conceptual or perceptual, with predictable and similar effects on people's relative judgements (Alter & Oppenheimer, 2009; Stanley et al., 2022). Take judgements of truth, for example: people are more likely to say a statement is true when it repeats, rhymes, when it is paired with a photo, when words or concepts related to that statement are semantically primed, and when it is presented in high contrast (Hasher et al., 1977; Kelley & Lindsay, 1993; McGlone & Tofighbakhsh, 2000; Newman, et al., 2012; Reber & Schwarz, 1999). Obviously these manipulations do not give people "diagnostic" information about truth, but the idea is that they all help to make it feel easier to process the information, and people then misattribute that feeling to the task at hand (Unkelbach &

Greifeneder, 2013). Here, subjects who watched a video with accompanying subtitles (and sometimes a title), or a structure that made the storyline easier to follow, anticipated better ability to understand Danish in new situations (E.g. Experiment 3). And though we did not teach people Danish, our data suggest we made it easier for them to understand a video clip. The fact that perceptual and conceptual manipulations of fluency can produce similar outcomes suggests an avenue for future work—testing the possibility that perceptual manipulations, such as the clarity of the video, the contrast of the subtitles, or the quality of the audio track, would also change people's confidence about their Danish skill (Newman & Schwarz, 2018).

These findings also align with work from the metacognitive literature showing that judgements of learning (JOLs) can be influenced by easier processing (Begg et al., 1989; Kornell et al., 2011; Koriat et al., 2004). Both perceptual, and conceptual manipulations similarly inflate people's assessments about what they will recall in the future. For instance, when people are presented with lists of words, they give higher JOLs to words played at a louder volume, presented in better visual clarity, or a larger font (features that make the words feel subjectively easier to process), even though their recall is no better than words played at quiet volumes, presented in poorer quality, or a smaller font (Rhodes & Castel, 2008; Rhodes & Castel, 2009; Yue et al., 2013). Similar patterns arise when people are given word lists containing words of differing familiarity (Begg et al., 1989). People rate common words as easier to imagine and understand, and they award them higher JOLs than rare words, despite recalling less common words. Our findings suggest subtitles also create a metacognitive illusion—producing easier comprehension of information that people misattribute as evidence of learning.

Our findings are reminiscent of the illusions of knowledge people demonstrate after searching Google for information about specific topics. Later, when these same people are barred from "googling" information about new topics, they overestimate how well they will be able to answer questions about those topics—a finding that suggests failure to recognize the influence of "outsourced" knowledge (Fisher et al., 2015). In addition, people develop similar illusions of knowledge when they

quickly view previews of web pages, an instance of the "hasty glance" phenomenon, or people's tendency to quickly make judgments about a situation, given minimal or incomplete information (Eliseev & Marsh, 2023). In another study, learning that scientists understand a given scientific phenomenon boosted people's appraisals of their own understanding of that phenomenon (Sloman & Rabb, 2016). Considered together, these findings suggest people conflate information out in the world with knowledge they possess. Likewise, our findings suggest people provided with subtitles conflated that information with skills they possessed. But our findings are puzzling given the same people also reported that learning a foreign language requires a great deal of expertise. It seems plausible, then, that our subjects demonstrated "knowledge neglect," a failure to retrieve or apply relevant knowledge (Fazio et al., 2013; see for a review, Marsh & Umanath, 2014). More specifically, our findings suggest people fail to retrieve, or apply, their demonstrated knowledge of the expertise involved in understanding a foreign language when evaluating their own ability to understand that language.

Although we focused on primarily English-speakers' estimates of their skill in Danish, we know of no theoretical basis to expect other languages would produce different patterns of results. But we do hope our results spur others to replicate and extend our work to other languages, especially highly disparate languages such as Japanese and English.

In addition, although we focused primarily on subtitles as the source of semantic context, research suggests many other aspects might be relevant in the effects we report here. These include, for instance, the rate of speech, the accent of the speaker, the number of ideas presented, and the syntactic and grammatical complexity of the speech (Bloomfield et al., 2010; Lower, 1998)[9]. These aspects largely bear on fluent processing. In general, when it comes to fluency, simple is better: slower speech, simpler words, less complexity (Oppenheimer, 2006). For example, complex syntactic structures are

---

[9] We thank an anonymous reviewer for this suggestion

harder for people to comprehend (less fluent) and less persuasive than simple language (Lowrey, 1998). Considering the role of these aspects in creating illusions of skill is a fruitful avenue for future research.

Although we were primarily interested in the relative difference between subtitles and no subtitles as a way of tapping into situations of greater and lesser semantic context, we did not include control conditions in which there was no exposure to the Danish language at all. As a result, we can make only relative claims and not absolute claims. That is, relative to a situation in which people saw Danish videos without subtitles, seeing those videos with subtitles increased people's assessments of their foreign-language ability. It is possible, then, that watching videos with and without subtitles both created disfluency and reduced people's self-reports of their ability to comprehend Danish from their baseline prior to participating in the experiment. In addition, a critic might argue that subtitles rarely pushed people's estimates of their abilities into a level we would interpret as "high." Instead, what's important here is not people's estimates in absolute terms, but the relative changes in self-reported ability after a single exposure to greater or lesser semantic context. Moreover these estimates are inflated relative to objective tests of vocabulary.

What then could be the effect of increasing exposure on confidence? We know from work on the misinformation effect that repeatedly hearing suggestions increases false memories of those suggestions (Zaragoza & Mitchell, 1996). In a similar vein, repeatedly imagining performing an action, such as breaking a toothpick, increases both people's likelihood of falsely reporting they actually performed that action and their confidence in that response (Goff and Roediger, 1998, Thomas & Loftus, 2002). Likewise, work on the repetition truth effect suggests that repeating statements increases people's likelihood to rate those statements as true (Dechêne et al., 2010). Perhaps most relevant here, we know that repeatedly watching experts perform tasks—say, watching an expert dart-thrower 20 times—increases people's confidence in their own dart throwing skills compared with people who watch the video just once, even though they get no better (Kardas & O'Brien, 2018). It's plausible that

with repeated watching—say, for example, watching an entire season of a foreign TV show—the illusion could compound. This possibility is an intriguing avenue for future research.

In sum, our findings have theoretical implications for our understanding of semantic context, processing fluency, and "knowledge neglect." Our findings also have practical implications. Here, we provide evidence of a metacognitive illusion such that ease of processing in the moment can distort people's judgements about how well they understand information, which in turn changes their evaluations of how well they could apply that understanding in new situations. These findings also add to the literature showing that presenting the same information to people in different ways can change people's calibrations of their own learning, performance, and abilities. For instance, students draw on feelings of ease as evidence of learning—but those feelings of ease can inflate students' evaluations of their own skills— as a result, they invest less time and effort during learning (Dreisbach & Fischer, 2011; Luna & Albuquerque, 2022; see, for a review, Chang & Brainerd, 2022). If semantic context eases processing of new information that people misattribute as evidence of their own skill, they may invest less time in their learning, potentially leading to poorer learning outcomes. The effects we report here are new and surprising examples of illusions of skill. Our data fit with the idea that these illusions occur because semantic context eases people's understanding of the videos they are watching, but they misattribute this ease of processing as evidence of their own abilities. Our work furthers our understanding about den mørke side af semantisk kontekst [10].

---

[10] the dark side of semantic context

References

Alter, A. L., & Oppenheimer, D. M. (2009). Uniting the tribes of fluency to form a

metacognitive nation. *Personality and Social Psychology Review*, *13*(3), 219–235.

https://doi.org/10.1177/1088868309341564

Ausubel, D. P. (1960). The use of advance organizers in the learning and retention of

meaningful verbal material. *Journal of educational psychology*, *51*(5), 267.

Begg, I., Duft, S., Lalonde, P., Melnick, R., & Sanvito, J. (1989). Memory predictions are based on

ease of processing. *Journal of Memory and Language*, *28*(5), 610-632.

https://doi.org/10.1016/0749-596X(89)90016-8.

Bisson, M. J., Van Heuven, W. J., Conklin, K., & Tunney, R. J. (2014). Processing of native and

foreign language subtitles in films: An eye tracking study. *Applied Psycholinguistics*, *35*(2), 399-

418.

Bloomfield, A., Wayland, S. C., Rhoades, E., Blodgett, A., Linck, J., & Ross, S. (2010). What makes

listening difficult? Factors affecting second language listening comprehension. *University of

Maryland Center for Advanced Study of Language*, 3-79.

Bower, G. H. (1978). Experiments on story comprehension and recall. *Discourse Processes*, *1*(3), 211-

231.

Bransford, J.D., & Johnson, M. (1972). Contextual prerequisites for understanding:

    Some Investigations of comprehension and recall. *Journal of Verbal Learning*

    and Verbal Behavior, *11*(6), 717-726.

Cardwell, B. A., Lindsay, D. S., & Garry, M. (2017). Uninformative photos can

    increase people's perceived knowledge of complicated processes. *Journal of*

    Applied Research in Memory and Cognition, *6*(3), 244–252.

    https://doi.org/10.1016/j.jarmac.2017.05.002

Cardwell, B. A., Newman, E. J., Garry, M., Mantonakis, A., & Beckett, R. (2017). Photos that increase

    feelings of learning promote positive evaluations. *Journal of Experimental Psychology:*

    Learning, Memory, and Cognition, *43*(6), 944–954. https://doi.org/10.1037/xlm0000358

Carpenter, S. K. (2012). Testing enhances the transfer of learning. *Current directions in psychological*

    science, 21(5), 279-283.

Carpenter, S. K., Olson, K. M. (2012). Are pictures good for learning new vocabulary in a foreign

    language? Only if you think they are not. *Journal of Experimental Psychology: Learning,*

    Memory, and Cognition, *38*, 92–101.

Carpenter, S. K., Wilford, M. M., Kornell, N., & Mullaney, K. M. (2013). Appearances can be

    deceiving: Instructor fluency increases perceptions of learning without increasing actual

    learning. *Psychonomic bulletin & review*, *20*(6), 1350-1356. https://doi.org/10.3758/s13423-

    013-0442-z

Carter, E. C. & McCullough, M. E. (2014). Publication bias and the limited strength model of self-

    control: has the evidence for ego depletion been overestimated? *Frontiers in Psychology, 5*,

    823. https://doi.org/10.3389/fpsyg.2014.00823)

Chang, M., & Brainerd, C. J. (2022). Association and dissociation between judgments of learning and memory: A meta-analysis of the font size effect. *Metacognition and Learning*. Advance online publication. https://doi.org/10.1007/s11409-021-09287-3

Dechêne, A., Stahl, C., Hansen, J., & Wänke, M. (2010). The Truth About the Truth: A Meta-Analytic Review of the Truth Effect. *Personality and Social Psychology Review*, *14*(2), 238–257. https://doi.org/10.1177/1088868309352251

Delacre, M., Lakens, D., & Leys, C. (2017). Why psychologists should by default use Welch's t-test instead of Student's t-test. *International Review of Social Psychology*, *30*(1). doi:10.5334/irsp.82

Delivett, C. P., Klepacz, N. A., Farrow, C. V., Thomas, J. M., Raats, M. M., & Nash, R. A. (2020). Front-of-pack images can boost the perceived health benefits of dietary products. *Appetite*, 155, 104831. https://doi.org/10.1016/j.appet.2020.104831

Dreisbach, G., Fischer, R. (2011). If it's hard to read . . . try harder! Processing fluency as signal for effort adjustments. *Psychological Research*, *75*, 376–383. doi:10.1007/s00426-010-0319-y

Eisenberg, M. L., Zacks, J. M., & Flores, S. (2018). Dynamic prediction during perception of everyday events. *Cognitive research: principles and implications*, *3*(1), 1-12. https://doi.org/10.1186/s41235-018-0146-z

Eliseev, E. D., & Marsh, E. J. (2023). Understanding Why Searching the Internet Inflates Confidence in Explanatory Ability. *Applied Cognitive Psychology*, 37(1). https://doi.org/10.1002/acp.4058

Fagerland, M. W., & Sandvik, L. (2009). Performance of five two-sample location tests for skewed distributions with unequal variances. *Contemporary clinical trials*, *30*(5), 490-496. https://doi:10.1016/j.cct.2009.06.007

Fazio, L. K., Barber, S. J., Rajaram, S., Ornstein, P. A., & Marsh, E. J. (2013). Creating illusions of knowledge: Learning errors that contradict prior knowledge. *Journal of Experimental Psychology: General*, *142*(1), 1–5.

Fisher, M., Smiley, A. H., & Grillo, T. L. (2022). Information without knowledge: the effects of Internet search on learning. *Memory*, *30*(4), 375-387. https://doi.org/10.1080/09658211.2021.1882501

Gernsbacher, M. A. (2015). Video captions benefit everyone. *Policy insights from the behavioral and brain sciences*, *2*(1), 195-202. https://doi.org/10.1177/2372732215602130

Goff, L. M., & Roediger, H. L. III. (1998). Imagination inflation for action events: Repeated imaginings lead to illusory recollections. *Memory & Cognition*, *26*(1), 20–33. https://doi.org/10.3758/BF03211367

Goldin-Meadow, S., & Alibali, M. W. (2013). Gesture's role in speaking, learning, and creating language. *Annual Review of Psychology*, *64*, 257-283. https://doi.org/10.1146/annurev-psych-113011-143802

Hasher, L., Goldstein, D., & Toppino, T. (1977). Frequency and the conference of referential validity. *Journal of Verbal Learning & Verbal Behavior, 16*(1), 107–112. https://doi.org/10.1016/S0022-5371(77)80012-1

Jordan, K., Zajac, R., Bernstein, D., Joshi, C., & Garry, M. (2022). Trivially informative semantic context inflates people's confidence they can perform a highly complex skill. *Royal Society Open Science*, *9*(3), 211977. https://doi.org/10.1098/rsos.211977

Kardas, M., & O'Brien, E. (2018). Easier seen than done: Merely watching others perform can foster an illusion of skill acquisition. *Psychological Science*, *29*(4), 521–536. https://doi.org/10.1177/0956797617740646

Kellerman, S. (1992). 'I see what you mean': The role of kinesic behaviour in listening and implications for foreign and second language learning. *Applied linguistics*, *13*(3), 239-258.

Kelley, C. M., & Lindsay, D. S. (1993). Remembering mistaken for knowing: Ease of retrieval as a basis for confidence in answers to general knowledge questions. *Journal of Memory and Language*, *32*, 1-24.

Kornell, N., Rhodes, M. G., Castel, A. D., & Tauber, S. K. (2011). The ease-of-processing heuristic and the stability bias: Dissociating memory, memory beliefs, and memory judgments. *Psychological Science*, *22*(6), 787–794. https://doi.org/10.1177/0956797611407929

Koriat, A., Bjork, R. A., Sheffer, L., & Bar, S. K. (2004). Predicting One's Own Forgetting: The Role of Experience-Based and Theory-Based Processes. *Journal of Experimental Psychology: General, 133*(4), 643–656. https://doi.org/10.1037/0096-3445.133.4.643

Laukkonen, R. E., Kaveladze, B. T., Tangen, J. M., & Schooler, J. W. (2020). The dark side of Eureka: Artificially induced aha moments make facts feel true. *Cognition, 196,* Article 104122. https://doi.org/10.1016/j.cognition.2019.104122

Lakshmanan, A., & Krishnan, H. S. (2011). The Aha! Experience: Insight and Discontinuous Learning in Product Usage. *Journal of Marketing*, *75*(6), 105–123. https://doi.org/10.1509/jm.10.0348

Lowrey, T. M. (1998). The effects of syntactic complexity on advertising persuasiveness. Journal of consumer psychology, 7(2), 187-206. https://doi.org/10.1207/s15327663jcp0702_04

Luna, Karlos & B. Albuquerque, Pedro. (2022). Do beliefs about font size affect retrospective metamemory judgments in addition to prospective judgments?: A registered report. *Experimental Psychology.* Advance online publication. https://doi.org/10.1027/1618-3169/a000549.

Marsh, E. J., & Umanath, S. (2014). Knowledge neglect: Failures to notice contradictions with stored knowledge . In D. N. Rapp and J. Braasch (Eds) *Processing Inaccurate Information: Theoretical and Applied Perspectives from Cognitive Science and the Educational Sciences* (p. 161-180). MIT Press.

Mason, W., & Suri, S. (2012). Conducting behavioral research on Amazon's Mechanical Turk. *Behavior Research Methods*, *44*(1), 1-23. https://doi.org/10.3758/s13428-011-0124-6

Mayer, R. E., & Gallini, J. K. (1990). When is an illustration worth ten thousand words?. *Journal of Educational Psychology*, *82*(4), 715.

McGlone, M. S., & Tofighbakhsh, J. (2000). Birds of a feather flock conjointly (?): Rhyme as reason in aphorisms. *Psychological Science*, *11*, 424-428.

Meyer, G. J., Finn, S. E., Eyde, L. D., Kay, G. G., Moreland, K. L., Dies, R. R., ... & Reed, G. M. (2001). Psychological testing and psychological assessment: A review of evidence and issues. *American Psychologist, 56*(2), 128.

Newman, E. J., & Schwarz, N. (2018). Good sound, good research: How audio quality influences perceptions of the research and researcher. *Science Communication*, *40*(2), 246-257. https://doi.org/10.1177/1075547018759345

Newman, E. J., Garry, M., Bernstein, D. M., Kantner, J., & Lindsay, D. S. (2012). Nonprobative photographs (or words) inflate truthiness. *Psychonomic Bulletin & Review*, *19*(5), 969-974. https://doi.org/10.3758/s13423-012-0292-0

Newman, E.J., Azad, T., Lindsay, D.S., & Garry, M. (2018). Evidence that photos promote rosiness for claims about the future. *Memory & Cognition*, *46*(8), 1223–1233. https://doi.org/10.3758/s13421-016-0652-5

Oppenheimer, D. M. (2006). Consequences of erudite vernacular utilized irrespective of necessity: Problems with using long words needlessly. Applied Cognitive Psychology, 20(2), 139–156. https://doi.org/10.1002/acp.1178

Oppenheimer, D. M. (2008). The secret life of fluency. *Trends in Cognitive Sciences*, *12*(6), 237–241. https://doi.org/10.1016/J.TICS.2008.02.014

Perego, E., Del Missier, F., Porta, M., & Mosconi, M. (2010). The cognitive effectiveness of subtitle processing. *Media Psychology*, *13*(3), 243-272, https://doi.org/10.1080/15213269.2010.502873

Reber, R., & Schwarz, N. (1999). Effects of perceptual fluency on judgments of truth. *Consciousness and Cognition*, *8*, 338-342.

Rhodes, M. G. & Castel, A. D. (2008). Memory predictions are influenced by perceptual information: Evidence for metacognitive illusions. *Journal of Experimental Psychology: General, 137*(4), 615–625. https://doi.org/10.1037/a0013684

Rhodes, M.G., Castel, A.D. (2009). Metacognitive illusions for auditory information: Effects on monitoring and control. *Psychonomic Bulletin & Review, 16*, 550–554. https://doi.org/10.3758/PBR.16.3.550

Royer, J. M. (1986). Designing instruction to produce understanding: An approach based on cognitive theory. In G. D. Phye & T. Andre (Eds.), *Cognitive classroom learning: Understanding, thinking, and problem solving* (pp. 83–113). Orlando, FL: Academic Press.

Schunk, D. H. (2012). Learning theories an educational perspective sixth edition. Pearson.

Schwarz, N. (1999). Self-reports: how the questions shape the answers. *American Psychologist, 54*, 93-105. https://doi.org/10.1037/0003-066X.54.2.93

Schwarz, N., Jalbert, M., Noah, T., & Zhang, L. (2021). Metacognitive experiences as information: Processing fluency in consumer judgment and decision making. *Consumer Psychology Review*, *4*(1), 4-25. https://doi.org/10.1002/arcp.1067

Sloman, S. A., & Rabb, N. (2016). Your understanding is my understanding: Evidence for a community of knowledge. *Psychological Science*, 27(11), 1451-1460. https://doi.org/10.1177/0956797616662271

Stanley, M. L., Whitehead, P. S., & Marsh, E. J. (2022). The cognitive processes underlying false beliefs. *Journal of Consumer Psychology*, *32*(2), 359-369. https://doi.org/10.1002/jcpy.1289

Sueyoshi, A., & Hardison, D. M. (2005). The role of gestures and facial cues in second language listening comprehension. *Language Learning*, *55*(4), 661-699. https://doi.org /10.1111/j.0023-8333.2005.00320.x

Thomas, A. K., & Loftus, E. F. (2002). Creating bizarre false memories through imagination. *Memory & Cognition*, *30*(3), 423–431. https://doi.org/10.3758/BF03194942

Toftness, A. R., Carpenter, S. K., Geller, J., Lauber, S., Johnson, M., & Armstrong, P. I. (2018). Instructor fluency leads to higher confidence in learning, but not better learning. *Metacognition and Learning*, *13*(1), 1-14. https://doi.org/10.1007/s11409-017-9175-0

Unkelbach, C., & Greifeneder, R. (2013). A general model of fluency effects in judgment and decision making. In. C.Unkelbach, & R. Greifeneder. *The experience of thinking: How the fluency of mental processes influences cognition and behaviour*, (pp. 11-32) London: Psychology Press.

Whittlesea, B. W. A. (1993). Illusions of familiarity. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *19*(6), 1235-1253. https://doi.org/10.1037/0278-7393.19.6.1235

Wiley, J. (2019). Picture this! Effects of photographs, diagrams, animations, and sketching on learning and beliefs about learning from a geoscience text. *Applied Cognitive Psychology. 33*, 9-19. https://doi.org/10.1002/acp.3495

Wiley, J., & Rayner, K. (2000). Effects of titles on the processing of text and lexically ambiguous words: Evidence from eye movements. *Memory & Cognition*, *28*(6), 1011–1021. https://doi.org/10.3758/BF03209349

Witherby, A. E., & Carpenter, S. K. (2022). The impact of lecture fluency and technology fluency on students 'online learning and evaluations of instructors. *Journal of Applied Research in Memory and Cognition*. Advance online publication. https://doi.org/10.1037/mac0000003

Yue, C. L., Castel, A. D., & Bjork, R. A. (2013). When disfluency is—And is not—A desirable difficulty: The influence of typeface clarity on metacognitive judgments and memory. *Memory & Cognition, 41*(2), 229–241. https://doi.org/10.3758/s13421-012-0255-8

Zacks, J. M. (2020). Event perception and memory. *Annual Review of Psychology*, *71*, 165-191. https://doi.org/10.1146/annurev-psych-010419-051101