

Intralingual live subtitling in EMI lectures in Flanders: Students' perceptions and performance

Yanou Van Gauwbergen, Isabelle S. Robert & Iris Schrijver

University of Antwerp

Abstract

Many universities are considering using English as a means of instruction (EMI), but students' limited proficiency in English could be a serious drawback. Live subtitling might help to overcome this language barrier. The aim of this article is to report on (1) how university students in Flanders perceive EMI lectures accompanied by intralingual live subtitles; and (2) whether these subtitles influence their performance. In this study, the impact of subtitling on students' perceptions and performance was investigated during six two-hour Research Skills lectures taught in English to students of Applied Linguistics who have Dutch as their mother tongue. The live subtitling was alternately produced through respeaking and through automatic speech recognition (ASR). The data were collected using (1) tests after each lecture about their perceptions and the content of the lecture (performance); (2) an online questionnaire about the students' demographics (e.g., mother-tongue and self-reported proficiency in English); and (3) online language tests, which consisted of a listening test and a vocabulary test. The results show that the impact of subtitles on students does not align with the findings of earlier research, as students' performance with subtitles was lower. This has implications for the possible implementation of live subtitling in education.

Keywords

Intralingual live subtitling, EMI, respeaking, ASR

CONTACT:

 Yanou Van Gauwbergen, yanou.vangauwbergen@uantwerpen.be, University of Antwerp, Grote Kauwenberg 18, 2000 Antwerp, Belgium

 Isabelle S. Robert, isabelle.robert@uantwerpen.be, University of Antwerp, Grote Kauwenberg 18, 2000 Antwerp, Belgium

 Iris Schrijver, iris.schrijver@uantwerpen.be, University of Antwerp, Grote Kauwenberg 18, 2000 Antwerp, Belgium

1. Introduction

Universities are increasingly choosing to offer classes in English. This practice is known as English as a medium of instruction (EMI) (Macaro et al., 2018). Three core elements – (1) English-medium instruction in (2) academic subjects other than the English language to (3) L2 users of English – are essential for EMI to take place. Other aspects are also discussed but will not be touched upon in this article for the sake of brevity (for more information, see Smith, 2023). However, it is important to note that in an EMI context English is not the subject being taught and that language development is not the primary goal (Paulsrud et al., 2021) of the teaching.

EMI is a relatively new area of research, with most publications on the subject having appeared in the last 15 years (Bolton et al., 2024). Although little attention was paid to it before the turn of the century, we cannot ignore it now owing to increasing student and teacher mobility. At first glance, EMI seems to have many advantages: an enhanced reputation for an education institution internationally; a larger number of international students; access to academic material published in English; opportunities for local students to gain future high-quality employment; and a higher level of English-language proficiency for the institution and/or the country in general (Macaro, 2024). However, some scholars characterize the phenomenon as a rapidly accelerating force, one being introduced possibly at the risk of destabilisation (Macaro, 2018), because there is not yet sufficient evidence that it promotes students' English proficiency or that learning through the medium of English does not compromise the acquisition of subject knowledge (Paulsrud et al., 2021). In addition, the question arises whether all students are capable of attending and learning in classes that are conducted in a language that is not their native language.

Consequently, the introduction and expansion of this recent phenomenon of EMI has linguistic and educational consequences that, remarkably, have long been ignored by higher education institutions and researchers. Instead of considering the proliferation of EMI programmes with a critical eye, according to Macaro (2024), too often there has been a persistent tendency to ignore or deny the issue. Introducing such programmes at universities, however, is a serious and complex process that presents many challenges and questions for both students and teachers (O'Dowd, 2018). According to Kirkpatrick (2017), without proper planning and preparation, EMI is doomed to failure; key stakeholders are therefore concerned about its hasty implementation (Macaro, 2018).

These concerns appear to be justified, because several studies have shown that EMI programmes, instead of having a positive impact on student learning outcomes, have had a negative effect (Hellekjaer, 2010; Al Zumor, 2019; Çağatay, 2019; de Vos et al., 2020; Gabriëls & Wilkinson, 2024; Macaro, 2024). These programmes could therefore possibly create new barriers for local students who lack sufficient English proficiency to follow a lecturer's exposition properly. In this regard, Macaro (2018) points out that the concept of "English language" is problematic: Which variety of English is being referred to in this context? British English, American English or some other variant such as Nigerian or Indian English, Australian or even South African English, or a mixture of them (Macaro, 2018)? Whatever variant may be used, it could create a language barrier for the L2 users of English. One of the aims of EMI programmes is to remove linguistic barriers or obstacles, primarily for foreign students who would otherwise have to attend lectures conducted in a local language with which they are unfamiliar. However, the removal of linguistic barriers for foreign students by introducing EMI might lead to the creation of new linguistic barriers for local students (Milligan, 2020; Sah & Li, 2020; Sung, 2020; Tsou & Baker, 2021; Wijesekera & Hamid, 2022; Machin et al., 2023). In addition, EMI could also serve as a barrier to some of the foreign students, since most of them are also L2 speakers of English (Galloway et al., 2020; Kuteeva, 2020).

To counterbalance the language barrier posed by EMI, Nachtrab and Mössner (2017) have suggested that lectures be accompanied by subtitles, either in the same language or in a different language (respectively providing intralingual and interlingual subtitling). The intralingual live subtitling of lectures – that is, providing translations in real-time, is an accessibility service already offered in, for example, Germany and Switzerland for students with disabilities such as deafness and different levels of hearing loss (Nachtrab & Mössner, 2017). If live subtitling is able to remove sensory barriers for deaf and hard-of-hearing students, it might also be a possible solution to removing linguistic barriers for students whose proficiency in understanding spoken English is not optimal, enabling them to comprehend an EMI class.

Very few studies exist on the benefits of intralingual live subtitling to comprehending an event or a lecture¹. Robert et al. (2021), for example, established that students performed significantly better when provided with live subtitles than without them. In a similar study, Van Gauwbergen et al. (forthcoming) obtained the same result. In both studies, subtitles were produced by a respeakerⁱ (see Lambourne, 2006 for a reference to the tools for providing live text). As for the reception of live subtitling in an educational context, there have been a few studies, but they have a narrow focus: they are about subtitles in English in the United States produced using automatic speech recognition (ASR). According to these studies, the live subtitles were of poor quality and the students perceived them negatively (e.g., as distracting) (Ryba et al., 2006). Furthermore, the quality of the subtitles generated was such that a significant amount of post-editing was required to render them suitable for use as lecture notes (Bain et al., 2002; Chan et al., 2019).

In contrast, many studies have been conducted on the effect of “prepared” (i.e., not live) subtitles in SLA (Second Language Acquisition) and on the comprehension of films and/or video as educational materials (e.g., Kruger & Steyn, 2013; Kruger et al., 2014; Liao et al., 2020) and these studies have generally drawn on the Multimedia Learning Theory of Mayer (e.g., 2014, 2021). Mayer assumes that the human mind is a dual-channel, limited-capacity and active-processing system (see also Paivio’s [1971] Dual Coding Theory). Mayer also argues that multimedia learning material has to be developed taking all three types of cognitive load (CL) into account, that is, intrinsic CL (effort required to represent the material in working memory, based on the complexity of learning material); extraneous CL (cognitive effort wasted on materials that do not support learning); and germane CL (effort required to understand the material, which is strongly affected by motivation).

Mayer has formulated a series of principles to minimize extraneous CL, manage intrinsic CL and optimize germane CL. For example, based on research with native-speakers, the modality principle is that people learn better from graphics together with spoken words than from graphics together with printed words (Mayer, 2009; Mayer & Pilegard, 2014). Also, according to the redundancy principle people learn better from the combination of graphics and spoken text than from that of graphics, spoken text and printed text (Mayer, 2009; Mayer & Fiorella, 2014). More recently, Mayer et al. (2020) have included the “subtitle principle” among the principles of multimedia instruction. This principle states that “people learn better from a video documentary in their second language when the words are printed (or printed and spoken) rather than spoken” (2020, p. 847). In other words, the subtitle principle reverses both the modality and the redundancy principle:

students performed better on a comprehension post-test if they viewed a video with printed words rather than a video with spoken words (i.e., reverse modality effect) or

¹ However, there have been some studies on the reception of intralingual live subtitling of the TV news (see Eugeni, 2008 for a wider overview).

a video with printed and spoken text rather than a video with spoken text alone (i.e., reverse redundancy effect) (Mayer et al, 2020, p. 848).

The reason for this is that spoken words are “transient whereas printed words can be revisited” (Mayer et al., 2020, p. 848). To express it differently, when students are presented with spoken words in a second language, they may find it difficult to perceive or identify them and could need to revisit them. In this case, as Mayer et al. (2020) state, printed words (i.e., subtitles) are more helpful because they are available for a longer duration (reverse modality). The reverse redundancy effect means that, in a foreign language, students benefit more from a video with printed (i.e., subtitles) *and* spoken text rather than a video with spoken text alone. However, the authors also state that adding subtitles is useful provided that the pace of a lesson is slow enough not to overload the students' working memory.

Finally, even more recently, in his multimedia learning model Mayer (2021) points to the need to incorporate the role of motivation more effectively (which affects germane CL) and metacognition. Metacognition is often referred to as the awareness and understanding of one's own thinking process (Anthony, 2021, p. 6883). Students who have metacognitive strategies can assess the way in which to apply a particular strategy to a specific task and when to do so. They are accordingly aware of their own learning and thinking processes, can adjust them as necessary and can evaluate them later. In this way, they will learn more efficiently (Anthony, 2021, p. 6883). Controlling cognitive processes is very important in multimedia learning. Indeed, owing to the limitations of working memory (the limited-capacity assumption), learners are obliged to choose the incoming information that they pay attention to and the extent to which they will have to try to connect that information to their existing knowledge in long-term memory.

In SLA, intralingual (and interlingual) subtitles have been shown to reduce cognitive load, thanks to the visual support they provide (Montero Perez, 2022). In other words, subtitles affect learning positively, which is in line with the modality principle, namely, that combining images with verbally similar information improves information processing. For example, performance studies in SLA have shown that subtitling holds significant potential for vocabulary acquisition, listening comprehension and content comprehension (for an overview, see Incalcaterra McLoughlin, 2018). In contrast, the reception and, in particular, the beneficial educational effects of intralingual subtitles have not been studied much outside of SLA. There are a few exceptions that focus on subtitled *video* lectures but not in a live setting. For example, Van Hoecke (2023), in one of his studies, found that subtitles supported the modality principle instead of the redundancy principle. Similarly, Kruger-Marais (2024) compared seven student participants in the Faculty of Natural and Agricultural Sciences at the University of Pretoria and concluded that the participants remembered certain concepts from the videos more effectively when their focus was on the subtitles themselves. Finally, Malakul and Park (2023) compared 79 Thai secondary school students and found that the auto-subtitles system which generates subtitles for English educational videos has greater potential to facilitate online learning when compared to edited subtitles. Therefore, the subtitles generated by the auto-subtitles system in English educational videos can facilitate students' learning comprehension, cognitive load and satisfaction.

2. Research questions and methodology

Drawing on the research described in the previous section, the aim of this study was to investigate the effect of intralingual live subtitles used in EMI lectures on the perception and performance (i.e., comprehension and memory) of Flemish students. Our main research questions were:

- **RQ1:** How do students perceive an EMI lecture with intralingual live subtitles?
- **RQ2A:** Do students perform better when intralingual live subtitles are provided?
- **RQ2B:** Does the subtitle production method influence the students' performance?
- **RQ3:** Does performance vary with the students' level of English proficiency, their academic motivation and their perception of the task?

2.1. Experimental design

To answer these research questions, we examined six two-hour EMI Research Skills lectures which were attended by students of Applied Linguistics at the University of Antwerp who have Dutch as their mother tongue. The study was conducted in the second semester of the second year of the BA in Applied Linguistics during the academic year 2021–2022. This was the first time that an EMI course was included in the curriculum that had until then been taught exclusively in Dutch (except for the language courses). The first three lectures were given by a teacher who is a native English-speaker, whereas the last three were taught by a Belgian teacher whose first language is Dutch. Both lecturers are professors with experience in Research Skills. This course was taken by 59 students; however, complete quantitative data could be collected for only 32 of them. Of these 32 students, 19 were studying English as a foreign language in the BA programme, 13 students were not.

Approval from the EASHW (Ethics Committee for the Social Sciences and Humanities, SHW_21_153) was obtained before the start of the study. Table 1 provides an overview of the experimental design used.

Additional surveys	Lecture 1	Lecture 2	Lecture 3	Lecture 4	Lecture 5	Lecture 6	Duration per lecture (min)
1: Consent form and demographics	Respeaking	No subtitles	Respeaking	ASR	No subtitles	No subtitles	25
	No subtitles	ASR	No subtitles	No subtitles	Respeaking	ASR	25
2: Online English proficiency tests	<i>Break</i>	<i>Break</i>	<i>Break</i>	<i>Break</i>	<i>Break</i>	<i>Break</i>	10
	Respeaking	No subtitles	Respeaking	ASR	No subtitles	No subtitles	25
3: Academic motivation scale	No subtitles	ASR	No subtitles	No subtitles	Respeaking	ASR	25
	<i>Online survey</i>	<i>Online survey</i>	<i>Online survey</i>	<i>Online survey</i>	<i>Online survey</i>	<i>Online survey</i>	10

Table 1. Experimental design

Two weeks before the first lecture with live subtitling, the students who were willing to participate in the study filled in a first online questionnaire that consisted of a consent form and a demographics section. They also took both an online English proficiency test and a motivation test (see “Additional surveys” in Table 1). Each of the six lectures consisted of two parts totalling about 50 minutes, with a break of 10 minutes between the parts. In each lecture, half of part 1 and half of part 2 were given with subtitles (about 25 minutes; see Table 1). The subtitles were alternately produced through respeaking and automatic speech recognition (ASR), but the two different production methods were not used during the same lecture.

The live subtitling produced through respeaking was provided by an experienced respeaker. For this project, she made use of the most up-to-date version of Dragon Professional Individual, a speech-recognition program. Using a Text-on-Top wireless captioning kit, the lecturer

projected their PowerPoint presentation together with the subtitles (scrolling and word per word, over two lines). She worked onsite with a microphone mask, sitting in the same room as the students and listening to the lecture directly and not through an additional audio device, since the lecture room used could not be changed. Before each corresponding lecture, the respeaker was given the relevant PowerPoint presentation so that she could familiarize herself with the terminology and prepare Dragon Professional Individual for the assignment. The subtitles provided via ASR were produced using the built-in feature in Microsoft PowerPoint. Throughout all the lectures, subtitled and unsubtitled sections alternated in a counterbalanced order so as to minimize the impact of their ordering (i.e., starting or ending with subtitles). At the end of each lecture, the students completed a perception and performance online questionnaire (QualtricsXM), using their laptops or mobile phones (see section 2.2 for the content and focus of the questionnaire).

2.2. Material

2.2.1. Demographics, English proficiency and academic motivation

The demographics questionnaire consisted of six questions: gender, age, Dutch as L1 or L2, self-reported proficiency in English, whether or not they were taking the course for the first time and whether or not they had some kind of hearing impairment. Of the 32 students, 6 were male and 26 were female; 13 students did not have English in their curriculum, but 19 did. Their mean age was 20.03 years; 78.1% reported having Dutch as their mother tongue (L1), whereas the remaining 21.9% claimed to be bilingual (i.e., Dutch combined with Portuguese, Polish, Chinese, Spanish, Amharic, Cantonese or Albanian). All of the students were taking the course for the first time. Regarding their English proficiency, the students were asked to self-assess their English-language proficiency for Interaction, Reading, Listening, Speaking and Writing on a scale from 1 (A1) to 6 (C2). This scale corresponds to the six different proficiency levels of the Common European Framework of Reference for Languages (Council of Europe, 2001). Their self-reported proficiency was between the B1 and B2 levels for speaking and between the B2 and C1 levels for interaction, reading, listening comprehension and writing.

These students were also asked to take two certified language tests. First, to assess their listening proficiency by means of a listening comprehension test, we used the listening part of the 50-minute Education First Standardized English Test (EFSET (EF)). Second, to assess their linguistic proficiency in general, we used the 3000, 4000 and 5000 levels of the updated Vocabulary Levels Test (VLT) (Schmitt et al., 2001), from which we calculated the sum of the three scores. The VLT is a measure of vocabulary size (Van Hoecke, 2023).

The scores for the listening test and the vocabulary test are shown in Table 2. The two averages are very high (72.94/100 and 71.28/90), which means that the students estimated their abilities highly in their self-assessment. There is, as could be expected, little variation in their scores, since the participants were all language students.

Measure	Listening comprehension (score out of 100)	Vocabulary comprehension (score out of 90)
Mean	72.94	71.28
Median	74.00	70.00
Standard deviation	10.44	10.27
Minimum	43.00	56.00
Maximum	100.00	90.00

Table 2. Students' mean scores for listening and vocabulary tests

However, there is a significant variation between the two student groups (i.e., those studying English versus those not studying English). This variation is indicated by the independent samples t-test ($t(31)=-2.132$, $p=.041$)ⁱⁱ, which favours the students studying English: those studying English averaged 74.31 for the vocabulary test and 74.84 for the listening test, whereas those not studying English averaged 66.85 for the vocabulary test and 70.15 for the listening test.

Since motivation can play a role in comprehension and learning in multimedia learning (see section 1), the students were also asked to take an academic motivation scale test (Kotera et al., 2021), which is designed to measure both intrinsic motivation and extrinsic motivation. Intrinsic motivation is linked to the desire to do something, whereas extrinsic motivation is linked instead to doing something in order to receive a reward. Table 3 indicates that the participants were relatively motivated – mainly extrinsically (average of 4.03 for those without English versus average of 3.96 for those with English) – since this result is higher than that for intrinsic motivation (average of 3.23 for those without English versus average of 3.26 for those with English). This implies, for example, that students go to university in order to obtain a more prestigious job later on or to have a better salary rather than, for example, deriving satisfaction or pleasure from studying.

Measure	Intrinsic academic motivation (score out of 5)	Extrinsic academic motivation (score out of 5)
Mean	3.25	3.99
Median	3.17	4.00
Standard deviation	0.65	0.62
Minimum	1.8	2.3
Maximum	5.00	5.00

Table 3. Students' intrinsic and extrinsic academic motivation

2.2.2. Content of lectures

The six lectures of the Research Skills course comprised the topics listed in Table 4. As explained above, the students were invited to take part in the experiment, which indicated a concrete and live example of the topics covered in the course, in particular data-collection methods.

Lecture	Topic
1	Research: Orientation and preparation
2	Finding and evaluating materials (Part 1)
3	Finding and evaluating materials (Part 2)
4	Terminological clarification and data-collection methods (Part 1)
5	Data-collection methods (Part 2)
6	Data-collection methods (Part 3) and quantitative data analysis

Table 4. Topics of the lectures

In the present study, we ensured that, for each lecture, the duration of the unsubtitled sections was similar to that of the subtitled sections. Furthermore, we made sure that the subtitled and unsubtitled sections of each lecture formed a cohesive, well-rounded unit of content. To achieve this, the lecturers ensured that the two sections covered two self-contained units of content.

The performance of the participants was assessed by their having to respond to 16 questions per lecture (eight for the subtitled condition and eight for the unsubtitled condition). We ensured

that each question followed the same format, consisting of five multiple-choice options with only one correct answer each. In the performance questionnaire, we aimed to integrate an equal number of questions about the content that had been covered in the subtitled sections and in the unsubtitled sections.

In addition, the students responded to a series of questions regarding perception. These questions focused on the students' perceptions of the class, that is, their perceived reading behaviour, the perceived helpfulness of the subtitles, the perceived quality of the subtitles, the perceived cognitive load and their perceived level of difficulty of each lecture. These questions about perception were based on previous research, including that by Ryba et al. (2006), Romero-Fresco (2009, 2011), Leppink and van den Heuvel (2015), Perego (2016), Di Giovanni (2018) and Robert et al. (2021). The questions are presented in detail below for the purposes of clarity and to indicate why the figures in section 3.1 will not attain the same value on the y-axis.

1. Perceived reading behaviour

How often did you look at the subtitles in today's lecture? (Likert scale from 1 to 5)

- ☐ Never
- ☐ Rarely
- ☐ Sometimes
- ☐ Often
- ☐ Always

2. Perceived helpfulness of the subtitles

The subtitles helped me to understand the class better. (Likert scale from 1 to 5)

- ☐ I completely disagree
- ☐ I disagree
- ☐ Neutral
- ☐ I agree
- ☐ I completely agree

3. Perceived quality of the subtitles

The content of the subtitles was most of the time: (Likert scale from 1 to 5)

- ☐ Incomplete
- ☐ Rather incomplete
- ☐ Sufficient
- ☐ Rather complete
- ☐ Complete

4. Perceived cognitive load

Perceived intrinsic cognitive load: (Likert scalebar from 0 to 10)

- ☐ The content of this lecture was very complex.
- ☐ The problem/s covered in this lecture was/were very complex.
- ☐ In this lecture, very complex terms were mentioned.
- ☐ I invested a very high mental effort in the complexity of this lecture.

Perceived extraneous cognitive load: (Likert scalebar from 0 to 10)

- o The explanations and instructions in this lecture were very unclear.
- o The explanations and instructions in this lecture were full of unclear language.
- o The explanations and instructions in this lecture were, in terms of learning, very ineffective.
- o I invested a very high mental effort in unclear and ineffective explanations and instructions in this lecture.

5. Perceived difficulty of the lecture

The content of today's lecture was: (Likert scale from 1 to 5)

- o Easy
- o Relatively easy
- o Neutral
- o Rather difficult
- o Difficult

3. Results and discussion

In this section, we report on and discuss the results to answer our research questions:

- **RQ1:** How do students perceive an EMI lecture with intralingual live subtitles?
- **RQ2A:** Do students perform better when intralingual live subtitles are provided?
- **RQ2B:** Does the subtitles production method influence the students' performance?
- **RQ3:** Does performance vary with the students' level of English proficiency, their academic motivation (see section 1) and their perception of the task?

3.1. Perception

To answer RQ1, perception was measured through a series of Likert-scale questions regarding (1) perceived reading behaviour, (2) perceived helpfulness of the subtitles, (3) perceived quality of the subtitles, (4) perceived cognitive load and (5) perceived level of difficulty of each lecture. The perception questions on a Likert scale of 5 are shown in Figure 1.

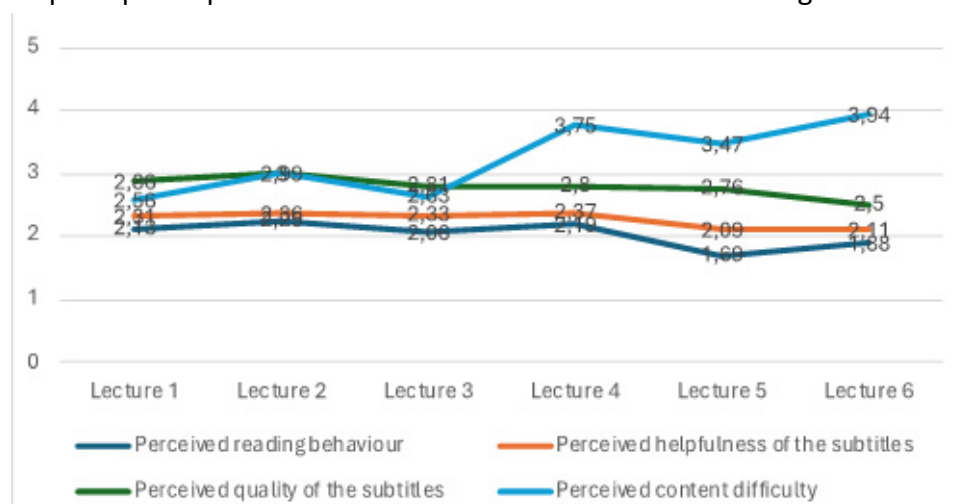


Figure 1. Perception questions on a Likert scale of 5

Regarding the perceived reading behaviour, the students indicated that they looked only rarely at the subtitles, as shown in Figure 1; towards the end of the lecture series, a tendency towards 'never' is observed. There was a significant difference between the lectures, as indicated by the

repeated ANOVA measures ($F=4.219$, $p=.001$)ⁱⁱⁱ. In summary, the students exhibited a largely indifferent stance towards the intralingual live subtitles, not looking much at the subtitles during the lectures (an average of 1.88/5 for students without English vs an average of 2.13/5 for those with English).

The perceived helpfulness is relatively low, too, and it remains constant throughout all the lectures. There was no significant difference between the lectures, as shown by the repeated ANOVA measures ($F=3.400$, $p=.853$). Therefore there is no perceived (substantial) difference between the subtitles through respeaking and the subtitles through ASR. In summary, there is largely no difference in the students' perception of the overall helpfulness of the subtitles (an average of 2.29/5 for students without English vs an average of 2.81/5 for those with English).

Moreover, the perception of quality is relatively neutral, but it reduces towards the end of the series. There was a significant difference across the lectures, as shown by the repeated ANOVA measures ($F=4.370$, $p<.001$). There was no significant difference regarding the perceived helpfulness of the subtitles, which does not point to a disturbing effect of the subtitles (an average of 3.23/5 for students without English vs an average of 3.33/5 for those with English).

Finally, regarding the perceived content difficulty, at first sight a clear difference between the classes of the two teachers can be observed (classes 1–3 by an English native-speaker and classes 4–6 by a non-native-speaker); this is confirmed by the repeated ANOVA measures ($F=20.979$, $p<.001$). In other words, the perceived content difficulty was not constant throughout the lecture series. In the beginning, the level of difficulty was considered 'neutral'; but towards the end it switched to 'difficult' (an average of 3.19/5 for students without English vs an average of 3.25/5 for those with English). Stated differently, the perceived level of difficulty of the content increased towards the end of the series of lectures, which is in line with the slight increase in the intrinsic cognitive load from lecture 4 (see Figure 2).

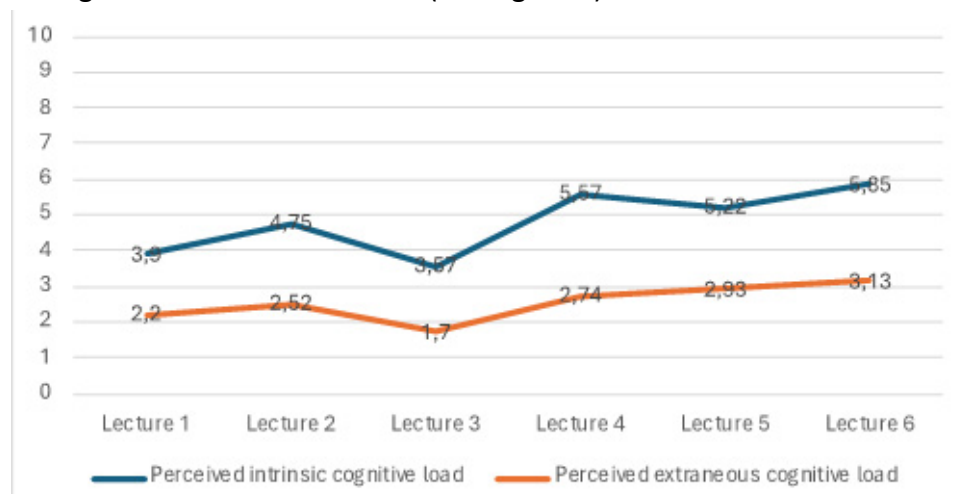


Figure 2. Perception questions on a Likert scale of 10

The perceived intrinsic CL fluctuated between approximately 4 and 6 on a scale of 10, with an increase from lecture 4, where teacher 2 began lecturing. The differences between the lectures are significant, as confirmed by the repeated ANOVA measures ($F=11.217$, $p<.001$). These results are in line with the perceived content difficulty, which makes sense, as intrinsic CL is linked to the content of a lecture. All in all, the students were rather neutral with respect to the perceived intrinsic CL (an average of 4.97/10 for students without English vs an average of 4.70 for those with English).

The extraneous load, in contrast, was perceived as being “low” in general and lower than the

intrinsic CL. The extraneous CL is linked to the way in which a lecture is given. At first sight, there seem to be no major differences across the lectures, which is an encouraging indication; but the repeated ANOVA measures ($F=5.496$, $p<.001$) indicate the contrary. However, the students' perceived extraneous CL was definitively lower than their perceived intrinsic CL, which does create a reassuring impression of the effectiveness of the lecture delivery method (an average of 2.47/10 for students without English vs an average of 2.58/10 for those with English).

3.2. Performance

Before answering research questions 2A (regarding the performance) and 2B (regarding the subtitle production method), we present some descriptive results, namely, the performance scores for the questions on different aspects of the content of the lectures and also per condition (i.e., either with or without subtitles), based on 32 students, as shown in Figure 3.

Overall, at first sight, we see no effect of the subtitles, or at least not a positive one: the students seem to perform better without subtitles, except in lecture 2. In all the other lectures, the students seem to perform worse with subtitles.

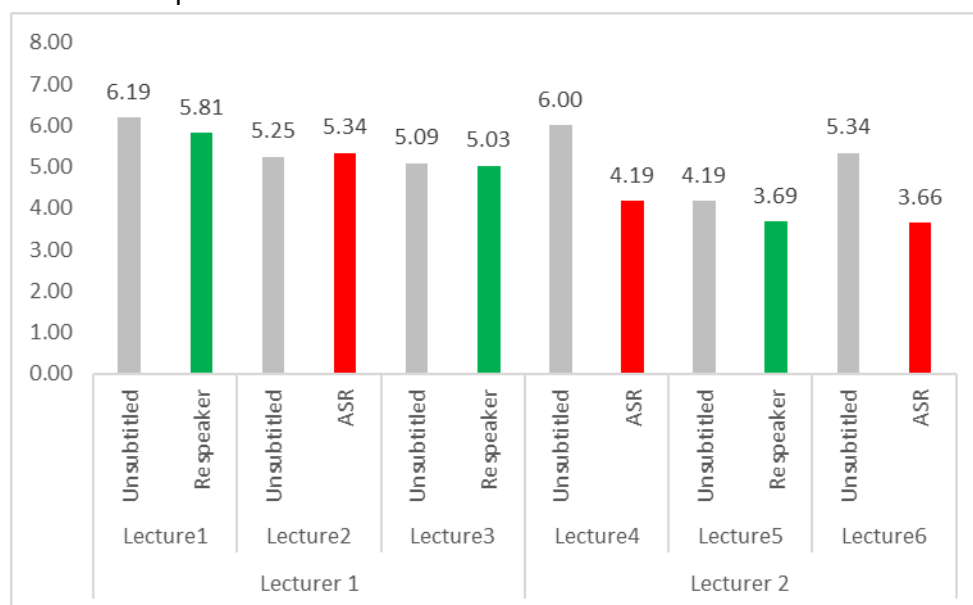


Figure 3. Performance scores per lecture and condition

In the first three lectures (lecturer 1, English native-speaker), the difference between the conditions seems limited, but in the last three (lecturer 2, non-native-speaker), it switches, and definitely in lectures 4 and 6, where the students seem to perform worse with subtitles. In addition, the figure shows that the discrepancies between the two types of subtitle production method are higher in lectures 4 and 6, where ASR was used.

To answer RQ2A (Do students perform better when intralingual live subtitles are provided?) and RQ2B (Does the subtitles production method influence the students' performance?), we started with a first Linear Mixed Model (Model 1) with the performance scores as dependent variable and the condition (subtitling by respeaker, subtitling by ASR, and no subtitles) as a fixed effect. We added the participant IDs as a random effect, since the performance scores are derived from the same participants. We also added the lectures and the teacher as random effects. This first model, namely, that with five parameters (performance score, condition, ID, lecture and teacher), had a $-2 \text{ Log Likelihood}$ of 1388.85^{iv}. The results show that the condition significantly predicted the performance score: $F(2,257.50) = 18.71$, $p < .001$. The estimates

of the fixed effects are shown in Table 5. It indicates that, compared to the scores in the ASR condition (intercept), those in the respeaking condition are significantly higher (+0,550) and that those in the unsubtitled condition are even significantly higher than those in the two subtitled conditions (+0,999).

Estimates of fixed effects ^a							
Parameter	Estimate	Std error	df	T	Sig.	95% confidence interval lower bound	95% confidence interval upper bound
Intercept	4.345	.151	381	28.828	<.001	4.049	4.641
Condition: Unsubtitled	.999	.168	253.653	5.958	<.001	.669	1.329
Condition: Respeaking	.550	.209	357.644	2.630	.009	.139	.961
Condition: ASR	0 ^b	0	.	.	.	.	.

^a Dependent variable: score.

^b This parameter is set to zero because it is redundant.

Table 5. Estimates of fixed effects (Model 1)

In other words, the participants performed significantly better in the unsubtitled condition than in the two subtitled conditions (RQ2A). This is not in line with the findings of Robert et al. (2021) or with those of other earlier research by Kruger and Steyn (2013), Kruger et al. (2014) and Liao et al. (2020), as our results suggest a redundancy effect rather than a modality effect (see section 1). Moreover, the participants performed less badly in the subtitled condition through respeaking than through ASR: their performance with respeaking is 0.55 higher than their performance with ASR (RQ2B).

We then controlled for the potential effect of the participants' studying English or not. In Model 2, we added the group to which students belonged, that is, those studying English versus those who do not. The model was not significantly better (-2 Log Likelihood of 1388.10; the difference in -2 Log Likelihood was 0.75, whereas the minimal difference with 1 degree of freedom [six parameters in Model 1 compared to five in Model 2] is 3.84 at $p=.05$). In other words, our better model before including additional variables to answer RQ3 was Model 1.

To answer RQ3 (Does performance vary with the students' level of English proficiency, their academic motivation – see section 1 – and their perception of the task?), we aimed to improve the model further by adding all the variables that are relevant to the research questions. Since we wanted to determine whether performance varies with the students' level of proficiency in English, their academic motivation and their perception of the task (RQ3), we added these variables one by one to determine whether the model was significantly better and, if so, whether the added variable was a significant predictor of the dependent variable (performance).

The responses to Model 3, including the level of English proficiency as measured through the listening test, were significantly better than those of Model 1. That is, the added variable was a significant predictor of performance: $F(1,192)=5.28$, $p=.023$. The -2 Log Likelihood was 1383.64; the difference in -2 Log Likelihood was 5.21, whereas the minimal difference with 1 degree of freedom (five parameters in Model 1 compared to six in Model 3) is 3.84 at $p=.05$. Model 3 therefore passed the test.

The responses to Model 4, including the level of English proficiency as measured by the vocabulary test, were not significantly better than those of Model 3 (-2 Log Likelihood of

1386.19); the added variable was therefore not a significant predictor of performance. Accordingly, English proficiency as measured through a listening test did have an impact on performance, which is in line with the findings in the literature (see Hellekjaer, 2010; Al Zumor, 2019; Çağatay, 2019; de Vos et al., 2020; Gabriëls & Wilkinson, 2024; Macaro, 2024). This finding suggests that language proficiency has an impact on performance. In other words, better comprehension facilitates learning.

The responses to Model 5A, including the level of intrinsic academic motivation, were not significantly better than those in Model 3 (-2 Log Likelihood = 1383.63), the added variable not being a significant predictor of performance either. The responses to Model 5B, including the level of extrinsic academic motivation, were also not significantly better than those in Model 3 (-2 Log Likelihood = 1383.64); nor was the added variable a significant predictor of performance. Although the literature suggests that academic motivation has an impact on performance (Mayer, 2021), this was not the case here, perhaps due to the format of the questionnaire (see the Conclusion), since the questions might have focused too generally on the academic world instead of on the studying experience of the students specifically.

In Model 6A we added the perceived reading behaviour as an additional fixed effect. The responses to Model 6A were not better than those in Model 3: the -2 Log Likelihood was 1383.24 and the difference is 0.40, which is less than the minimal difference with 1 degree of freedom (six parameters in Model 3 compared to seven in model 6A) of 3.84 at $p=.05$. In Model 6B we added the perceived usefulness of the subtitles as an additional fixed effect. The responses to Model 6B were also not better than those in Model 3 (-2 Log Likelihood of 1383.61). In Model 6C we added the perceived quality of the subtitles as an additional fixed effect. The responses to Model 6C, too, were not better than those in Model 3 (-2 Log Likelihood of 1382.91, i.e., a small difference between the two models).

In Model 6D we added the intrinsic CL as an additional fixed effect. In this case, the responses were significantly better than those in Model 3: the -2 Log Likelihood was 1358.92, a difference of 24.72, which is more than the minimal difference with 1 degree of freedom (six parameters in Model 3 compared to seven in Model 6D) of 3.84 at $p=.05$. However, in the case of this model, the effect of the listening comprehension test was no longer significant. Consequently, in Model 6E, we added the perceived difficulty of the lecture as an additional fixed effect *instead of* the intrinsic cognitive load, since both variables should measure the same even though the scales were different. The responses to Model 6E were also better than those in Model 3: the -2 Log Likelihood was 1353.95, a difference of 29.69, which is more than the minimal difference with 1 degree of freedom (six parameters in Model 3 compared to seven in Model 6E) of 3.84 at $p=.05$. In other words, the level of difficulty of a lecture was a significant predictor of performance ($F(1,199.94)=32.88, p<.001$), with a significant negative effect being indicated of perceived difficulty on performance. Consequently, we decided to retain Model 6E at this stage.

In the next model, Model 6F, we added the perceived extraneous load as an additional fixed effect. The responses to Model 6F were not better than those of Model 6E: the -2 Log Likelihood was 1350.17, the difference 3.78, which is less than the minimal difference with 1 degree of freedom (seven parameters in Model 6E compared to eight parameters in Model 6F) of 3.84 at $p=.05$.

Finally, we added interactions to Model 6E and created Models 7 and 8 – Model 8 becoming the best model as measured by Log Likelihood. In Model 7, we added the interaction between the condition and the listening comprehension test, but the responses to this model were not better (-2 Log Likelihood of 1353.71). The findings in Model 8, including an interaction between the condition and the perceived content difficulty, were better (-2 Log Likelihood

of 1344.26, i.e. a difference of 9.69, which is more than the minimal difference with two degrees of freedom [seven parameters in Model 6E compared to nine parameters in Model 8] of 5.99 at $p=.05$). In this model, listening comprehension was not a significant predictor, whereas the interaction between condition and perceived content difficulty was significant. The finding that perceived content difficulty interacts significantly with condition, while listening comprehension does not, suggests that task difficulty plays a more central role in shaping performance than baseline listening ability does. In other words, when the perceived content difficulty increases, the performance decreases in all conditions, but more so in the ASR condition than in the respeaking condition, and more so in both subtitled conditions than in the unsubtitled condition. Consequently, the ASR condition seems to be more sensitive to perceived content difficulty than the respeaking condition.

This result is interesting and supports the idea that ASR subtitles may add to the cognitive load under more demanding listening conditions. However, while this interpretation is plausible, possible confounds could also have had an impact: variation in lecturer or content structure across lectures, for instance. In addition, the limitations of the small sample size and possible residual imbalances should also be considered (see also 'limitations' in the Conclusion). Estimates of the fixed effects are shown in Table 6.

Estimates of fixed effects ^a							
Parameter	Estimate	Std error	df	T	Sig.	95% confidence interval lower bound	95% confidence interval upper bound
Intercept	6.632	.851	314.463	7.796	<.001	4.958	8.305
[Condition_3=,00]	-1.145	.652	269.886	-1.755	.080	-2.429	.140
[Condition_3=1,00]	-1.315	.793	369.162	-1.658	.098	-2.875	.245
[Condition_3=2,00]	0 ^b	0	—	—	—	—	—
Eng_Prof100_Listening	.013	.008	186.233	1.691	.092	-.002	.028
Perceived_ContentDifficulty	-.909	.159	377	-5.715	<.001	-1.222	-.596
[Condition_3=,00] * Perceived_ContentDifficulty	.571	.181	253.117	3.148	.002	.214	.928
[Condition_3=1,00] * Perceived_ContentDifficulty	.437	.240	369.291	1.817	.070	-.036	.910
[Condition_3=2,00] * Perceived_ContentDifficulty	0 ^b	0	—	—	—	—	—

a Dependent variable: score.

b This parameter is set to zero because it is redundant.

Table 6. Estimates of fixed effects (Model 8)

In short, a possible reason for the lower performance in the subtitled condition could be the perceived difficulty of the lecture (which did not always have the same structure and may also have had an impact). As indicated before, the subtitles were produced differently. Moreover, in lectures 1–3 the lecturer was a native-speaker whereas in lectures 4–6 she was not. The perceived reading behaviour of the students towards the subtitles was also very low, which could also explain their lower performance in the subtitled condition: if students do not look at the subtitles, it is obvious that the subtitles cannot help them. The reason why they do not look at the subtitles might be their (perceived) language proficiency. Another conclusion could be that, if students read the subtitles, their attention is split, meaning that they focus less on what is heard because they are distracted by the subtitles; alternatively, the subtitles become

redundant due to the students' level of language proficiency. These are valid hypotheses that are supported by the relevant literature, including that of Mayer et al. (2020) and Van Hoecke (2023). The caveat seems to be that, although the participants in this study are not native speakers, they do not need the subtitles and are able to follow the lectures without difficulty. However, they do become distracted by the subtitles once the lecture content becomes more difficult, since the subtitles are redundant and therefore overload their working memory. This is the conclusion Mayer (2002) also reached, claiming that subtitles may overload working memory.

4. Conclusion

This study investigated the impact of intralingual live subtitles on both the perceptions and the performance of students (via a comprehension test) during six EMI lectures at a university in Flanders. As far as perception is concerned, although the students were satisfied with the quality of the subtitles and did not find them cognitively demanding, they regarded them as not particularly useful. Consequently, they did not look at them frequently, which suggests an indifferent attitude.

Regarding overall performance across all lectures, the students' performance with subtitles was lower than their performance without them. At first sight, listening comprehension seemed to influence their performance positively. Academic motivation, however, did not seem to have an impact; nor did perception, apart from the perceived level of difficulty of a lecture. A Linear Mixed Model analysis showed that the only significant predictor of performance was the interaction effect between condition (unsubtitled, subtitling via respeaking and subtitling via ASR) and the perceived level of difficulty of the lecture content. In other words, performance decreased as perceived content difficulty increased across all three conditions but with a stronger reduction in the ASR condition compared to the respeaking condition or the unsubtitled condition. The negative impact of subtitles on these students' performance does not align with the findings of previous research (Malakul & Park, 2023; Van Hoecke, 2023; Kruger-Marais, 2024). In other words, the present study does not support the reverse redundancy effect described by Mayer et al. (2020) but instead suggests a redundancy effect.

As is generally the case in empirical research, our research has some limitations. First, the academic motivation survey may not have been sufficiently adapted to the specific target audience, as the questions were fairly general rather than focused on the specific context of the surveyed students. Other possible limitations, such as the need to use different content in a within-group design, the small number of participants, and the reliance on a convenience sample, may also have had an impact. Moreover, the complexity of the design and the limited sample size may have reduced the generalizability of the findings. For example, to what extent can the students' perceptions be meaningfully interpreted, especially when trying to isolate the effects of different subtitling methods? Finally, the working conditions of the respeaker might also have had an impact on the quality of the subtitles, as they were not optimal for producing accurate subtitles.

Can it then be assumed that the English proficiency of Flemish university students is sufficient and that the introduction of an EMI programme will not create linguistic barriers for them? Such a conclusion may be premature. This experiment was conducted with language students whose English proficiency appears to be high enough to follow lectures given in English. Consequently, it is essential to replicate the experiment with students from other academic backgrounds whose English proficiency may be less developed.

5. References

- Al Zumor, A. (2019). Challenges of using EMI in teaching and learning of university scientific disciplines: Student voice. *International Journal of Social Sciences and Educational Studies*, 5(3), 1–17. <https://doi.org/10.23918/ijsses.v5i3p1>
- Anthonyssamy, L. (2021). The use of metacognitive strategies for undisrupted online learning: Preparing university students in the age of pandemic. *Education and Information Technologies*, 26, 6881–6899. <https://doi.org/10.1007/s10639-021-10518-y>
- Bain, K., Basson, S., & Wald, M. (2002). *Speech recognition in university classrooms*. Paper presented at the Proceedings of the ACM Conference on Assistive Technologies, ASSETS 2002, July 8–10, 2002, Edinburgh.
- Bolton, K., Botha, W., & Lin, B. (2024). English-medium instruction in higher education worldwide. In K. Bolton, W. Botha, & B. Lin (Eds.), *The Routledge Handbook of English-Medium Instruction in Higher Education* (pp. 1–17). Routledge. <https://doi.org/10.4324/9781003011644-1>
- Çağatay, S. (2019). The pros and cons of English-medium instruction in higher education in EFL context. *Journal of Foreign Language Education and Technology*, 4(2), 206–241.
- Chan, S. W., Kruger, J.-L., & Doherty, S. (2019). Comparing the impact of automatically generated and projected subtitles on cognitive load and learning in a first- and second-language educational context. *Linguistica Antverpiensia*, 18, 237–272. <https://doi.org/10.52034/lanstts.v18i0.506>
- de Vos, J. F., Schriefers, H., & Lemhöfer, K. (2020). Does study language (Dutch versus English) influence study success of Dutch and German students in the Netherlands? *Dutch Journal of Applied Linguistics*, 9(1/2), 60–78. <https://doi.org/10.1075/dujal.19008.dev>
- Di Giovanni, E. (2018). The reception of professional and non-professional subtitles: Agency, awareness, and change. *Cultus*, 11, 18–37.
- Eugeni, C. (2008). Respeaking the TV for the Deaf: For a real special needs-oriented subtitling. *Studies in English Language and Literature*, 21, 37–47.
- Gabriëls, R., & Wilkinson, R. (2024). English-medium instruction in higher education in the Netherlands. In K. Bolton, W. Botha, & B. Lin (Eds.), *The Routledge Handbook of English-Medium Instruction in Higher Education* (pp. 161–175). Routledge. <https://doi.org/10.4324/9781003011644-14>
- Galloway, N., Numajiri, T., & Rees, N. (2020). The “internationalisation”, or “Englishisation”, of higher education in East Asia. *Higher Education*, 80(3), 395–414. <https://doi.org/10.1007/s10734-019-00486-1>
- Hellekjaer, G. O. (2010). Lecture comprehension in English-Medium Higher Education. *HERMES - Journal of Language and Communication in Business*, 23(45), 11–34. <https://doi.org/10.7146/hjlc.v23i45.97343>
- Incalcaterra McLoughlin, L. (2018). Audiovisual translation in language teaching and learning. In L. Pérez-González (Ed.), *The Routledge Handbook of Audiovisual Translation* (pp. 483–515). Routledge. <https://doi.org/10.4324/9781315717166-30>
- Kirkpatrick, A. (2017). The languages of Higher Education in East and Southeast Asia: Will EMI lead to Englishisation? In B. Fenton-Smith, P. Humphreys, & I. Walkinshaw (Eds.), *English Medium Instruction in Higher Education in Asia-Pacific* (pp. 21–36). Springer. https://doi.org/10.1007/978-3-319-51976-0_2
- Kotera, Y., Conway, E., & Green, P. (2021). Construction and factorial validation of a short version of the Academic Motivation Scale. *British Journal of Guidance & Counselling*, 51(2), 274–283. <https://doi.org/10.1080/03069885.2021.1903387>
- Kruger, J.-L., Hefer, E., & Matthew, G. (2014). Attention distribution and cognitive load in a subtitled academic lecture: L1 vs. L2. *Journal of Eye Movement Research*, 7(5), 1–15. <https://doi.org/10.16910/jemr.7.5.4>
- Kruger, J. L., & Steyn, F. (2013). Subtitles and eye tracking: reading and performance. *Reading Research Quarterly*, 49(1), 105–120. <https://doi.org/10.1002/rrq.59>
- Kruger-Marais, E. (2024). Subtitling for language acquisition: Eye tracking as predictor of attention allocation in education. *International Journal of Language Studies*, 18(2), 129–150.
- Kuteeva, M. (2020). Revisiting the “E” in EMI: Students’ perceptions of standard English, lingua franca and translanguaging practices. *International Journal of Bilingual Education and Bilingualism*, 23(3), 287–300. <https://doi.org/10.1080/13670050.2019.1637395>
- Lambourne, A. (2006). *Subtitle respeaking: A new skill for a new age*. inTRAlinea Special Issue: Respeaking.
- Leppink, J., & van den Heuvel, A. (2015). The evolution of cognitive load theory and its application to medical education. *Perspectives on Medical Education*, 4, 119–127. <https://doi.org/10.1007/S40037-015-0192-X>
- Liao, S., Kruger, J.-L., & Doherty, S. (2020). The impact of monolingual and bilingual subtitles on visual attention, cognitive load, and comprehension. *JoSTrans*, 33, 70–98.
- Macaro, E. (2018). *English Medium Instruction*. Oxford University Press. <https://doi.org/10.30687/978-88-6969-227-7/001>
- Macaro, E., Curle, S., Pun, J., An, J., & Dearden, J. (2018). A systematic review of English medium instruction in higher education. *Language Teaching*, 51(1), 36–76.

- Macaro, E. (2024). English-medium instruction: Theoretical and applied perspectives. In K. Bolton, W. Botha, & B. Lin (Eds.), *The Routledge Handbook of English-Medium Instruction in Higher Education* (pp. 21–34). Routledge. <https://doi.org/10.4324/9781003011644-3>
- Machin, E., Ament, J., & Pérez-Vidal, C. (2023). Motivation and investment: Exploring the choice of English-medium instruction for mid-degree undergraduates in Catalonia. *Journal of English-Medium Instruction*, 2(1), 79–100. <https://doi.org/10.1075/jemi.22002.mac>
- Malakul, S., & Park, I. (2023). The effects of using an auto-subtitle system in educational videos to facilitate learning for secondary school students: learning comprehension, cognitive load, and satisfaction. *Smart Learning Environments*, 10(4), 1–17. <https://doi.org/10.1186/s40561-023-00224-2>
- Mayer, R. E. (2014). Introduction to Multimedia Learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 1–24). Cambridge University Press.
- Mayer, R. E. (2002). Multimedia learning. *Psychology of Learning and Motivation*, 41, 85–139. [https://doi.org/10.1016/S0079-7421\(02\)80005-6](https://doi.org/10.1016/S0079-7421(02)80005-6)
- Mayer, R. E. (2021). *Multimedia learning: Third edition*. Cambridge University Press.
- Mayer, R. E., Fiorella, L., & Stull, A. (2020). Five ways to increase the effectiveness of instructional video. *Educational technology research and development*, 68, 837–852. <https://doi.org/10.1007/s11423-020-09749-6>
- Milligan, L. O. (2020). Towards a social and epistemic justice approach for exploring the injustices of English as a medium of instruction in basic education. *Educational Review*, 74(5), 927–941. <https://doi.org/10.1080/00131911.2020.1819204>
- Montero Perez, M. (2022). Second or foreign language learning through watching audio-visual input and the role of on-screen text. *Language Teaching*, 55, 163–192. <https://doi.org/10.1017/S0261444821000501>
- Nachtrab, M., & Mössner, S. (2017). Speech-to-text and online respeaking in Switzerland. In S. J. Jekat & G. Massey (Eds.), *Barrier-free Communication: Methods and Products* (pp. 16–22). Winterthur: ZHAW.
- O'Dowd, R. (2018). The training and accreditation of teachers for English medium instruction: an overview of practice in European universities. *International Journal of Bilingual Education and Bilingualism*, 21(5), 553–563. <https://doi.org/10.1080/13670050.2018.1491945>
- Paivio, A. (1971). *Imagery and verbal processes*, Holt, Rinehart, and Winston.
- Paulsrud, B., Tian, Z., & Toth, J. (2021). *English-Medium Instruction and Translanguaging*. Multilingual Matters. <https://doi.org/10.21832/9781788927338>
- Perego, E. (2016). Is subtitling equally effective everywhere? A first cross-national study on the reception of interlingually subtitled messages. *Across Languages and Cultures* 17, 203–227. <https://doi.org/10.1556/084.2016.17.2.4>
- Robert, I. S., De Meulder, A., & Schrijver, I. (2021). Live subtitling for access to education: A pilot study of university students' reception of intralingual live subtitles. *JoSTrans*, 36a, 53–78.
- Romero-Fresco, P. (2009). More haste less speed: Edited versus verbatim respoken subtitles. *VIAL*, 6, 109–133.
- Romero-Fresco, P. (2011). *Subtitling through speech recognition: Respeaking*. St Jerome.
- Ryba, K., Mclvor, T., Shakir, M., & Paez, D. (2006). Liberated learning: Analysis of university students' perceptions and experiences with continuous automated speech recognition. *E-Journal of Instructional Science and Technology*, 9(1), 1–19.
- Sah, P. K., & Li, G. (2020). Translanguaging or unequal languaging? Unfolding the plurilingual discourse of English medium instruction policy in Nepal's public schools. *International Journal of Bilingual Education and Bilingualism*, 25(6), 2075–2094. <https://doi.org/10.1080/13670050.2020.1849011>
- Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behaviour of two new versions of the Vocabulary Levels Test. *Language Testing*, 18(1), 55–58. <https://doi.org/10.1177/026553220101800103>
- Smit, U. (2023). English-Medium Instruction (EMI). *ELT Journal*, 77(4), 499–503. <https://doi.org/10.1093/elt/ccad018>
- Sung, C. C. M. (2020). Investing in English-mediated practices in the EMI university: The case of cross-border mainland Chinese students in Hong Kong. *Lingua*, 243, 1–17. <https://doi.org/10.1016/j.lingua.2020.102919>
- Tsou, W., & Baker, W. (2021). *English-medium instruction translanguaging practices in Asia: Theories, frameworks and implementation in higher education*. Springer. <https://doi.org/10.1007/978-981-16-3001-9>
- Van Hoecke, S. (2023). *Subtitles for access to education: The impact of subtitles, subtitle language and linguistic proficiency on cognitive load, comprehension, reading and processing in different styles of asynchronous, online university lectures* [Unpublished doctoral dissertation]. University of Antwerp.
- Wijesekera, H. D., & Hamid, M. O. (2022). The dynamics of bilingual education in post-conflict Sri Lanka. In L. Adinolfi, U. Bhattacharya, & P. Phyak (Eds.), *Multilingual education in South Asia: At the intersection of policy and practice* (pp. 16–38). Routledge. <https://doi.org/10.4324/9781003158660-2>

Websites

Council of Europe (2001). Self-assessment grid: Table 2 (CEFR 3.3): Common reference levels. <https://rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?docu mentId=090000168045bb52>
Education First. Standardized English listening test. <https://www.efset.org/ef-set-50/ta>

- ⁱ Different methods exist for producing these so-called (intralingual) live subtitles. Respeaking is a “technique in which a respeaker listens to the original sound of a (live) program or event and respeaks it, including punctuation marks [...], to a speech recognition software, which turns the recognized utterances into subtitles displayed on the screen with the shortest possible delay” (Romero-Fresco, 2011, p. 1). However, other production methods could also be used: (1) different variants of fast typing; (2) trained automatic speech recognition (t-ASR) of the speaker’s voice, thus without the intervention of a respeaker, but with speech recognition (SR) software trained with the speaker’s voice; and (3) fully automatic speech recognition (ASR) (i.e., without training of the speech recognition software).
- ⁱⁱ Based on the Shapiro-Wilk test for normality ($p=.708$) and Levene’s test for equality of variances ($p=.510$).
- ⁱⁱⁱ All of the tests (i.e., including the following) were in accordance with Mauchly’s test of sphericity ($p>.05$).
- ^{iv} This Log Likelihood is neither good nor bad; it simply serves to do the calculation for the next model(s).

Acknowledgements

We would like to thank the reviewers for their comments and suggestions, and John Linnegar for the proofreading.



This work is licensed under a Creative Commons Attribution 4.0 International License.