The Use of ASR-CAI Tools and their Impact on Interpreters' Performance during Simultaneous Interpretation

Pannapat Tammasrisawat
Chalermprakiat Center of Translation and Interpretation,
Chulalongkorn University, THAILAND

Nunghatai Rangponsumrit
Translation, Interpreting and Intercultural Communication Research Unit
Chulalongkorn University, THAILAND

ABSTRACT
This study aims to investigate the impact of Computer-Assisted Interpreting tools powered with Automatic Speech Recognition technology (ASR-CAI tools) on the simultaneous interpreting (SI) process and the interpreters' overall performance. We conducted an experiment with four trainee interpreters to observe how SI was performed with and without the support of an ASR-CAI tool. The results showed that the use of an ASR-CAI tool led to a significant reduction in the error rate and omissions and improved the quality of terminology rendition. The paper also discusses the impact of the ASR-CAI tool on interpreters' processing capacity and cognitive effort during SI. In addition, potential benefits and limitations of the ASR-CAI tool were analyzed to provide a better understanding regarding the tool's usability so that such tools can be better integrated into the interpreting process.

KEYWORDS: ASR-CAI tools, automatic speech recognition, cognitive effort, computer-assisted interpreting, simultaneous interpreting, terminology rendition

1. Introduction

Since the introduction of computer-assisted interpreting (CAI) technology in the early 2000s, many interpreters have made use of CAI tools to assist them during different phases of their

workflow, mainly in organizing terminological data during the preparation phase and in accessing glossaries in the actual interpreting phase in the booth. Early empirical analyses on the use of CAI tools during simultaneous interpretation (SI) suggest an improvement of terminology rendition during the interpreting process, but the main drawback remains in the way that the tools require manual operation when looking up terminology, which can potentially add to the interpreters' cognitive load during the interpreting process (Pisani and Fantinuoli 2021: 138). In recent years, automatic speech recognition (ASR) has been proposed as a means to improve the lookup mechanism for interpreters to reduce additional cognitive effort when performing a glossary search. Since then, several studies have been conducted to investigate the integration of ASR in CAI tools (Fantinuoli 2016, 2017a, 2017b; Desmet et al. 2018; Canali 2019, cited in Fantinuoli 2022), with InterpretBank being featured in most studies as a prototype of ASR-CAI integration, which provides an interpreter with real-time transcription of the speech delivered by a speaker and also provides suggestions regarding the translation of terminology, as well as numerals and their units of measurement. The real-time transcription provided by ASR offers the added advantage of serving as a written reference, assisting interpreters especially when confronted with numerical data.

So far, existing studies (Fantinuoli 2017b; Desmet et al. 2018; Canali 2019, cited in Fantinuoli 2022; Fantinuoli and Montecchio 2022) have focused on evaluating the integration of ASR in CAI tools by measuring the precision and recall scores for terminology and identification of numbers, but few studies (Prandi 2018; Pisani and Fantinuoli 2021) have evaluated the impact of ASR-CAI tools on the interpreting process. Most experiments (Fantinuoli 2017b; Desmet et al. 2018; Canali 2019, cited in Fantinuoli 2022; Fantinuoli and Montecchio 2022) centered on a product-based analysis –focusing on the transcription accuracy and latency provided by ASR, and/or the extent to which they may improve interpreters' rendition quality, particularly in accuracy in number rendition, but there are still a limited number of studies (Prandi 2018; Pisani and Fantinuoli 2021) on how the use of ASR-CAI tools may impact interpreters' processing capacity as well as their overall performance in SI. Although all the above-mentioned studies conducted on ASR-CAI tools suggest that ASR proves effective in providing interpreters support during the interpretation of speeches dense in numbers and terminology, few experiments (Prandi 2018; Pisani and Fantinuoli 2021) have implemented the process-oriented method in studying the usability of

ASR-CAI tools and how the use of such tools may affect the overall interpreting process, with particular emphasis on interpreters' cognitive load while interpreting.

Considering the research gaps mentioned above, this study set out to investigate the usability of an ASR-CAI tool and its impact on interpreters' overall performance during the simultaneous interpreting process, both from the product-based and the process-based perspectives. The aim of this research was threefold: 1) to investigate the differences in accuracy of SI performed with and without the support of the ASR-CAI tool and to test the usefulness of the functions provided by the tool, namely the real-time transcriptions and translation suggestions for terminology, 2) to explore interpreters' processing capacity and cognitive effort during SI performed with the ASR-CAI tool, and 3) to study interpreters' perception towards the usability of the ASR-CAI tool to see how the use of such tools affects the interpreting process.

We hope that further investigation to evaluate the impact of ASR-CAI tools on interpreters' performance and processing capacity will help bridge existing research gaps and provide useful insights on how ASR-CAI tools can be better integrated into the interpreting process. for enhanced interpreting quality (as described by Bühler 1986) and tools' usability.

2. CAI Tools with ASR Integration: The State of the Art

Nowadays, more and more interpreters have opted to use CAI tools to assist them during the preparation stage and to facilitate the glossary search process in the booth. Although one key feature of CAI tools lies in their ability to support interpreters in accessing terminology during SI, earlier studies (Fantinuoli 2017b; Pisani and Fantinuoli 2021) pointed out certain shortcomings of the use of CAI tools, with the main drawback being that the tools' database has to be queried manually. As a result, this can potentially add to interpreters' cognitive effort required in concurring tasks simultaneous interpreters have to perform in the booth such as listening, analysis, translation, speech production, monitoring, and memory retention. This added cognitive load can interrupt the overall interpreting process, since, as Gile (2009: 182) pointed out, interpreters work close to saturation most of the time.

Fantinuoli (2017b) proposed that this disadvantage could be addressed by automating the querying system through the use of ASR. He conducted a pilot study featuring a prototype for the integration of ASR into CAI tools to test its precision and recall of terminology retrieval and identification of numbers. The results confirmed that ASR could effectively provide support for users during the interpretation of speeches that were dense in numbers and terminology, which consequently led to a reduction of the error rate and omissions among the interpreters who participated in the study.

Proposed as a means to enhance state-of-the-art computer-assisted interpreting tools (e.g. InterpretBank), ASR made it possible for machine-learning techniques to be integrated into the workflow of professional interpreters. As the CAI tool reads the transcription provided by the ASR system and automatically presents the interpreter with numerals and entries from the terminology database, it acts like an electronic boothmate, providing useful information to the colleague whenever necessary (Fantinuoli 2017b: 27). Following Fantinuoli's proposition regarding the possibility of ASR integration, many scholars have tried to develop theoretical frameworks for empirical studies in the area of ASR-supported CAI tools, with most studies (Fantinuoli 2017b; Defrancq and Fantinuoli 2020; Pisani and Fantinuoli 2021) focusing on the tools' performance when dealing with problem triggers such as numbers and terminological data, and exploring the extent to which the tools can improve interpreters' rendition quality.

Defrancq and Fantinuoli (2020) reported on a small-scale experiment with in-booth CAI, which tested the usefulness of real-time transcriptions with numbers using InterpretBank ASR. The study indicated that the system's precision was high and its latency low enough to fit interpreters' ear–voice span (EVS). The results of the study showed that 96% of the numbers were displayed correctly and that the tool's precision was higher than interpreters' accuracy levels reported in experimental and corpus-based research (Braun and Clarici 1996; Lamberger-Felber 2001; Mazza 2001; Pinochi 2009; Timarová 2012; Korpal 2016; Desmet, Vandierendonck, and Defrancq 2018; Frittella 2019; Collard 2019). Defrancq and Fantinuoli concluded that ASR, therefore, had the potential to improve interpreters' accuracy in number rendition.

Another similar research conducted by Pisani and Fantinuoli (2021) adds on the findings of Defrancq and Fantinuoli by measuring the impact of ASR on number rendition in SI. In their experiment, Pisani and Fantinuoli used a real-life ASR-enhanced CAI tool in which typical issues of ASR such as latency and mistranscriptions were not eliminated. This allowed them to draw conclusions on the potential and limits of ASR technology. The experiment confirmed that ASR proved effective in providing interpreters support during interpretation of speeches dense in numbers. The support of ASR was also reported to help reduce omissions and approximations, and helped interpreters avoid phonetic perception errors.

The results of both studies confirmed the usefulness of ASR and allowed for a more complete picture of user-machine interaction in the context of real-time CAI support. However, both experiments focused on the product-based analysis of ASR-CAI tools and the quality of number rendition. Both studies mainly assessed participants' performance in terms of accuracy, with Pisani and Fantinuoli's study (2021) delving further into how the use of ASR-CAI tools may affect interpreters' cognitive capacity. However, they did not explore how the results may vary if the tool is presented with other kinds of information such as translation suggestions for specialized terminology.

An early attempt at analyzing the allocation of cognitive resources while working with CAI tools during SI can be identified in exploratory research by Bianca Prandi (2017). In her research, Prandi aimed to develop a research methodology through an exploratory study which implemented both product/process-based measures to investigate the local variations in cognitive load while interpreters performed a glossary query using a CAI tool in comparison to electronic glossaries. Prandi pointed out that, when working with CAI tools, interpreters were expected to perform fewer manual-spatial and visual-spatial sub-tasks since they only had to type in and visually locate the term needed. In comparison, electronic glossaries such as Word or Excel tables would require interpreters to position the cursor in the search field, type and press the enter button, scroll up and down or press the "forward" button to locate the term needed, and delete the term before starting a new search (2017:83).

The results from Prandi's research showed that CAI tools did perform better compared to electronic glossaries. However, the analysis remained focused at terminological level, and it

was deemed necessary to expand the analysis to the sentence level. In addition, in Prandi's subsequent work (2018) on the use of InterpretBank in the booth, she also proposed that the integration of ASR in CAI tools such as InterpretBank would lower additional cognitive load as no manual-spatial response would be needed.

With regard to existing studies and findings on the use of ASR-CAI tools during SI, it remains unclear how the support of CAI tools with ASR integration may affect interpreters' processing capacity and the interpreting process. This paper hopes to bridge the gap regarding the use of ASR-CAI tools by implementing both product/process-oriented approaches to assess participants' overall performance and their cognitive effort during SI of speeches dense in terminology. By using combined analysis methods to evaluate participants' renditions of SI performed with the support of ASR-CAI tools, the study set out to explore the potential benefits and limitations of such tools as well as their usability during the simultaneous interpreting process.

3. Research Methodology

To investigate the influence of ASR-CAI tools on interpreters' performance and their impact on the interpreting process, we conducted a small-scale experiment with four trainee interpreters, using remote Zoom meetings due to covid-19 restrictions. Our objective was to observe how SI was performed with and without the support of an ASR-CAI tool. At the end of the experiment, we interviewed the participants to obtain feedback on their perception and experience in using the ASR-CAI tool. Both product/process-based methods were then implemented to analyze the data obtained from the experiment and interview. The product-based method examined the terminological quality and interpreting errors (as classified by Barik 1971) in the participants' rendition, whereas the process-based method analyzed the participants' accounts of their experiences and corroborated those accounts with the participants' observable interactions with the ASR-CAI tool support.

3.1 Tools

For the ASR-CAI tool of choice, this experiment used InterpretBank, a web-based ASR-supported CAI tool which transcribes the speech delivered by a speaker in real-time and automatically provides interpreters with numerals and translation options for terminology

drawn from the tool's terminology database. For this experiment, we prepared a glossary beforehand on the tool's database to support the tool's terminology lookup function.

Figure 1 shows a screenshot of InterpretBank's user interface. Real-time transcriptions of the source text are displayed at the top of the screen. On the left is the terminology section where translation suggestions for terminology are presented. Lastly, numerals are displayed on the right together with their units of measurements, with the newest information being displayed on top and highlighted in red.

Figure 1: InterpretBank's user interface

3.2 Participants

The experiment involved trainee interpreters from the Master of Arts Program in Interpretation at Chulalongkorn University, which is the sole conference interpreting degree program in Thailand. We aimed to recruit all students who were enrolled in the fourth and final semester of the program and had Thai as their A language and English as their B language. At the time this experiment was conducted in 2022, there were only six students in the program. As the principal investigator of this study was one of the six students, she was excluded from the study. Additionally, one of the six had Thai as their B language, which also led to their exclusion. Consequently, we ended up with a total of four participants. None

of them had prior experience using ASR-CAI tools during SI nor had they received training on how to use InterpretBank prior to this experiment.

3.3 Speeches

The experiment used two educational videos selected from YouTube. Both speeches were dense in terminology and comparable in terms of topics, information density, and the delivery style. The speeches were on the topics of (1) assisted reproductive technologies and (2) the condition of endometriosis. Both speeches were chosen based on the assumption that the participants were likely to have minimal familiarity with the selected topics. Each speech lasted about ten minutes with an average delivery rate of 160 words per minute. The first speech contained 61 specialized terms while the second speech contained 56 specialized terms. Since both topics are related to reproductive health, some specialized terms (eight) were present in both speeches.

3.4 Procedure

Due to covid-19 restrictions, the experiment was conducted remotely via Zoom meeting. Since it was conducted one on one with each participant, we did not find it necessary to enable the interpretation feature in Zoom. The performance of each participant was recorded by the participants themselves using a dedicated audio recording application of their choice. In addition, the entire experiment was also video-recorded in Zoom.

The participants were given instructions about the structure of the experiment and were informed that they would be interpreting two speeches which were dense in terminology from English to Thai. They were then given basic training on how to use InterpretBank. A short video was played to demonstrate how the tool operated and to show the types of information that the tool provided. This was also done so that participants could gain familiarity with how the information would be displayed on the user interface.

The participants were informed about the topics of the two speeches right before the experiment started and were given a briefing on the topic of each speech. Since the selected speeches were highly technical, the participants were given materials containing background information and glossary documents for both speeches to study 15 minutes before each

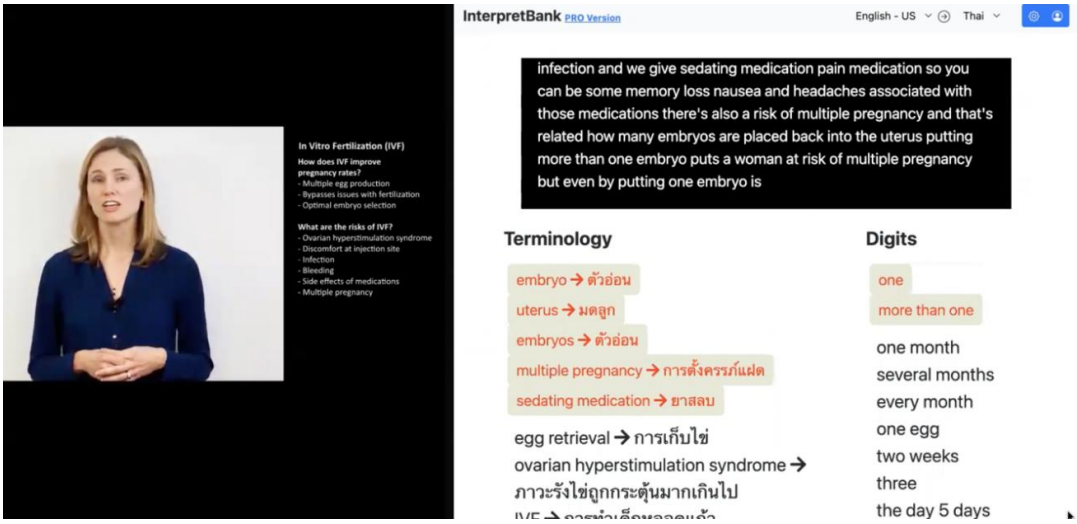
session started. However, they were not given time to prepare their own glossary or do any further research on each topic beforehand.

The materials for background information included two pages of information in Thai which covered all the main points being presented in the speeches. In addition, the materials also gave explanations and clarifications for the specialized terms that were found in both speeches. The glossary documents were given to the participants in the form of Excel files, featuring all the terminology presented in the speeches together with their translation equivalents.

The experiment was divided into two sessions, one for each speech.

a) The first session was carried out with the support of InterpretBank, in which the participants were provided with real-time transcriptions of the speech as well as suggestions for numerals and terminology translations. Providing that the participants had the support of InterpretBank in this session, they were asked not to consult the glossary document given to them during the interpreting process. In terms of the setup, the screen was divided into two areas as shown in Figure 2. The video of the speaker was positioned on the left-hand side while the user interface of InterpretBank was displayed on the right-hand side where the participants could see the information provided by the tool.

Figure 2: Screen setup for the experiment



b) In the second session, the participants were asked to perform SI without the support of InterpretBank. Instead, they were allowed to consult the provided glossary document throughout the entire session.

Since the two selected speeches were about reproductive health and shared eight specialized terms, we anticipated that the participants would likely perform better in the second session due to familiarity with the topic and terms gained during the first session (the so-called 'practice effect'). To prevent any potential bias associated with ASR- CAI tools leading to improved interpreter performance solely due to practice, we provided InterpretBank support during the first session. We also deliberately used the speech that contained a slightly higher number of specialized terms in the first session for the same reason.

At the end of the second session, we conducted a one-on-one interview with the participant to obtain feedback on the tool's usability. We aimed to investigate the participants' perception on the use of InterpretBank in the following aspects:

- The extent to which InterpretBank helps improve overall performance.
- Issues or difficulties that occurred during the SI performed with the support of InterpretBank.
- The usefulness of visual suggestions provided by the tool.
- Participants' perception regarding the tool's usability.
- Reports of any increased cognitive effort when working with ASR-CAI tool.

The subjective perceptions and experience reported by the participants during the one-on-one interviews would serve as a crucial starting point for our process-based analysis as we expected it to shed some light on the impact the ASR-CAI tool had on their interpreting process as well as their perception regarding the tool's usability.

3.5 Data analysis

Each participant's renditions of the two speeches, along with their interview, were transcribed manually as we could not find any application that could provide a satisfactory transcription

in Thai. The transcription included annotations for details such as pauses, hesitations and mispronunciations to allow us to observe the cognitive struggle of the participants. Subsequently, the data were analyzed both from the product- and process-based perspectives as follows:

3.5.1 Product-based analysis

To assess the impact of ASR-CAI tools on interpreters' performance, we first focused on the analysis of the product of the interpreting process, that is the participants' renditions. The transcripts of the participants' renditions were analyzed in terms of terminology quality and translation accuracy to determine whether the support of InterpretBank helped improve terminological precision and the quality of overall performance.

Terminology quality

While there are numerous studies that have assessed the accuracy of interpreters' number renditions (e.g. Desmet et al. 2018; Defrancq and Fantinuoli 2020; Pisani and Fantinuoli 2021), we have found very few that investigated terminology quality (Fantinuoli 2017b; Prandi 2018). While Fantinuoli (2017b) assessed the performance of a prototype ASR-CAI tool employing a dichotomous classification of correctness or error, Prandi (2018)'s measured precision in terminology renditions of a group of student interpreters with a three-way classification of terminology renditions: close (no information loss), acceptable (some information loss) and unacceptable (serious deviation).

Upon examining our data, we found instances that could not be categorized by Prandi (2018)'s classification. Therefore, we decided to assess the terminology quality in the participants' renditions by detecting mistranslations and use of English words, which means that terms were left untranslated and repeated as they were said in the source text. In this case, the English words that are widely used or borrowed into Thai would not be counted as errors if the participants left them as in the source language. Table 1 shows examples of the terms featured in the speech about assisted reproductive technologies, demonstrating which terms were considered as acceptable to be used in English and which terms required translation.

Table 1: Terminology from Speech 1 categorized by their requirements for translation

Translations not required	Translations required
Blastocyst stage	Cervix
Cystic fibrosis	Egg retrieval
Endometriosis	Fertility treatment
IUI – Intrauterine Insemination	Ovulation
IVF – In Vitro Fertilization	Unexplained infertility
LH surge	Uterus
Sperm wash	Vagina

Translation accuracy

For the assessment of translation accuracy, the analysis was expanded to the sentence level to examine whether the support of InterpretBank would lead to an improvement of participants' overall performance. In analyzing translation accuracy, the participants' renditions were annotated following the first and arguably the best known classification of errors encountered in SI by Henri C. Barik's (1971), which are generally classified as omissions, additions, and substitutions.

Omissions

Omissions were counted as errors when ideas in the source text were completely missing in the rendition, resulting in loss of information in the target text. The following paragraph shows an example of segments from the source text in which the participants were found to commit omissions in their renditions. In this example, the omitted ideas are shown in brackets.

"IUI is also known as sperm wash. It's done in cases of male factor infertility [or in cases where woman couples, single woman, or same-sex couples are] using donor sperm, [and it's also done in cases of unexplained infertility as well.]"

Additions

When information was presented in the target text where no reference to it could be found in the source text, such errors were considered as additions.

For example, in one segment which said, "The useful thing about doing this surgery is that if you find an endometrial tissue, it's possible to put some treatments in place such as cauterizing that tissue or excising it out," one participant correctly conveyed the key message in their rendition. However, they also added information about how endometrial tissue could cause damage to the body. In such cases when information was presented in the target text with no reference to the ideas presented in the source text, the error was considered as addition.

Substitutions

Errors that were considered as substitutions were when ideas in the source text were altered, resulting in inconsistency with the original sense, contradictions, ambiguity, or misinterpretations in the target text.

For example, during SI of the speech about endometriosis, one segment in the source text originally said, "There's no particular endometriosis gene that's been found, but we do find that it does have a tendency to run in families." However, one participant misinterpreted the segment by saying that endometriosis had no tendency to run in families. In cases when such errors occurred, those errors would be counted as substitutions.

3.5.2 Process-based analysis

The study of interpreting process, especially in the topic of how CAI tools impact the interpreting process, generally attempts to investigate interpreters' cognitive activity while they are interpreting. As "[c]ognitive activity is not directly observable, […] researchers must work with a range of methods to triangulate potential cognitive behavior" (Mellinger 2019: 36). Some used psycho-physiological measures such as heart rate and blood pressure (e.g. Korpal 2016) while others employed pauses and eye tracking data to shed light on hypotheses about cognitive load (Prandi 2017, 2018). Another possibility is to resort to retrospective

verbalization (via questionnaire or interview) to gain insight into participants' working process after it has finished (e.g. Pisani and Fantinuoli 2021). "These verbal protocols can be triangulated with the actual performance of interpreters to provide a more complete picture of the interpreting process" (Mellinger 2019: 37).

In this study, we employed three methods to elucidate the cognitive activities of our participants. First, we examined the influence of the CAI tool's functionality on our participants. To this purpose, we observed how the tool's core features, including data visualization, real-time transcription and translation suggestions – operated during the experiment and scrutinized the transcriptions of participants' interpreted renditions to discern the impact of these features. Secondly, we analyzed the interview transcripts to extract information in relation to the tool's usability and difficulties encountered during the interpreting sessions. In addition, to gain insight into the participants' cognitive effort we drew on similar research conducted by Ewa Gumul (2019) in which she aimed to measure the level of correspondence between the problems related to increased cognitive effort reported by interpreters and problem indicators identified in the target texts. In the present study, the target text segments in which the participants reported experiencing increased cognitive load were analyzed in search of the following problem indicators which have been proposed to reflect increased cognitive effort during SI.

- Pauses exceeding two seconds
- Omissions leading to the loss of information
- Mispronunciations
- Hesitation markers

4. Results and Discussion

4.1 Product-based analysis

In this section, we will present and discuss the results from the analysis of the transcriptions of the participants' interpreted renditions in relation to terminology quality and translation accuracy.

4.1.1 Terminology quality

Terminology quality here refers to accuracy of terminology rendition, with a specific focus on the occurrence of mistranslations and the use of English words, which means that terms were left untranslated and repeated as they were said in the source text. The study found that, in terms of terminology quality, the participants committed fewer mistranslations of specialized terms and used fewer English words in their renditions when they had the support of InterpretBank. For example, the terms "uterus" and "cervix" were presented in both speeches. When the participants performed SI with the support of InterpretBank, all of them were able to provide correct translations for both terms. However, during SI performed without the tool's support, several participants either left the two terms in English or gave incorrect translations for them.

Figures 3 and 4 show that the use of InterpretBank helped improve the terminology quality in all participants' renditions. On average, the participants committed 48% fewer mistranslations and used 72% fewer English word during SI performed with InterpretBank. The results indicated that the support of ASR-CAI tools such as InterpretBank led to a significant reduction of mistranslations in terminology rendition as well as reduced the use of English words as a coping strategy among all participants, resulting in a higher level of precision in terminology quality.

Figure 3: Mistranslations detected

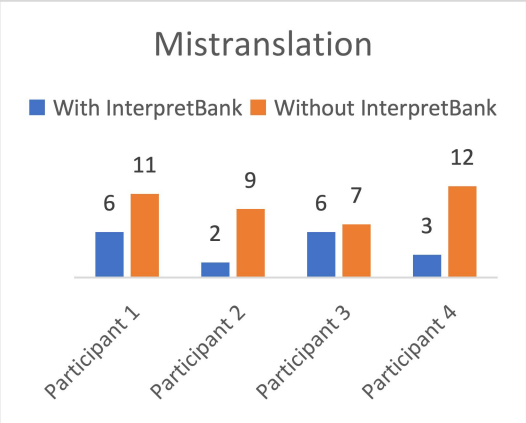
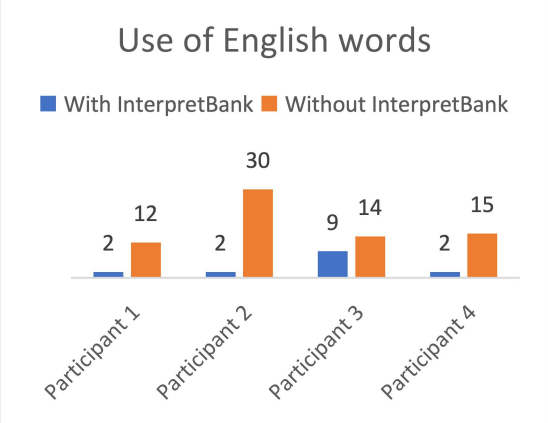


Figure 4: Use of English words



4.1.2 Translation accuracy

To assess translation accuracy, the analysis was extended to the sentence level in order to identify errors including omissions, additions and substitutions. Upon investigation, the study

found that the participants committed fewer omissions, additions, and substitutions during SI performed with the support of InterpretBank. Figures 5, 6, and 7 show the differences in the error rate of SI performed with and without the support of ASR-CAI tool.

Figure 5: Omissions detected

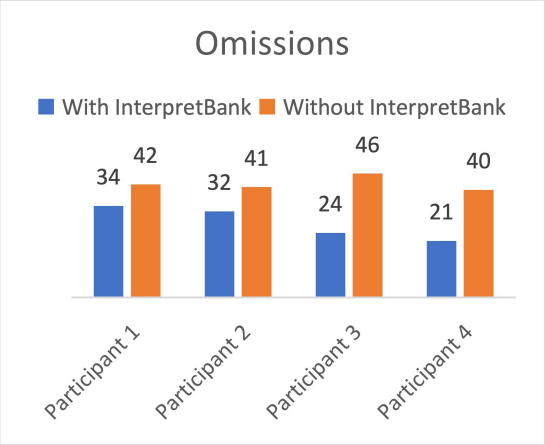


Figure 6: Additions detected

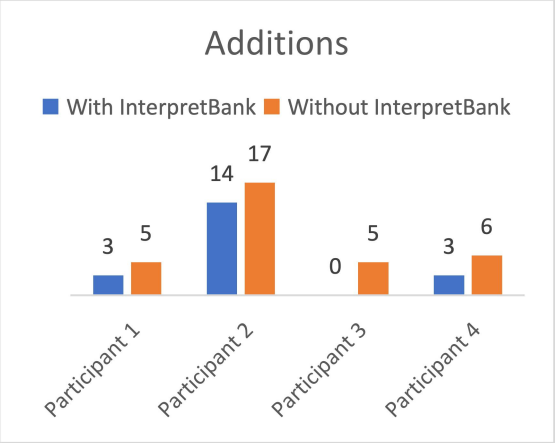
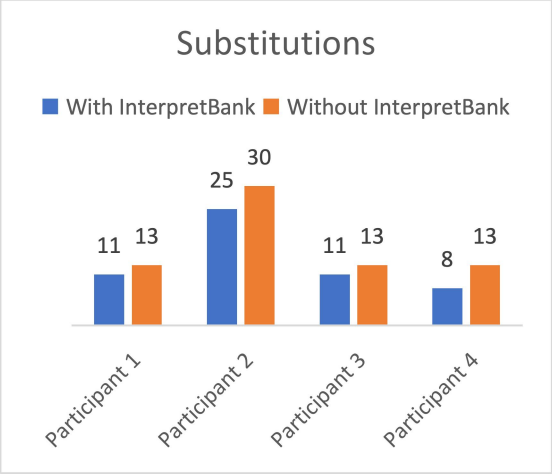


Figure 7: Substitutions detected



According to the data presented, the participants were found to commit fewer errors when they had the support of InterpretBank during SI. On average, InterpretBank was found to help reduce omissions by 33%, reduce additions by 51.9%, and reduce substitutions in the participants' renditions by 21%. In this regard, the study concluded that the use of ASR-CAI

tools proved to help reduce errors in all participants' renditions with the reduction of additions being the most evident.

The usefulness of InterpretBank's support during SI was further confirmed as the features of the two speeches used in the experiment were taken into account. With a slightly smaller number of specialized terms in the second speech and the familiarity with the topic and terms gained during the first session, the participants should have performed better during SI of the second speech. However, the data showed that the participants actually committed more mistranslations and errors and used more English words in their renditions of the second speech when they had no support from ASR-CAI tool.

Despite their lack of experience in dealing with the topic and the terminology being featured in the speech, the participants evidently performed better during SI of the first speech when they had the support of InterpretBank, resulting in a lower error rate in all participants' renditions.

This proved that ASR-CAI tools such as InterpretBank have the potential to serve as an effective support for interpreters during SI of specialized texts which are dense in terminology, and further confirmed that such tools could really help improve interpreters' overall performance and the quality of their terminology rendition.

4.2 Process-based analysis

In this section, we will present the results obtained from three methods we employed to shed light on the impact of the CAI-tool on the interpreting process, namely the observable impacts of CAI-Tool's features on interpreters, the participants perceptions, and the analysis of cognitive effort.

4.2.1 Impact of the CAI tool's core features on the participants

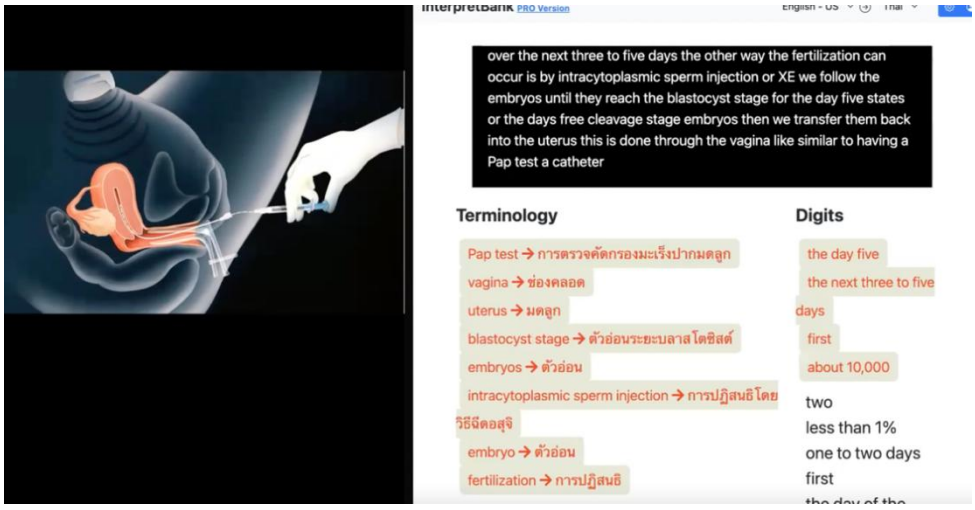
In this section, the functions provided by InterpretBank, including data visualization, real-time transcription, and translation suggestions of specialized terms, will be discussed in relation to how they affected the participants' performance and the interpreting process.

Data visualization

Regarding the use of InterpretBank during SI, data visualization was found to have significant impact on the participants' rendition quality. For example, in cases where the software failed to provide any translation suggestions for the terminology featured in the source text, the participants often struggled to come up with their own translations for the terminology presented. An example could be found in the segment which said, "IUI is also known as sperm wash. It's done in cases of male factor infertility or in cases where woman couples, single woman, or same-sex couples are using donor sperm." During this segment, InterpretBank failed to retrieve the translation for the term "male factor infertility" from the terminology database. As a result, no translation suggestion for the term was provided for the participants. This caused two participants to pause for several seconds and ended up omitting the term in their renditions. In addition, the delay also affected their renditions of the subsequent segments, resulting in omissions of information in the target text.

Furthermore, data visualization was also found to play a role in cases when multiple suggestions were displayed at the same time on the user interface as shown in Figure 8.

Figure 8: InterpretBank's interface with multiple suggestions displayed



The segment displayed in the transcription section contained eight specialized terms, and all of them were displayed at the same time on the interface together with the suggestions for

their translations. In such segments where multiple translation suggestions were shown, most participants reported having difficulty identifying the translations they needed. As a result, they were found to commit omissions and other types of errors such as additions and substitutions in their renditions.

According to this finding, the study concluded that data visualization could significantly impact the interpreting process as well as interpreters' rendition quality. In cases when ASR-CAI tool failed to provide any translation suggestions or when multiple suggestions were displayed at the same time on the interface, the interpreting process could potentially get interrupted, causing interpreters to commit errors in their renditions.

Real-time transcriptions

An assessment regarding the usefulness of real-time transcriptions revealed that the ASR system still had limitations relating to its precision. In analyzing the quality of real-time transcriptions, the study assessed the transcriptions provided by InterpretBank in search of errors in each segment. In this study, minor mistranscriptions which did not affect the overall meaning of the source text were overlooked. In this regard, only mistranscriptions which potentially resulted in contradictions, ambiguity, or misinterpretations were taken into account.

The study found that, in terms of terminology, the ASR system was able to correctly transcribe and identify 80% of the terminology featured in the glossary. However, when the transcriptions were assessed at the segment level, mistranscriptions were found in 50% of the segments. In these instances, the software was found to provide the participants with inaccurate transcriptions, which consequently led participants who relied on them to commit mistranslations in their renditions.

For example, when one participant was performing SI with the support of InterpretBank, the software mistranscribed one segment which said "IVF can be performed for a number of different reasons" as "I have to be performed for a number of different reasons." This caused confusion for the participants and consequently led them to commit errors in the target text.

Furthermore, the transcriptions were often found to shift back and forth before they became stabilized due to the real-time correction function of the ASR system.

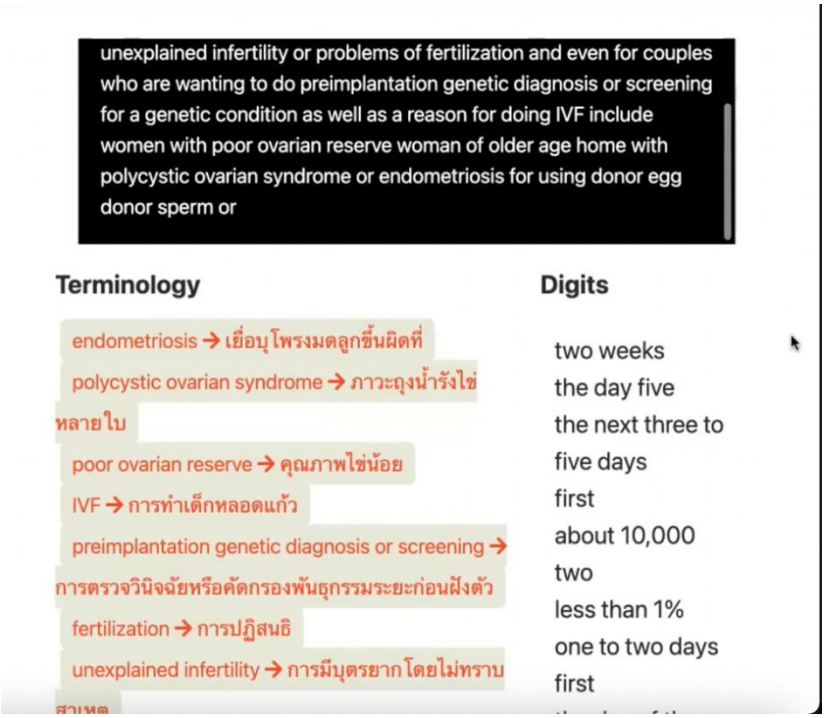
Considering the issues associated with the ASR system identified in this study, it seemed that real-time transcriptions still need improvement to be of effective support for interpreters. Admittedly, real-time transcriptions provided by ASR-CAI tools such as InterpretBank could provide a certain level of support for interpreters during the interpreting process. However, interpreters should still make sure not to rely completely on the transcriptions provided to them by the tool. Instead, they should keep in mind to use ASR-CAI tools as merely supportive tools rather than letting the tools' support replace their skills entirely.

Translation suggestions

Upon investigation, translation suggestions proved effective in providing support for the participants during SI of speeches dense in terminology, resulting in a higher level of precision in terminology rendition. The study found that the participants were able to integrate most of the suggested translations into their renditions in a correct context during segments that did not feature a high number of specialized terms, which led to an accurate rendition of the whole segments. On the other hand, suggested translations were often improperly integrated during segments that were particularly dense in terminology, resulting in errors such as additions or substitutions in the target text.

In this study, when a segment contained over five specialized terms which appeared close to one another, such segments were considered as dense in terminology. Figure 9 shows a screenshot of InterpretBank's interface, featuring the transcriptions for two segments that were considered dense in terminology. Both segments contained six specialized terms, all appearing in close range of one another.

Figure 9: Screenshot of segments dense in terminology



The following paragraph shows the full transcriptions of the two segments mentioned. The highlighted terms were included in the glossary and were displayed on the interface together with their translations.

"IVF can be performed for a number of different reasons, which include male factor infertility, tubal factor infertility, for women and couples with unexplained infertility or problems with fertilization, and even for couples who are wanting to do pre-implantation genetic diagnosis or screening for genetic condition as well."

"Additional reasons for doing IVF would include women with poor ovarian reserve, women of older age, women with polycystic ovarian syndrome or endometriosis or using donor egg or donor sperm or even embryo donation."

During such segments, the participants were able to integrate the translation suggestions for terminology into their renditions, but they also committed additions and substitutions in the process, resulting in an inaccurate rendition of the whole segments.

Based on this finding, the study concluded that translation suggestions had the potential to improve the participants' rendition quality in terms of terminology. However, they did not always lead to an accurate rendition of the segment in which the terminology was embedded.

Partly, the density of information provided by InterpretBank might have added to the participants' processing load. This additional task of input processing consequently required the participants to put additional effort into managing their processing capacity. This could potentially lead to processing capacity mismanagement considering that all participants had no prior experience in using ASR-CAI tools such as InterpretBank during SI. Thus, they lacked the familiarity and expertise in how to best handle the additional information provided to them, which caused them to commit errors in their renditions.

In addition, the occurrence of errors may also be ascribed to the participants' over-reliance on the support of ASR-CAI tool during SI. When the participants could not keep up with the speaker, they might resort to using the suggested translations for terminology as pointers for the content they missed. This led the participants to integrate translation suggestions into their renditions with no clear idea of the exact context that those terms appeared in, causing them to commit errors in their renditions at the sentence level while precision was still maintained at the terminological level.

4.2.2 Participants' perceptions

This section reports and discusses the results obtained from the participants' interviews. In general, the participants described the support of InterpretBank as helpful when dealing with speeches dense in terminology. They reported that they benefitted the most from the translation suggestions for terminology, whereas real-time transcriptions offered the least benefit as they were reported to be more of a distraction for most participants during the interpreting process. The participants who relied on the transcriptions for support reported having difficulty correcting their renditions in a prompt manner when the software provided them with incorrect transcriptions of the source text, which consequently led them to commit mistranslations in their renditions.

In terms of the user interface, the participants stated that the sections for each type of information were well-divided and that they had no difficulty identifying the information they needed from each section. However, two participants expressed that they sometimes struggled with how the data was presented, specifically when multiple translation suggestions appeared at the same time on the interface during segments rich in terminology. In these instances when several suggestions were displayed in the terminology section, almost all participants reported experiencing an increase in cognitive load, which consequently led them to commit omissions and other types of errors in their renditions.

Interestingly, one participant pointed out that the support of ASR-CAI tools such as InterpretBank had the potential to offer psychological benefits to interpreters through the availability of visual suggestions provided by the tool. The participant expressed that, by having the translation suggestions available when needed, it helped reduce the cognitive task of having to manually perform a glossary search. Moreover, this also helped reduce their cognitive pressure while performing SI of speeches dense in terminology.

In conclusion, the suggested translations for terminology were used most often by the participants which evidently led to an improvement of terminology quality in their renditions. Real-time transcriptions appeared to offer the least benefit and oftentimes posed as a source of distraction for the participants, leading them to commit mistranslations in the target text. In addition, most participants reported experiencing an increased cognitive effort when multiple translation suggestions were displayed at the same time on the interface. However, all participants emphasized that they might have been able to make use of the tool's functions more effectively had they had more familiarity with how the tool operated as well as more practice in how to use the tool.

4.2.3 Analysis of the cognitive effort

In this section, we corroborated the participants' reports during interviews with analysis of cognitive problem indicators such as long pauses, hesitation markers and mispronunciation. Upon investigation, the study found that the majority of the participants' reports regarding increased cognitive effort coincided with the problem indicators found in their renditions. In these instances, pauses, mispronunciations, and hesitation markers were identified in the

target text segments in Thai in which the participants reported experiencing an increase in cognitive load.

Most problem indicators were found in segments that were particularly dense in terminology, which consequently caused InterpretBank to display multiple translation suggestions for the terminology presented in those segments. This correlated with the participants' feedback in which they reported experiencing increased cognitive effort when the tool displayed multiple translation suggestions on the interface during segments dense in terminology, causing them to commit errors and omissions in their renditions.

The study concluded that the problem indicators found in the target text could indicate an increase in cognitive effort and processing problems experienced by the participants during the interpreting process. The increased cognitive effort experienced by the participants could partly be ascribed to the use of ASR-CAI tool during SI. The additional information provided by InterpretBank together with the density of information in the source text could potentially add to the participants' processing load and make them more susceptible to experiencing cognitive overload, causing them to commit errors such as additions, substitutions, and omissions in their renditions of subsequent segments.

5. Conclusion

The study confirmed that InterpretBank proved to be effective in providing interpreters support during SI of speeches dense in terminology. The tool's support evidently led to an improvement of terminology quality and a reduction of errors in all participants' renditions.

The paper discussed potential benefits and limitations of the functions provided by InterpretBank, namely the real-time transcriptions and translation suggestions for terminology. It was found that while the participants significantly benefitted from the translation suggestions, real-time transcriptions seemed to pose as a distraction due to issues associated with the ASR system such as latency and imprecision of the transcriptions, which led the participants who relied on them to commit errors in their renditions. In addition, the paper also investigated the influence of ASR-CAI tool on participants' processing capacity

and cognitive effort during SI, in which data visualization was discussed in relation to increased cognitive load and the participants' rendition quality.

The findings presented above can help shed light on the usability of ASR-CAI tools and their impact on interpreters' performance and cognitive effort during SI. Nevertheless, further research should be conducted to study the impact of ASR-CAI tools on interpreters' cognitive load with new methods tested to provide a clearer picture of how such tools can impact interpreters' processing capacity during the interpreting process. In addition, it would be beneficial if similar studies could be conducted with different population samples such as professional interpreters or interpreters with prior experience in using ASR-CAI tools to see how the use of such tools may impact their performance during SI.

Despite the limitations that InterpretBank presents, it is evident that the support of such tools can help improve interpreters' rendition quality and overall performance. Although the information provided by the tool appears to add to interpreters' cognitive load in certain circumstances, this issue can be minimized if interpreters are given opportunities to gain more familiarity with the tool's functions, after which they can devise their own strategies to better manage the additional information provided to them by ASR-CAI tools.

As technology makes its way into the field of interpretation, the implementation of CAI technologies has also become more prevalent. An integration of ASR in CAI tools presents interpreters with new possibilities in how interpretation can be performed and improved. Given that there are only a limited number of studies in this area, we hope that the findings which have been discussed here can provide a better understanding regarding the usability of ASR-CAI tools so that interpreters can make better use of the functions provided and better integrate them into the interpreting process for the most effective outcome.

References

Barik, Henri C. (1971) 'A description of various types of omissions, additions, and errors of translation encountered in simultaneous interpretation', *Meta,* 16*(4)*, 199-210.

Braun, Susanne, and Andrea Clarici (1996) 'Inaccuracy for numerals in simultaneous interpretation: Neurolinguistic and neuropsychological perspectives', *The Interpreters' Newsletter* 7: 85–102.

Bühler, Hildegund (1986) 'Linguistic (semantic) and extra-linguistic (pragmatic) criteria for the evaluation of conference interpretation and interpreters', *Multilingua* 5(4), 231-235.

Canali, Sara (2019) *Technologie und Zahlen beim Simultandolmetschen: Utilizzo del riconoscimento vocale come supporto durante l'interpretazione simultanea dei numeri,* MA thesis. Università degli studi internazionali di Roma.

Collard, Camille (2019) *A Corpus-based Study of Simultaneous Interpreting with Special Reference to Sex*. PhD dissertation. Ghent University.

Defrancq, Bart and Claudio Fantinuoli (2020) 'Automatic speech recognition in the booth: Assessment of system performance, interpreters' performances and interactions in the context of numbers', *Target,* 33*(*1*)*, 73-102.

Desmet, Bart, Mieke Vandierendonck and Bart Defrancq (2018) 'Simultaneous interpretation of numbers and the impact of technological support', in Claudio Fantinuoli *(*ed*.) Interpreting and Technology*, Berlin: Language Science Press, 13-17.

Fantinuoli, Claudio (2016) 'InterpretBank. Redefining computer-assisted interpreting tools', *Proceedings of the 38th Conference Translating and the Computer,* London: Editions Tradulex, 42-52.

Fantinuoli, Claudio (2017a) 'Computer-assisted preparation in conference interpreting', *Translation & Interpreting,* 9*(2)*, 24-37. DOI:10.12807/ti.109202.2017.a02

Fantinuoli, Claudio (2017b) 'Speech recognition in the interpreter workstation', *Proceedings of the 39th Conference Translating and the Computer*, Geneva: Editions Tradulex, 25-34.

Fantinuoli, Claudio (2022) 'Conference interpreting and new technologies', in Michaela Albl-Mikasa and Elisabet Tiselius (eds.) *Routledge Handbook of Conference Interpreting*, 508-522. DOI: 10.4324/9780429297878-44

Fantinuoli, Claudio and Maddalena Montecchio (2022) 'Defining maximum acceptable latency of AI-enhanced CAI tools', *arXiv preprint arXiv:2201.02792.*

Frittella, Francesca Maria (2019) '"70.6 billion world citizens": Investigating the difficulty of interpreting numbers', *Translation and Interpreting* 11 (1): 79–99.

Gile, Daniel (2009) *Basic Concepts and Models for Interpreter and Translator Training*, Amsterdam: John Benjamins.

Gumul, Ewa. (2019) 'Evidence of cognitive effort in simultaneous interpreting: Process versus product data', *Beyond Philology*, 16(4), 11-45. DOI: 10.26881/bp.2019.4.01

Korpal, Paweł (2016) *Linguistic and Psychological Indicators of Stress in Simultaneous Interpreting*, PhD dissertation. Adam Mickiewicz University.

Lamberger-Felber, Heike (2001) 'Text-oriented research into interpreting – Examples from a case-study', *Hermes* 26: 39–64.

Mazza, Cristina (2001) 'Numbers in simultaneous interpretation', *The Interpreters' Newsletter* 11: 87–104.

Mellinger, Christopher D. (2019) 'Computer-Assisted Interpreting Technologies and Interpreter Cognition: A Product and Process-Oriented Perspective', *Revista Tradumàtica. Tecnologies de la Traducció*, 17, 33-44.

Pinochi, Diletta (2009) 'Simultaneous interpretation of numbers: Comparing German and English to Italian. An experimental study', *The Interpreters' Newsletter* 14: 33–57.

Pisani, Elisabetta and Claudio Fantinuoli (2021) 'Measuring the impact of automatic speech recognition on number rendition in simultaneous interpreting', in Caiwen Wang and Binghan Zheng (eds.) *Empirical Studies of Translation and Interpreting*, New York: Routledge, 181-197. DOI:10.4324/9781003017400-14

Prandi, Bianca (2017) 'Designing a multimethod study on the use of CAI tools during simultaneous interpreting', *Proceedings of the 39th Translating and the Computer*. London: Editions Tradulex, 76-88.

Prandi, Bianca (2018) 'An exploratory study on CAI tools in simultaneous interpreting: Theoretical framework and stimulus validation', in Claudio Fantinuoli *(*ed.*)* *Interpreting and Technology*, Berlin: Language Science Press, 28-59.

Timarová, Sarka (2012) *Working Memory in Simultaneous Interpreting*, PhD dissertation. KU Leuven.