CHAPTER 3

Towards AI-enhanced computer-assisted interpreting

Claudio Fantinuoli
Johannes Gutenberg University Mainz, Germany | KUDO Inc.

Over the last 20 years, digital tools for interpreters have evolved from easy-to-use glossary creation and management solutions to complex workstations that cover most stages of the interpreter workflow, from assignment preparation to terminology lookup during interpretation. Currently, computer-aided interpretation (CAI) tools are starting to offer advanced features based on recent developments in artificial intelligence, such as automatic speech recognition, machine translation, etc., sometimes becoming part of integrated platforms, for example for the provision of remote interpretation.

This chapter provides a historical overview of CAI tools focusing on current technological developments in the artificial intelligence (AI) space. It begins by discussing the rationale for creating tailor-made interpreter support tools in the light of interpreters' specific needs. Against this background, the author offers a chronological overview of the development of CAI technology and a comparison and categorization of the available tools. Previous research on computer-assisted interpreting is briefly presented. It follows a general introduction to key AI topics as applied to the field of interpreting technology. The conclusions address potential avenues for future developments and challenges.

Keywords: computer-assisted interpreting (CAI), artificial intelligence, machine learning, interpreting, simultaneous interpreting

1. Introduction

In recent years, interest in computer-aided interpreting (CAI) tools, particularly but not exclusively in the field of simultaneous interpreting, has increased significantly. CAI tools are applications specifically designed to assist professional interpreters in at least one of the several sub-processes of interpreting, such as knowledge acquisition and management, lexicographical memorization, real-

https://doi.org/10.1075/ivitra.37.03fan

time terminology access, and so on. The tools available so far differ greatly both in the number of functionalities offered and in their architecture. These can be simple terminology management spreadsheets available on the user's computer or complex applications deployed in the cloud. Recently, advanced approaches to natural language processing and machine learning, especially deep learning, have been integrated, opening up new opportunities to develop advanced and intelligent tools. Advances in language-related technologies such as machine translation, speech recognition, and language modeling have the potential to transform many aspects of assistive technology in the interpreting field.

CAI tools are particularly relevant to improve the work experience of professional interpreters and to help them maintain or improve the quality of their service. This is especially true given a widespread need to streamline processes and compensate for shorter time-to-events, a trend that started years ago and is gaining momentum. In this context, the emergence of distance interpreting, especially remote simultaneous interpretation (RSI), seems to make the use of CAI tools even more relevant. The transition from an analog to a digital workspace, from the physical booth and the hardware console to the immateriality of a digitalized and artificial environment, has opened up new and natural ways to integrate CAI tools into the interpreter workflow.

While this AI-driven technologization process may help improve some aspects of the profession, many questions remain. They concern both the possibilities and limitations of what technology can do for the profession and how that technology affects the work of interpreters. This chapter will cover the former aspect of the interpreter-technology interaction. Due to the breadth of the subject covered in this chapter, which lies at the intersection of at least two disciplines, namely interpreting studies and computer science, the scope of this chapter is limited: introducing a non-specialist reader to the key technologies and applications relevant to the domain without the use of technical terms or jargon. It is left as an exercise for the reader to delve deeper into the topics using the references given at the end of the chapter.

The remainder of this chapter is organized as follows. Section 2 briefly discusses the stages of the interpreting workflow where such tools can be used. Section 3 gives an overview of the history of CAI tools, categorizing them into three groups according to their features and scope. Section 4 introduces the main language-related technologies used in AI-enhanced CAI tools. Then, Section 5 introduces the key features of AI-enhanced tools. The focus is in particular on automatic glossary creation (5.1), supporting functions for simultaneous interpreting (5.2) and for the consecutive modality (5.3). Underexplored assistive technologies (e.g., speech translation) are briefly presented in Section 5.4. The subject of interpreter management is dealt with in Section 6. Section 7 presents some ethical aspects of using AI. Finally, Section 7 concludes the chapter and addresses potential avenues for future developments and challenges.

2. Interpreting workflow and CAI tools

The interpreting workflow can be typically divided at least into three parts (cf. Gile 2009; Kalina 2007, Will 2009): before, during and after the event. Given the spontaneous character of speech and the time constraints that are typical of the interpreting process, knowledge acquisition occurs primarily *before* the conference. This is the phase in which preparatory work is performed (cf. Gile 2009, Stoll 2009, Will 2009). In this phase, tools are used to collect information, process it, create multilanguage glossaries and deep-dive into the topic of the event, both from a subject-matter as well as from a linguistic perspective.

While knowledge acquisition does not stop after the preparation phase, the retrieval of relevant information becomes particular relevant *during* the event. Digital tools that allow easy lookup of data are central to this activity. Finally, *after* the event, the reorganization and systematization of information plays a major role. This is where knowledge management tools come into play.

A few words need to be said about these three phases.

2.1 Preparation

In a typical specialized conference, interpreters work for specialists who share knowledge that is wholly or partially unknown to non-experts in the field. Communication is therefore characterized by a gap between the interpreter and the participants (cf. Gile 2009, Will 2009, Kucharska 2009). To fill this gap, interpreters need to prepare for the conference topic hours or even days in advance. This preparatory phase, in particular the role of specialist terminology and the strategies for its extraction and management, has been considered central to improving the quality of interpretation and helping interpreters to overcome the inherent difficulties of the interpreting process (cf. Pöchhacker 2016).

Since interpreters work for specialists whose knowledge is completely or partially unknown to outsiders, the knowledge gap introduced above manifests itself at least at two levels: the level of subject matter and of linguistic knowledge. Even if there is consensus on the crucial role and on some basic principles of preparation, for example on the fact that interpreters need an overall thematic knowledge into which terminology is embedded (Will 2009), daily approaches to preparation may diverge. For example, some support the idea that advance knowledge acquisition should focus on extra-linguistic information while others prioritize linguistic preparation, in particular to the terminological challenges (cf. Gile 2009).

In recent years, scholars have stressed the idea that preparation always needs to encompass linguistic and extra-linguistic knowledge. Knowledge in interpreting is described as a combination of language, content and situational expertise, moving from simple and sparse data to the establishment of a complex knowledge system (cf. Kalina 2007, Rütten 2007, Will 2009).

2.2 In-process

One of the most peculiar features of CAI tools is the ability to support the search for specialized terminology or other units of interest while interpreting. This functionality is generally considered as a backup strategy when other interpreting strategies, such as paraphrasing or the use of synonyms, are not viable and would lead to miscommunication and to a general degradation of the interpreter's performance. However, it can become of strategic importance if there is a need to use a specific terminology. This is the case, for example, when clients demand that interpreters adopt their own terminology, even when alternatives would be perfectly fine from a communicative point of view.

On the same line, the decreasing time to event observed in the new digitized world can limit the time for the interpreter to memorize certain word combinations, leading to a decrease in customer satisfaction. Interpreters can look up a term, generally in an event-specific database, while interpreting, while helping the boothmate or simply during breaks, perhaps to find the translation of a recurring term used in a previous speech.

While CAI tools have been designed with the ergonomical principle of reducing the cognitive effort needed to perform a search and retrieve the results, one of the main limits of such tools has been that they require the interpreter to allocate a specific amount of cognitive capacity to manually perform the search and to integrate the result of such operation into their delivery. Considering that simultaneous interpreting is a cognitively demanding task that is generally performed at the limit of cognitive saturation, a technological means of automating the search mechanism could have the potential to reduce cognitive load, benefiting the overall interpreting process. In this context, the integration of automatic speech recognition to automatise the lookup process may increase the usability of CAI tools. In fact, ASR has been regarded as a technology "with considerable potential for changing the way interpreting is practiced" (Pöchhacker, 2016, p. 188). Different to classic CAI tools that require manual input to get a translation for a given terminological unit, an ASR-enhanced CAI tool is able to automatise this process, with obvious advantages at the level of human – machine interaction.

2.3 Post-event

Interpreting does not end when the microphone is switched off. Some activities are usually performed after the end of a meeting. Notably, interpreters may want to update their informational assets, such as documents and glossaries, adding for example notes, correcting translations, or adding new terms. Especially when it comes to recurring meetings, both in terms of topics and clients, such information is considered useful to improve the quality of interpretation at future events. In some international organizations, where the number of interpreted meetings is high, complex informational workflows are set up to include post-meeting follow-ups. CAI tools are vital in this process as they can automate or at least streamline such activities.

A particular area of interest in the post-meeting scenario is evaluating interpreter performance. This can be done in informal follow-up discussions with teammates, with users of the interpreting services, or through structured feedback where the quality of rendition and overall performance is assessed by experienced interpreters. This kind of activity is performed nowadays exclusively by humans, and no tool has been developed to automate, at least to some extent, this process. The reasons are manifolds: on the one hand it is difficult to define in a formal and operative way the concept of quality in interpretation (cf. Pöchhacker 2002, Tiselius 2009, Kalina 2020). This challenge is not only known among interpreting scholars, but also among trainers and evaluators (cf. Behr 2013). On the other hand, a robust evaluation of spoken language translation requires a level of understanding of language, communicative situation, and context, to name just a few, that machines still do not possess (Bender and Koller 2020). Notwithstanding the limitations, some element of evaluation could be automized, as it will be introduced in Section 5.4.

3. History of computer-assited interpreting tools

In the last 20 years some effort has been put into the development of digital tools to support professional interpreters, following a similar path that has been taken in the translation profession.

Computer-assisted interpreting (CAI) tools fall into the category of process-oriented technologies (Fantinuoli 2018), as they influence the cognitive processes involved in interpreting with the general goal of improving the quality and the productivity of interpreters, while keeping the additional cognitive load, especially during interpretation, as low as possible.

At specialized conferences and events, interpreters do translate the clients' spoken words, often rich in technical jargon, even though they do not share their same level of subject matter knowledge as speakers and listeners. With this backdrop, CAI tools have been initially designed drawing from professional interpreters' experiences as well as on studies carried out on the topic with the goal to support interpreters in acquiring specialized knowledge, organizing semantic information, and making it available while interpreting.

Much of the initial scientific work on the subject of CAI tools goes back to German scholars. Rütten (2007) describes some of the characteristics and structure of an ideal CAI tool, thereby laying the foundation for the development of software programs for interpreters. According to Rütten, CAI tools should provide the user with modules for conducting online and offline research, managing documents for conference preparation, extracting and analyzing terminology, organizing and managing terminology and memorizing it. In view of the time pressure under which interpreters work (cf. Rütten 2003), such a tool should also include a system for quick and precise queries in the terminology database. Stoll (2009) describes a futuristic interpreter workplace and emphasizes the importance of ergonomics in the human-machine relationship.

In 2011, Fantinuoli expands on the ideas outlined in the aforementioned publications and presents a new generation of computer-aided interpreting tools that utilize advanced approaches to natural language processing, such as automatic corpus creation and terminology extraction. The integration of AI-based technologies in a CAI tool is proposed for the first time years later (Fantinuoli 2017c), with the presentation of the first prototype of a Virtual Boothmate, a tool that integrates automatic speech recognition to suggest in real-time some problem triggers for simultaneous interpreters, namely terminology and numbers.

In the last years, seminal publications and the availability of early prototypes have generated some interest in such technologies, particularly in relation to the impact of such tools on the quality of the interpreter's performances. While some of these studies take a descriptive approach (Costa et al. 2016; Schild Ortiz and Cavallo 2018), others apply empirical methods to investigate how the use of CAI tools affects the quality of interpretation, both during conference preparation (Xu 2018) and during simultaneous interpretation (Defrancq and Fantinuoli 2020; Pisani and Fantinuoli 2021; Prandi 2022).

Following the path of increasing automation of CAI tools, the initial interest in the effect of manually looking up glossaries in the booth (cf. Prandi 2017) has recently been overthrown by the study of automatic suggestions using artificial intelligence (Defrancq and Fantinuoli 2020, Pisani and Fantinuoli 2021, Prandi 2022).

Notwithstanding the promises made to the interpreting profession in recent years, the impact of the CAI tools remains modest. The number of available tools is limited and its design and development have been traditional limited to academic projects[1] or start-ups.[2]

There are multiple reasons for this. Most obviously, the development of CAI tools suffers from the marginality of interpreting activity in terms of industry size and business opportunities. Since the number of potential users is relatively small and the economic benefits of using such tools are far less tangible than, for example, with written translation, no great effort has been made to develop real-life applications. This is the reason of the projects undertaken in the past have been research-driven and arose out of academic interest rather than market demand.

More recently, in fact, CAI tools have received a new wave of attention in the context of other major technological developments, most notably advances in artificial intelligence and prospective use of such technologies into remote simultaneous interpretation applications, and the related business opportunities tight to their integration. Looking ahead, the ubiquity of AI-based technologies such as speech recognition, machine translation, etc., and their greater accessibility, both integrated into interpreter-agnostic applications and into tools specifically designed for the profession, is likely to lead to wider use of such developments.

From a mere technical and product-based perspective, CAI tools can be categorized according to several criteria depending on the workflow phases they cover, the needs they are designed to satisfy, the technological sophistication they have, etc. For the purposes of this chapter, CAI tools can be broadly classified into three into groups: first-generation CAI tools, which were first proposed about 20 years ago, second-generation CAI tools, which have been developed until the rise of artificial intelligence, and more recently, third-generation CAI tools designed around AI features.

3.1 First-generation

First-generation CAI tools are programs designed to manually create and manage terminology in an interpreter-friendly way. The first tools date from the turn of the millennium. They are very simple in terms of architecture and functionality and support interpreters in managing multilingual glossaries similar to MS Word or Excel lists. They do not provide for any other specific supporting activity of the interpreting process, such as information retrieval or real-time support. They

1. See for example Corpas Pastor (2022).

2. See for example InterpretBank (www.interpretbank.com) and Interpreter's Help (www.interpretershelp.com).

were mainly desktop applications designed to store and retrieve terminological data from a database.

First-generation CAI tools can be treated as a simplified version of traditional terminology management systems commonly used by terminologists and translators, with simple input structures and user-friendly search features. For example, to search the database or a subset of it, the user generally types a string of text (the term or part of it) into the search form and presses the Enter key. No advanced search algorithm specifically designed to take into account the time constraints of the (live) interpretation task (such as spell correction, progressive search in one or more glossaries) are contemplated.

Although such tools have been widely viewed as a first step towards computer-aided optimization of some aspects of the interpreting task – for example, to make the use of paper glossaries in the booth obsolete or to facilitate the reusability of previously collected terminology data – their success has remained limited and their impact on the profession very marginal.

3.2 Second-generation

Second-generation CAI tools aimed to expand the limited scope of the first tools developed to date by integrating a more holistic approach to terminology and knowledge management for interpreting tasks. They built on initial academic research and studies on terminology and knowledge management in interpreting (e.g. Rütten 2007, Will 2009) as well as on the application of advanced approaches in natural language processing and computer linguistics to the interpreting task (e.g., Fantinuoli 2006).

They offer more advanced functionalities that go beyond basic terminology management, such as features to organize preparatory material, retrieve information from corpora or other resources (both online and offline), learn conceptualized domains, etc. Second-generation tools exploited more advanced computational approaches to offer a supporting toolset suitable for different phases of the interpreting process, from preparation to interpretation. For the preparatory phase, for example, they comprise terminological lookup in online resources, automatic terminology extraction from preparatory documents, concordance, glossary memorization, and so forth.

The lookup mechanisms of such tools have been designed for the simultaneous modality and are different from the ones implemented in translation-oriented terminology tools. In order to reduce the cognitive load needed to look up a term, this kind of CAI tools use algorithms designed to reduce the number of strokes needed to input the search word, to correct typing errors, to discriminate results according to the conference topics, their relevance, etc.

With the advancements and the optimization potentials offered by second-generation CAI tools, the interest for such tools has increased and some of the tools gained a certain degree of popularity, with early adopters among freelance interpreters and language service providers.

3.3 Third-generation

Third generation CAI tools represent an evolution of digital programs for interpreters as they aim to integrate recent developments in the field of artificial intelligence as applied to natural language. Examples are automatic speech recognition, machine translation, etc. Third generation tools are currently the focus of several commercial, academic, and institutional projects (see also Zhang et al., this volume). To better put such tools in perspective, Section 4 will present the basic concepts used in language technologies applied to the domain, while Section 5 will focus on the practical application of such technologies as integrated in CAI tools.

4. Language technologies and artificial intelligence

This section introduces the basic concepts of language technologies and artificial intelligence relevant to interpreting.

4.1 Artificial intelligence and machine learning

Artificial intelligence (AI) is a branch of computer science concerned with building smart machines capable of performing tasks that are typically associated with some form of intelligence when performed by humans. In contrast to General or Strong AI (intelligent machines that are indistinguishable from the human mind), Narrow or Weak AI is used to refer to systems designed to handle a single or limited task.

Examples of Narrow AI have become extremely common in our society in the last few years, ranging from systems to perform medical diagnosis to algorithms able to suggest the next film to watch. Almost every area of life today makes use of some form of Narrow AI, both in an explicit way, i.e., with the user is aware of its use, for example in text generation algorithms used by editors, as well as implicitly, where it integrates into more complex applications, e.g., voice assistants, holiday booking platforms, etc.

Recent advances in AI have been made possible by developments in machine learning and deep learning, a field of computer science that aims to teach machines how to learn and perform tasks without being explicitly programmed to do so. Machine learning is an approach that involves building models, a mathematical representation of reality, by means of exposing the machine to data and let it "learn" from them. In particular, models try to make accurate prediction given a specific input. In visual system, for example, exposing the machine to a great number of pictures of animals paired with their species will allow the algorithm to make a prediction of the most probable species an animal belongs to when exposed to a previously unseen picture of an animal. In this respect, machine learning emulates a very human principle, i.e., a learning process driven by experience.

For this reason, given the right data and a clearly defined predictive task, most activities that require some sort of analysis and decision taking can be performed with different levels of quality by a machine. It is important to point out that to perform well, and in many cases to outperform humans, machines do not have to imitate human intelligence, but can follow completely different approaches (Floridi 2014).

Machine learning can be applied to human language too. This is the area of activity of Natural Language Processing, as described in the next session.

4.2 Natural language processing, understanding and generation

Natural Language Processing (NLP) is a discipline in computer science that aims at automating the manipulation of natural language to achieve some specific goals and to enable computers to understand human language in both written and verbal forms. NLP has been around for 50 years or so, bringing about many everyday applications such as word spelling correctors, and the like. Lately, NLP has been deeply influenced by machine learning (ML) and deep learning (DL) advancements so that most NLP applications are nowadays based on ML. From pipeline components, such as tokenization, stemming, part-of-speech tagging, syntactic parsers, etc., to complete applications, such as machine translation, summarization, autocompletion etc., NLP makes use of large quantity of language data to create general language models, such as BERT (Devlin et al. 2019) or GPT-3 (Brown et al. 2020). Such models are mathematical representations of a language that can be used to perform several tasks and be integrated in higher-level pipelines, for example automatic speech recognition and machine translation. CAI tools make abundance use of NLP, for example to match terminology in a digital boothmate (see 5.2), to translate terminologies, etc.

Natural Language Understanding (NLU) is a subfield of natural language processing that aims at allowing machines to develop some sort of understanding of the language and the communication process. They do this by using syntactic and semantic analysis of text and speech to determine the meaning of a sentence or of a text. NLU establishes a relevant ontology: a data structure which specifies the relationships between words and phrases, disambiguating synonymies, etc.

Defining what NLU is from a linguistic and a philosophical point of view (Bender and Koller 2020) is very difficult. This leads in many cases to confusion about the extent of understanding that a system reveals. Famous intelligent systems such as IBM's Watson, that won in 2011[3] the Jeopardy game against humans, do not possess, for example, any capability of "understanding", at least in the sense that is intuitively defined by humans, even less the sign any "intelligence". However, this is not a limitation for many applications built on top of NLP and NLU, since it is clear by now that many smart systems can be built without the need for the machine to manifest any intelligence (Floridi 2014). NLU can be used by CAI-tools to produce a quality translation of a speech, to better match terminology in a digital boothmate, to extract Named Entities, such as proper names, etc.

Natural Language Generation (NLG) is another subfield of natural language processing. While natural language understanding focuses on computer understanding and comprehension, natural language generation enables computers to produce language, typically in writing. In more technical terms, NLG is the process of producing a human language text response based on some data input. This text can also be converted into a speech format through text-to-speech synthesis. NLG can be used by CAI-tools to produce a summarization of a speech while maintaining the integrity of the information, for example.

4.3 Automatic speech recogintion

Automatic speech recognition (ASR), i.e. the process of transcribing spoken language into written text, is a practical NLP application. The common use cases of ASR range from dictation applications and voice assistants to software for the analysis of customers sentiments, video captioning, etc. In general terms, modern ASR systems are a combination of acoustic modelling and linguistic modelling. They therefore combine a knowledge representation of phonemes in a particular language with the probabilistic rules of that language. In other words, ASR systems make educated guesses on how to transcribe words by assessing its acoustic-based suggestions against correct syntactic, semantic, and tonal rules. Various components can be added to the pipeline, such as text-based recurrent neural network models to add unspoken punctuation,[4] etc. More recently, end-to-end approaches have emerged in the domain. Such approaches map in a single language model the

3. See https://www.ibm.com/ibm/history/ibm100/us/en/icons/watson/
4. https://ai.googleblog.com/2019/10/on-device-captioning-with-live-caption.html

audio and the desired transcription, simplifying considerably the system architecture (Amodei et al. 2016).

For certain languages and in specific contexts, typically speeches with a high degree of formality, clear pronunciation, and good quality of the acoustic signal, ASR performs like human transcribers, with word errors rates (WER) as low as 3–4 % (Filippidou and Moussiades 2020). The quality with dysfunctional or unconventional spoken language, with specific domains, with unfavorable acoustic conditions, and for low-resourced languages. To partially compensate for these shortcomings, baseline ASR models can be customized with new data to account for specific situations. So, for example, specific vocabulary (terms and phrases) can be presented to the language models in order to reduce the out of vocabulary effect, increasing the quality of the transcription.

Inference within ASR systems can be performed in batch or in a real-time (streaming). In the first case, the entire audio is processed by the ASR and a final transcription is produced by the tool. In the latter, the audio is processed while the speech is still unfolding, producing temporary partial transcriptions that may be corrected by the system as soon as more context gets available. Since ASR is a computational demanding task, ASR is typically deployed on the cloud where dedicated powerful machines are available. However, the simplification of the inferencing algorithms and the size reduction of models is making possible to run ASR also on edge devices (offline). This development is supposed to boost a new series of applications, for example mitigating concerns on confidentiality of data (see 7).

4.4 Machine translation

Machine translation (MT) is the process of converting a written text from one language into written text in another language. While in the past MT has been traditionally focused on rule-based, statistical approaches, or a combination of them, MT follows nowadays the ML paradigm. Baseline systems are trained by ingesting a large amount of text and their translations in an end-to-end fashion, deriving a mathematical representation of the translation process. Several additional layers of transformation can be applied to such models, such as target language adaptation or terminology enforcement, to name just a few, with the goal to improve translation quality and trim its results to specific needs.

Machine translation can be applied as an agent of interlinguistic communication, as a support for professional translators, or as a component of other applications, such as CAI tool, for example to translate terminology lists (see 5.1). In several contexts, MT has reached an unprecedented level of precision, making its use ubiquitous in everyday life as well as in the professional workflow of many translators or other language service providers. Besides intrinsic limits such scarce NLU abilities, absence of world knowledge, cultural awareness and the like, the potential of MT in real-world scenarios is still unfolding its whole potential. Much attention is devoted lately in building applications around MT that able to satisfy specific user cases. The technology needed to build and deploy MT engines is getting simpler, and many pre-trained models are available in the open-source space.[5]

4.5 Machine interpreting

Machine interpreting (MI), also known as speech-to-text or speech-to-speech translation, is the process of converting a spoken text from one language into written or spoken text in another language in real-time, translating the original while it is still unfolding. MI systems have the potential to be used in live communicative settings for language access, such as institutional events, lectures, conferences, etc. and to make multilingual content accessible in real-time, thus increasing inclusion and participation when human services for language accessibility are not available, such as live interlingual subtitling (Romero-Fresco and Pöchhacker 2017) or conference interpreting (Pöchhacker 2016). MI, especially the speech-to-text variation, can also be integrated into CAI tools and be used to augment professional interpreters, for example to offer suggestions in real-time to the human interpreter (see 5.4).

Although the history of MI is quite long, it is only recently that much effort has been made in this sub-discipline of MT, with conferences, evaluation campaigns and major organizations, such as Google, Meta, etc. working on it (cf. Jia, Weiss et al. 2019, Jia, Gu et al. 2021, Jian, Ramanovich et al. 2021). Many challenges related to the high complexity that is typical of the spoken language, both in the professional and in the everyday context, are still to be solved. Much progress has been done in the last few years (cf. Anastasopoulos et al. 2021), however major advancements are required to make the technology enter real-life settings, as recent user-centric evaluations reveal (cf. Karakanta et al. 2021, Fantinuoli and Prandi 2021).

There are two approaches to MI: the cascading and the end-to-end approach. The former comprises a combination of speech recognition, machine translation, and, if required, speech synthesis (cf. Sudoh et al. 2020). This is the technology typically used in applications at the moment of writing. This approach can profit from already existent technologies and extensive dataset to train the models.

5. A list of freely available machine translation models is available here https://huggingface.co

The end-to-end approach, on the contrary, directly translates the source speech in the target language, either in textual form or in a speech, without the need to use intermediate representations of what has been said, for example the transcription. While the end-to-end approach is still in its infancy, it promises to simplify the way translation systems are built. The major limitation of this approach, at the moment, is data scarcity (cf. Sperber and Paulik 2020).

4.6 Summarisation

Text summarization is an NLP application that aims at reducing the size of a text by maintaining the key concepts and information expressed in the original using some form of heuristics or statistical methods. Summarization is typically achieved with two concurrent approaches: the extractive and the abstractive approach. The extractive approach aims at identifying units of text (generally sentences) in the original document that contain information worth to be included in the summarized version. The whole of the retained sentences will form the summary. The abstractive approach, by contrast, employs more powerful natural language processing techniques to 'understand' text and generate new summary text.

Summarization, as any other complex language processing technique, such as translation, suffers from computerized language processing approaches having limited understanding capabilities, as described in Section 4.2. Notwithstanding this intrinsic limitation, depending on the specific goals of the task, summarization results seem good enough to be used in some real-world applications, for example to get the gist of a document, to classify a text, etc. This is the role that such feature can play inside of CAI tools.

5. AI-enhanced CAI tools

As introduced in Section 2.3, the newest versions of CAI tools are characterized by a higher degree of automation compared to the previous ones. By integrating new advancements in ML-based NLP, AI-enhanced tools aim at partially or fully automatizing some aspects of the interpreting workflow, from the preparation work to in-process and post-event activities (see Section 3). In the next sessions, several application fields will be described. Some of them have already be implemented in commercial products, others are just potentials applications that have not seen a concrete implementation so far. In some cases, some applications have not been included here and are discussed in the dedicated chapters of this book.

5.1 Automatic glossary creation

Due to the intrinsic challenges of multilingual terminology, such as lexical ambiguity, domain specificity, etc. (cf. Steurs and Tryczynska 2021), the output of an automatic glossary building engine is typically a draft and the user, namely the interpreter, is required to validate the terms and their translation, editing the solution proposed or adding new entries based on her particular needs (Fantinuoli 2017a).

There are a multitude of approaches to multilingual glossary creation, and infinite variations of their constituent components. A typical workflow will be composed of the following parts:

- Corpus creation
- Term extraction
- Term translation
- Glossary evaluation

5.1.1 *Corpus creation*

The corpus building step aims to collect textual material on a specific topic in order to extract the relevant terminology from it. Depending on the subsequent approaches to terminology extraction, the corpus created in this phase can be monolingual (Fantinuoli 2018b), bilingual comparable (Jia, Gu et al., 2021; Corpas Pastor and Gaber 2021) or bilingual parallel (Haque et al. 2014) corpora.

Corpora can be created manually, for example by collecting relevant documents through web searches, or in automatic and semi-automatic ways, for example by using the web scraping approach as introduced in BootCaT (Baroni and Bernardini 2004) or similar software. In this case, the interpreter defines the topic by means of keywords and uses a search engine to locate download, and process documents (Fantinuoli 2006, 2018b). When the terminology and phraseology of a specific client is needed, a typical way to create a thematic corpus is to scrape the content of a single webpage. Other approaches using the web as a semantic source of information for interpreting tasks are possible (cf. Fantinuoli, Marchesini et al., 2022).

5.1.2 *Term extraction*

Monolingual terminology extraction from a text corpus is the process of identifying the relevant terminology by means of statistical and linguistic approaches, aiming at maximizing the level of precision and recall. From a user perspective, the majority of extracted terms should be potentially useful while the number of malformed terms or general words should be kept to a minimum. In interpreting, this poses several challenges because it is difficult to generalize what interpreters may consider useful or not, depending on their knowledge background, attitude and so forth (cf. Fantinuoli 2006; Fantinuoli, Marchesini et al., 2022).

From a computational perspective, several statistical measures have been defined to compute the degree of *termhood* of candidate terms, i.e., the probability that a candidate is to be considered a proper term. From the other side, to define, identify and recognize terms pure linguistic properties have been applied, using linguistic filtering techniques aiming to identify specific syntactic term patterns such as noun + noun and adjective + noun for English terms. Hybrid approaches, finally have tried to combine together these two paradigms, taking into account both linguistic and statistical hints to recognize proper terms. The application of these approaches in the interpreter workstations are discussed for example in Fantinuoli, Marchesini et al. (2022).

5.1.3 *Term translation*

Terminology translation plays a crucial role in domain-specific interpreting. The task of translating a monolingual terminology list can be solved with different approaches depending on the overall architecture of the tool, for example:

- the automatic lookup in pre-existing terminological repositories, integrated in the tool or freely available on the web
- the use of corpora of parallel documents to extract on-the-fly translation candidates, for example using word-similarity algorithms
- the use of machine translation

All approaches suffer from one major shortcoming: the issue with lexical ambiguity and the need of contextual information in order to find a suitable translation candidate. It is common knowledge that terms can be translated in different way depending on many factors, such as domain, context, etc. To mitigate this issue, context aware approaches to term translation can be used. This can be achieved for example by using machine translation on the entire segment containing the term to be translated, thus using context as a disambiguation feature, or by simply focusing the translation generation to specific domains. This is the case, for example, by applying bilingual terminology extraction from domain-specialized corpora.

5.1.4 *Glossary review*

Glossary review is the process of assessing the quality of a bilingual or multilingual glossary and editing it in order to achieve the desired result. The evaluation of terminology translation, despite its importance in the industry, has been a less examined area in interpreting research. Term translation quality is usually performed by the interpreters. Typically, the review process will lead to the deletion of entries that are not considered useful, for example because considered too general, well-known, etc. or because they are out-of-domain or malformed. During the review process interpreters are called to improve the suggestions proposed by the machine.

A certain degree of automation can be obtained for the evaluation of terms and translations, adopting several approaches ranging from the combination of suggestions generated by different systems to the computation of word-vectors for the entry pairs (cf. Haque et al. 2019; Bakaric et al. 2021).

5.2 Artificial boothmate

An area of particular interest for the application of artificial intelligence in the interpreting workflow, especially but not exclusively in the simultaneous modality, is the possibility to have tools that automatically suggest, in real-time, problem triggers, such as numbers, terminology, and proper names (Fantinuoli 2017b). An Artificial Boothmate (ABM) aims at increasing the quality of the rendition in terms of precision and accuracy for those elements that have been suggested in literature as very challenging for the human interpreter, such as numerals and terminology (Braun and Clarici 1996; Gile 2009; Setton and Dawrant 2016). While humans are very good at making sense of information, machines are superior in terms of memory and information retrieval capacities. An artificial boothmate aims at leveraging this ability of machines, giving the interpreter the possibility to concentrate on what she can do best, elaborating meaning and transferring it into the target language. At the moment of writing, three ABM tools have been developed: InterpretBank (Fantinuoli 2017a, 2017c), SmartTerp (Rodríguez et al., 2021) and KUDO Interpreter Assist (Fantinuoli et al. 2022a).

Over the years, a handful of empirical studies have been carried out to test the feasibility of the human-machine interaction in the simultaneous modality. They have focused in particular on the effectiveness of ASR-support during the interpretation of numbers (Desmet et al. 2018, Defrancq and Fantinuoli 2020, Pisani and Fantinuoli 2021), and first studies have been conducted on terminology (Prandi 2022). While the results are still provisional, they seem to suggest the positive effect of real-time support both in terms of increasing accuracy as well as reducing omissions.

The architecture of an ABM comprises generally three components: (a) an automatic speech recognition (ASR) engine to transcribe in real-time the speech uttered by the speaker, (b) a language model (LM) to retrieve the units of interest from the unfolding transcription and match them with the translations curated in a glossary or translated by means of machine translation, and (c) a user interface to display the extracted information to the interpreter.

There are possible variants of the architecture introduced above. For example, it is possible to use direct speech-to-text translation and present the interpreter the running raw translation of the speech or use this translation to select the units of interest. From an architectural perspective, a central role is played by the retrieval mechanisms of the LM. While digits and proper names are usually retrieved by means of automated Named Entity Recognition (NER) based on machine learning approaches, terminology can be retrieved both by means of NER or by looking up terminological units in the user's database. The former offers the advantage of not requiring a dataset, i.e., a glossary compiled for a specific event, which is obviously an advantage in reducing preparation time, but the latter allows for customization and verification of the results as it is audited data used for a specific client, event or interpreter.

Besides user-centric evaluations, both in terms of human-machine interaction and of interpreting process analysis, ABMs can be evaluated by means of technical assessments. This aims at evaluating parameters such as latency, precision, and recall. These values are good indicators of the performances of the tool as far as their technical implementation and information retrieval strategies are concerned. This evaluates the robustness of the ASR (or similar) engine, the ability of the ML to generalize and extract relevant information, and the delay with which these operations are performed, a critical aspect in the simultaneous interpreting modality (cf. Fantinuoli and Montecchio 2022).

5.3 Artificial notepad

An artificial notepad is a tool specifically designed for the use in the dialogic or consecutive modality. Artificial notepads for interpreting extend the support provided by general note-taking applications (cf. Goldsmith 2018, and also Goldsmith, this volume) with AI-based functions. Unlike ABMs, whose goal is to suggest specific parts of the speech in the simultaneous modality, the focus of an artificial notepad is on the consecutive modality, where the transcription of the original speech is annotated with information useful for producing a detailed and precise rendition in the target language.

The transcription can be used to perform sight-translation or as a supportive tool for the notes taken by the interpreter, alleviating memory constraints and leading to a higher precision of the renditions. Because of the technological constraints, this kind of support seems to be suitable for well-structured speeches, while it may be detrimental for spontaneous speeches, rich in disfluencies, corrections and badly structured.

The potential range of support offered in this kind of tools may be very broad, ranging from real-time transcription of the source speech segmented in paragraphs, highlight of units of interests, such as proper names or numbers, and the conversion of units of measurement between metrical systems. Terms or phrases in the transcription can be machine translated on-the-fly. At the moment of writing, the number of applications is limited, and consequently no targeted empirical studies have been conducted on them.[6]

5.4 Underexplored use of AI

The fast and exponential advancements in several areas of artificial intelligence are still unfolding their potential for supportive applications in the domain of human interpretation. Many uses of AI still need to be explored. While an Artificial Boothmate is designed around the idea to limit the number of supportive inputs for the interpreter, the use of raw speech recognition or speech translation could prove to be an effective means to decrease interpreters cognitive load and improve performances. Here the equation between added and freed cognitive capacities needs to be experimentally evaluated.

Another major area of interest for applying AI to the interpretation domain is the automatic assessment of interpreter performances. Automatic assessment based on metrics such BLEU, METEOR or BERT, can be useful to support trainers, recruiters or in the scope of self-evaluation at the end of an event. The validity of automatic metrics is primarily dependent on a strong correlation with human assessments. Lu and Han (2022) demonstrated that such metrics have a moderate-to-strong correlations with the human-assigned scores across the assessment scenarios.

Language use and spoken language translation are however complex phenomena. While machines seem not to be able to evaluate interpretation from a communicative perspective, narrower type of assessments are possible. For example, it is conceivable to verify the adherence to a specific terminology, the accuracy of number rendition, or the omissions of consistent part of the source speech in a completely automatic way, generating a report for the interpreter to consume. This report could contain direct links between the transcription and the audio recording of the original speech and of the translation. At the moment of writing, no applications to automate the evaluation of interpreted speeches has been developed or presented to the public.

6. https://cai.uni-mainz.de/asr-pad/

6. Interpreter management systems

Interpreter Management Systems (IMS) are programs designed to streamline and make more efficient the process of scheduling and booking interpreters, especially inside of language service providers or international organizations. Their main goal is to match the right interpreters for an assignment, based on factors such as language combination, level of seniority, qualifications, domain expertise, previous work, temporal availability, and so forth. In a traditional approach to interpreter management, such criteria are matched by expert project managers using conventional database queries. AI-driven IMS can fully automatize this process or offer support to human managers in selecting the right interpreter.

Since one of the most successful applications of AI-driven systems is ranking, such systems have the potential to become a viable solution for managing interpreting projects, especially in contexts where the number of assignments and interpreters is very high. A management system based on machine learning is able to perform a selection using a higher set of criteria than humans, triangulating data with information and feedback on past events and improving its own selection algorithm over time. The ethical impact of the use of AI-driven systems is briefly introduced in Section 7.

7. Ehtics of AI use in interpreting

As in all areas of life, both private and professional, the use of AI raises ethical questions that need to be addressed. In this section, we limit ourselves to a short list of some of the most common ethical issues that every user of AI should be aware of in the interpreting profession. In particular, we will focus on confidentially of data and system bias.

The confidentiality issue is raised by the processing of data on the cloud and not on the user's computer. The reason is that most AI applications, especially in the language domain, still require enough computational power that is not available on consumer's devices. While great efforts have been made to reduce the size of language models and make them fit for use at the edge, today it is cloud systems that make it possible to handle the complexity of inferring a model and allow for scalability and fast response times. The use of web-based applications needs to be pondered as far as data confidentiality and privacy is concerned.

AI-driven tools are subject to algorithmic bias, i.e., a phenomenon which occurs when an algorithm produces results that are systemically prejudiced due to erroneous assumptions in the machine learning process. In interpreting a major role is played by tools processing spoken texts, for speech recognition or speech translation. Here the potential impact of bias is prominent. Not only has ASR quality issues with low resource languages or uncommon language variations, but there is growing evidence that performances deteriorate with underrepresented language features connected with ethnicity, gender, age, social background, etc. (cf. Koenecke et al. 2020, Feng et al. 2021), which can lead to impairments in the use of such tools and even to behavioral and psychological consequences of people affected by it (Mengesha et al. 2021).

Bias problems have been identified also in the use of AI in human resources (HR) applications (cf. Mujtaba and Mahapatra 2019). In this area, bias appears when AI screens and downranks applicants whose demographic traits – however irrelevant to the position – differ from those in the original data set. The automation of the selection process of interpreters, as introduced in Section 6, for example with the purpose to add interpreters to a company database or to match interpreters for a specific event, may suffer from such bias. The implementation of algorithms in this area should therefore be done under scrutiny of experience data scientists to avoid or neutralize gender or racial bias.

8. Conclusions

This chapter has provided an overview of the AI technologies that have entered the profession in recent years or will enter in the near future. While one noticeable impact of AI will come from the widespread use of automated means of translating spoken language, the use of AI-driven CAI tools has the potential to maximize the collaborative potential between artificial intelligence and the interpreters, supporting interpreters perform better and stay relevant in a changing professional world.

Historically, CAI technologies have had limited impact on the profession and most scholars, practitioners and developers have viewed them in isolation, for example as separate entities from mainstream applications such as remote interpreting tools. Chances are they will intermingle at any time soon and will become part components of a new interpreter ecosystem. So, for example, remote simultaneous interpreting platforms may integrate computer-assisted tools to make specialized terminology accessible to the entire team, mainstream web conferencing platform will integrate RSI, and so forth.

Technology is one of the main drivers of change in the professional environment. We are still in a transitional phase of this technologization process. Professionals are advised to stay abreast, approach these changes critically but openly and contribute where possible to shaping the future of their profession. Academics are recommended to intensify the study of the influence of technology on the interpreting process and on technology-mediated multilingual communication.

References

Amodei, D., Ananthanarayanan, S., Anubhai, R., Bai, J., Battenberg, E., Case, C., Casper, J., Catanzaro, B., Cheng, Q., Chen, G., Chen, J., Chen, J., Chen, Z., Chrzanowski, M., Coates, A., Diamos, G., Ding, K., Du, N., Elsen, E., Engel, J., Fang, W., Fan, L., Fougner, C., Gao, L., Gong, C., Hannun, A., Han, T., Johannes, L., Jiang, B., Ju, C., Jun, B., LeGresley, P., Lin, L., Liu, J., Liu, Y., Li, W., Li, X., Ma, D., Narang, S., Ng, A., Ozair, S., Peng, Y., Prenger, R., Qian, S., Quan, Z., Raiman, J., Rao, V., Satheesh, S., Seetapun, D., Sengupta, S., Srinet, K., Sriram, A., Tang, H., Tang, L., Wang, C., Wang, J., Wang, K., Wang, Y., Wang, Z., Wang, Z., Wu, S., Wei, L., Xiao, B., Xie, W., Xie, Y., Yogatama, D., Yuan, B., Zhan, J. & Zhu, Z.. (2016). Deep Speech 2: End-to-End Speech Recognition in English and Mandarin. Proceedings of The 33rd International Conference on Machine Learning Research 48: 173–182 Available from https://proceedings.mlr.press/v48/amodei16.html.

doi Anastasopoulos, Antonios, Ondřej Bojar, Jacob Bremerman, Roldano Cattoni, Maha Elbayad, Marcello Federico, Xutai Ma, et al.. 2021. "FINDINGS OF THE IWSLT 2021 EVALUATION CAMPAIGN." In Proceedings of the 18th International Conference on Spoken Language Translation (IWSLT 2021), 1–29. Bangkok, Thailand (online): Association for Computational Linguistics, 2021.

doi Bakaric, Marija Brkic, Nikola Babic, and Maja Matetic. 2021. "Application-Based Evaluation of Automatic Terminology Extraction." *International Journal of Advanced Computer Science and Applications* 12 (1): 18–27.

Baroni, Marco, and Silvia Bernardini. 2004. "BootCaT: Bootstrapping corpora and terms from the web". In *Proceedings of LREC 2004*, 1313–1316. Lisbon: ELDA.

Behr, Martina. 2013. "Zur Einbeziehung kommunikationswissenschaftlicher Erkenntnisse bei der Evaluation von Dolmetschleistungen." In *Quality in interpreting: Widening the scope.* Volume 1, edited by Olalla García Becerra, E. Macarena Pradas Macías and Rafael Barranco-Droege, 83–100. Granada: Comares.

doi Bender, Emily M., and Alexander Koller. 2020. "Climbing towards NLU: On Meaning, Form, and Understanding in the Age of Data." In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 5185–5198. Online: Association for Computational Linguistics.

Braun, Sabine, and Andrea Clarici. 1996. "Inaccuracy for Numerals in Simultaneous Interpretation: Neurolinguistic and Neuropsychological Perspectives." *The Interpreters' Newsletter* 7: 85–102.

Brown, Tom B., Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, et al.. 2020. "Language Models Are Few-Shot Learners." *ArXiv:2005.14165 [Cs]*, July 22, 2020.

doi Corpas Pastor, Gloria. 2022. "Technology Solutions for Interpreters: The VIP System." *Hermēneus. Revista de Traducción e Interpretación* 23: 91–123.

Corpas Pastor, G. and Mahmoud Gaber. 2021. "Extracción de fraseología para intérpretes a partir de corpus comparables compilados mediante reconocimiento automático del habla". In *Sistemas fraseológicos en contraste: Enfoques computacionales y de corpus*, edited by Gloria Corpas Pastor, Rosario Bautista Zambrana and C. Hidalgo-Ternero, 271–291. Granada: Comares.

doi Costa, Hernani, Gloria Corpas Pastor, and Isabel Durán Muñoz. 2016. "A Comparative User Evaluation of Terminology Management Tools for Interpreters." In Proceedings of the Workshop on Computational Terminology (CompuTerm'14), 68–76, 2014.

doi 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. International Journal of Translation Studies* 33 (1): 73–102.

Desmet, Bart, Vandierendonck, Mieke, and Bart Defrancq. 2018. "Simultaneous interpretation of numbers and the impact of technological support". In. *Interpreting and Technology*, edited by Claudio Fantinuoli, 13–27. Berlin: Language Science Press.

Devlin, Jacob, Chang, Ming-Wei, Lee, Kenton, and Kristina Toutanova. 2019. "BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding." *ArXiv:1810.04805 [Cs]*, May 24.

Fantinuoli, Claudio. 2006. "Specialized Corpora from the Web for Simultaneous Interpreters." In *Wacky! Working Papers on the Web as Corpus*, edited by Marco Baroni and Silvia Bernardini, 173–90. Bologna: GEDIT.

doi Fantinuoli, Claudio. 2017a. "Computer-Assisted Preparation in Conference Interpreting." *Translation & Interpreting* 9 (2): 24–37.

Fantinuoli, Claudio. 2017b. "Computerlinguistik in Der Dolmetschpraxis Unter Besonderer Berücksichtigung Der Korpusanalyse." In *Annotation, Exploitation and Evaluation of Parallel Corpora*, edited by Oliver Čulo, Slivia Hansen-Schirra and Stella Neumann, 111–46. Berlin: Language Science Press.

Fantinuoli, Claudio. 2017c. "Speech Recognition in the Interpreter Workstation." In Proceedings of the Translating and the Computer 39 Conference, London, 367–377. Geneva: Editions Tradulex.

Fantinuoli, Claudio. 2018a. "Computer-Assisted Interpreting: Challenges and Future Perspectives." In *Trends in E-Tools and Resources for Translators and Interpreters*, edited by Gloria Corpas Pastor and Isabel Durán Muñoz, 153–174. Leiden: Brill.

Fantinuoli, Claudio. 2018b. "The use of comparable corpora in interpreting practice and training. " *The Interpreters' Newsletter* 23: 133–149.

doi Fantinuoli, Claudio, and Bianca Prandi. 2021. "Towards the Evaluation of Automatic Simultaneous Speech Translation from a Communicative Perspective." In Proceedings of the 18th International Conference on Spoken Language Translation (IWSLT 2021), 245–54. Bangkok, Thailand (online): Association for Computational Linguistics.

Fantinuoli, Claudio, Marchesini, Giulia, Landan, David, and Lukas Horak. 2022. "KUDO Interpreter Assist: Automated Real-Time Support for Remote Interpretation." *ArXiv:2201.01800 [Cs]*, January 5, 2022. [Online] Available at http://arxiv.org/abs/2201.01800.

Fantinuoli, Claudio, and Maddalena Montecchio. 2022. "Defining Maximum Acceptable Latency of AI-Enhanced CAI Tools." *ArXiv:2201.02792 [Cs]*, January 8, 2022. [Online] Available at http://arxiv.org/abs/2201.02792.

Feng, Siyuan, Kudina, Olva, Halpern, Bence Mark, and Odette Scharenborg. 2021. "Quantifying Bias in Automatic Speech Recognition." *ArXiv:2103.15122 [Cs, Eess]*, April 1, 2021. [Online] Available at http://arxiv.org/abs/2103.15122.

Filippidou, Foteini, and Lefteris Moussiades. 2020. "A Benchmarking of IBM, Google and Wit Automatic Speech Recognition Systems. *Artificial Intelligence Applications and Innovations* 583: 73–82.

Floridi, Luciano. 2014. *The Fourth Revolution How the Infosphere Is Reshaping Human Reality*. Oxford: Oxford University Press.

Gile, Daniel. 2009. *Basic Concepts and Models for Interpreter and Translator Training: Revised Edition*. 2nd ed. Amsterdam: John Benjamins.

Goldsmith, Joshua. 2018. "Tablet Interpreting: Consecutive Interpreting 2.0." *Translation and Interpreting Studies* 13 (3): 342–65.

Haque, Rejwanul, Penkale, Sergio, and Andy Way. 2014. "Bilingual termbank creation via log-likelihood comparison and phrase-based statistical machine translation". In Proceedings of the 4th International Workshop on Computational Terminology (Computerm): 42–51.

Haque, Rejwanul, Hasanuzzaman, Mohammed, and Andy Way. 2019. "TermEval: an automatic metric for evaluating terminology translation in MT". In: CICLing 2019: the 20th International Conference on Computational Linguistics and Intelligent Text Processing, 7–13 Apr 2019, La Rochelle, France.

Jia, Hao, Shuqin Gu, Yangbin Shi, Xiangyu Duan, Zhongkai Hu, Yuqi Zhang, and Weihua Luo. "Bilingual Terminology Extraction from Non-Parallel E-Commerce Corpora." *ArXiv:2104.07398 [Cs]*, April 15, 2021. [Online] Available at http://arxiv.org/abs/2104.07398.

Jia, Ye, Tadmor Ramanovich, Michelle, Remez, Tal, and Roi Pomerantz. 2021. "Translatotron 2: Robust Direct Speech-to-Speech Translation." *ArXiv:2107.08661 [Cs, Eess]*, December 3, 2021. [Online] Available at http://arxiv.org/abs/2107.08661.

Jia, Ye, Ron J. Weiss, Fadi Biadsy, Wolfgang Macherey, Melvin Johnson, Zhifeng Chen, and Yonghui Wu. 2019. "Direct Speech-to-Speech Translation with a Sequence-to-Sequence Model." *ArXiv:1904.06037 [Cs, Eess]*, June 25, 2019. [Online] Available at http://arxiv.org/abs/1904.06037.

Kalina, Sylvia. 2007. "Microphone Off' – Application of the Process Model of Interpreting to the Classroom." *Kalbotyra* 57 (3): 111–21.

Kalina, Sylvia. 2020. "Quality in Interpreting." *Handbook of Translation Studies Online*. John Benjamins. Accessed March 12, 2022. [Online] Available at https://benjamins.com/online/hts/articles/qua2.

Karakanta, Alina, Papi, Sara, Negri, Matteo, and Marco Turchi. 2021. "Simultaneous Speech Translation for Live Subtitling: From Delay to Display." In Proceedings of the 1st Workshop on Automatic Spoken Language Translation in Real-World Settings (ASLTRW), 35–48. Virtual: Association for Machine Translation in the Americas, 2021. [Online] Available at https://aclanthology.org/2021.mtsummit-asltrw.4.

Koenecke, Allison, Nam, Andrew, Lake, Emily, Nudell, Joe, Quartey, Minnie, Mengesha, Zion, Toups, Connor, Rickford, John, Jurafsky, Dan, and Sharad Goel. 2020. "Racial Disparities in Automated Speech Recognition." *Proceedings of the National Academy of Sciences* 117 (14): 7684–89. [Online] Available at

Kucharska, Anna. 2009. *Simultandolmetschen in Defizitären Situationen. Strategien Der Translatorischen Optimierung*. Leipzig: Frank & Timme.

Lu, Xiaolei, and Chao Han. 2022. "Automatic assessment of spoken-language interpreting based on machine-translation evaluation metrics A multi-scenario exploratory study." *Interpreting*. (Online first article). [Online] Available at https://www.jbe-platform.com /content/journals/10.1075/intp.00076.lu

doi Mengesha, Zion, Heldreth, Courtney, Lahav, Michal, Sublewski, Juliana and Elyse Tuennerman. 2021. "I Don't Think These Devices Are Very Culturally Sensitive.' – Impact of Automated Speech Recognition Errors on African Americans." *Frontiers in Artificial Intelligence* 4: 725911.

doi Mujtaba, Dena F. and Nihar R. Mahapatra. 2019. "Ethical Considerations in AI-Based Recruitment." In 2019 IEEE International Symposium on Technology and Society (ISTAS), 1–7. Medford, MA, USA: IEEE, 2019.

doi Pisani, Elisabetta, and Claudio Fantinuoli, 2021. "Measuring the Impact of Automatic Speech Recognition on Interpreter's Performances in Simultaneous Interpreting." In *Empirical Studies of Translation and Interpreting: The Post-Structuralist Approach*, edited by Wang Caiwen and Zheng Binghan, 181–197. New York: Routledge.

doi Pöchhacker, Franz. 2002. "Quality Assessment in Conference and Community Interpreting." *Meta* 46 (2): 410–25.

doi Pöchhacker, Franz. 2016. *Introducing Interpreting Studies*. Manchester: Routledge.

Prandi, Bianca. 2017. "Designing a Multimethod Study on the Use of CAI Tools during Simultaneous Interpreting." In Proceedings of the 39th Conference Translating and the Computer, 76–88. Geneva: Tradulex.

Prandi, Bianca. 2022. Computer-Assisted Simultaneous Interpreting: A cognitive-experimental study on terminology". PhD thesis University of Mainz.

Romero-Fresco, Pablo, and Franz Pöchhacker. 2017. "Quality assessment in interlingual live subtitling: The NTR model." *Linguistica Antverpiensia, New Series: Themes in Translation Studies*, 16, 149–167.

Rütten, Anja. 2003. "Computer-Based Information Management for Conference Interpreters Or How Will I Make My Computer Act Like an Infallible Information Butler?" In Proceedings of the International Conference Translating and the Computer. London, 2003. [Online] Available at http://www.mt-archive.info/Aslib-2003-Rutten.pdf.

Rütten, Anja. 2007. *Informations- Und Wissensmanagement Im Konferenzdolmetschen*. Frankfurt am Main: Peter Lang.

doi Schild Ortiz, Luís Eduardo, and Patrizia Cavalli. 2018. "Computer-Assisted Interpreting Tools (CAI) and Options for Automation with Automatic Speech Recognition." *TradTerm* 32: 9–31.

doi Setton, Robin, and Andrew Dawrant. 2016. *Conference Interpreting: A Complete Course*. Amsterdam and Philadelphia: John Benjamins.

doi Sperber, Matthias, and Matthias Paulik. 2020. "Speech Translation and the End-to-End Promise: Taking Stock of Where We Are." *ArXiv:2004.06358 [Cs]*, April 14, 2020. [Online] Available at http://arxiv.org/abs/2004.06358.

doi Steurs, Frieda, and Katarzyna Tryczynska. 2021. "European Labour Law and its Challenges in Multilingual Terminology and Translation: A Case Study." *Sendebar* 32: 219–237.

Stoll, Christoph. 2009. *Jenseits Simultanfähiger Terminologiesysteme*. Trier: Wvt Wissenschaftlicher Verlag.

Sudoh, Katsuhito, Kano, Takatomo, Novitasari, Sashi, Yanagita, Tomoya, Sakti, Sakriani, and Satoshi Nakamura. 2020. "Simultaneous Speech-to-Speech Translation System with Neural Incremental ASR, MT, and TTS." *ArXiv:2011.04845 [Cs]*, November 11, 2020. [Online] Available at http://arxiv.org/abs/2011.04845.

Tiselius, Elisabet. 2009. "Revisiting Carroll's Scales." In *American Translators Association Scholarly Monograph Series*, edited by Claudia V. Angelelli and Holly E. Jacobson, XIV: 95–121. Amsterdam: John Benjamins.

Will, Martin. 2009. *Dolmetschorientierte Terminologiearbeit. Modelle Und Methode*. Tübingen: Gunter Narr Verlag.

Xu, Ran. 2018. "Corpus-Based Terminological Preparation for Simultaneous Interpreting." *Interpreting* 20 (1): 29–58.