36

Conference interpreting and new technologies

Claudio Fantinuoli

Introduction

Conference interpreting has traditionally been viewed both by the general public and by practitioners as an activity that is not particularly related to technology. Yet, as in many other professions, various technologies have played a crucial role in shaping the way interpreters work today. The history of technological transformation is long and dates back to the very birth of the profession.

It is common knowledge to define technology as the branch of knowledge that deals with the creation and use of technical means to support people in performing specific tasks. In this sense, technology has been part of conference interpreting from the very beginning. Already in its most traditional form, the consecutive modality, interpreting has required tools, albeit simple ones, such as paper, pen and an annotation system, to allow for a complete and precise rendition of long speeches. However, it is in the 1920s that more advanced technological innovations started to play a central role in the profession. In those years, wired systems for the transmission of sound allowed the transformation of a mode of interpreting—*chuchotage*—in the mainstream way to deliver oral translation at multi-language conferences. The simultaneous modality, faster and less cumbersome for its users, fitted perfectly into a world characterized by increasing international contacts and by an ever-growing acceleration of human experiences. Technology, therefore, radically changed interpreting as an activity, and contributed to establish it as the profession we know today.

More recently, the digitization and digitalization processes occurring in our society are once again changing the landscape of most professions (e.g. Susskind & Susskind 2017). As is often the case, many changes that took place inside a particular profession were not caused by technologies designed to satisfy its peculiar needs, but were brought by innovations that did not target specific professions, such as email, video-conferencing, and the like. Yet, nonetheless, they had a profound impact on professions, for example in accelerating the pace of work organization (e.g. Rosa 2015). In many cases, at first, its members did not even notice the far-reaching transformation that such changes were causing, mainly because such transformations took time to unfold. Only in the retrospective did the extent of these changes appear crystal clear to the keen eye.

DOI: 10.4324/9780429297878-44

We can look at information and knowledge as an example. Information-creation, information-seeking, and information-dissemination habits have been radically changed by information technology and the internet (Susskind & Susskind 2017). Of course, interpreting has not been immune to this change. The wealth of multilingual and specialized information that became available in digitalized form with the surge of the internet profoundly modified interpreters' habits in terms of how and when they acquired and accessed knowledge. In this respect, it is plausible to think that the technological transformation introduced above has challenged some long-standing assumptions about interpreting and interpreters. In the context of an increased specialization of multilingual events, the possibility of retrieving a potentially infinite number of preparatory texts at any time and virtually everywhere or to access comprehensive multi-language resources may have shifted the assumption of the interpreter as an "omniscient" person with a vast stock of encyclopaedic and world-knowledge to draw upon (e.g. Pöchhacker 2016; Riccardi 2000) to a professional that on top of a broad and eclectic general knowledge has the skill to 'prepare' for a specific topic or event. The consequences of this shift brought by technological change for knowledge, memory and cognition are possibly overwhelming, but exactly how this has affected the way we deal with information has not been subjected to systematic scrutiny so far (Fantinuoli 2019).

Recently, the digitalization of professional work has also motivated companies to develop new tools specifically designed to fit the various workflows related to interpreting. The deployment of technology in the context of this profession has advanced at a much faster pace than in the past, and the technologies involved have diversified enormously. It is therefore useful to operate a preliminary distinction among the several areas of technology that are directly connected with conference interpreting. Broadly speaking, it is possible to identify at least four main areas.

In the first area, we find digital tools to enhance the training process, the so-called *computer-assisted interpreting training tools*. In the second area, we identify tools that have been developed to augment the interpreting process, i.e. to improve the preparation and performance of interpreters. They are generally called *computer-assisted* or *computer-aided interpreting tools*. The third area of interest is *remote* or *distance interpreting*. In this area, solutions have been proposed to allow new ways for the delivery of interpreting services, both in consecutive and simultaneous mode, by means of audio-video links. Finally, the last area comprises emerging solutions to completely automatize the interpreting process. This is the realm of *machine interpreting*.

It goes without saying that all these areas are impacting in one way or another the profession, determining a shift in the way interpreting is performed, delivered and perceived by stakeholders and the general public. After years of a somewhat irrational aversion (e.g Pym 2011) to technological transformation,[1] the general attitude of interpreters to such innovations has recently changed, as recent technical reports and position papers (e.g AIIC 2018; European Commission 2019) and, more generally, the wider adoption of such technologies among professionals seem to confirm. Unexpected events, such as the Covid pandemic in 2020, have undoubtedly helped to accelerate this process. The magnitude of change that this will bring by, however, is certainly difficult to predict (Fantinuoli 2019).

The next sections will provide a brief overview of the above-mentioned technologies. The remainder of this chapter is organized as follows: the next section will focus on technologies used in interpreter training, then the tools to augment the interpreting process are discussed, and finally the topic of machine interpreting is introduced. The topic of remote interpreting will be discussed in a separate chapter (see Seeber & Fox, Chapter 35, in this volume).

Technology in educational and training contexts

One of the areas in which digital technologies have been widely recognized as extremely beneficial by both researchers and practitioners is training (see Kalina & Barranco-Droege, Chapter 24, in this volume). Over the last decades, several projects have been devoted to the development of digital means to support interpreter training and to experimentally evaluate its effectiveness. Some of the tools are still used in formal training at universities or in self-study and have become de facto an integral part of the interpreting training landscape. Tools specifically developed in this area, such as speech repositories (see below) or authoring tools (see below), are commonly referred to as computer-assisted interpreter training (CAIT) tools.

The rationale behind CAIT is straightforward: exploiting the multimedia capabilities of Information and Communication Technology (ICT) to enhance the teaching and learning of interpreting in various ways, such as incrementing students' motivation, improving training workflows, allowing distance teaching, and so forth (e.g. Mayor & Ivars 2007). This is particularly important in the light of several facts. On the one hand, training has moved from a teacher-centered towards a learner-centered approach, placing strong emphasis on students' autonomous practice and group work on the premises of socio-constructivist teaching (e.g. Fantinuoli & Prandi 2018; Stengers et al. 2018; Vygotsky 1978). In such a context, CAIT tools may be used to overcome several limitations of traditional training, such as the fact that students rarely do have access, for example, for self-study, to suitable learning support or appropriate materials, meaning that the work they "do in their self-study hours is often unstructured and unmonitored" (Sandrelli 2002: 190). Furthermore, they have been suggested as a possible solution to overcome the lack of situated learning activities. For example, they are well suited to support the creation of exercises that simulate the real conditions of professional life. On the other hand, CAIT tools may be the solution of choice to enable distance learning, both as a complementary form of delivering training next to the classical in-situ approach as well as a complete on-line alternative, such as practised by on-line universities or training institutions.

Empirical research has revealed that aside from streamlining the teaching process, CAIT tools may alleviate the anxiety of performance typical of interpreting students, facilitate self-paced and self-monitored practice and encourage independent practice among students (e.g. Deysel & Lesch 2018; Gorm Hansen & Shlesinger 2007; Kerremans & Stengers 2017).

CAIT tools can be tentatively grouped in at least four main areas (e.g. Sandrelli & Jerez 2007): (1) speech repositories; (2) authoring tools; (3) immersive digital learning environments; and (4) remote teaching platforms.

Speech repositories

The most widespread form of digital tools to enhance interpreting training are speech repositories. They provide trainers and students with suitable materials for classroom use and self-study. Such repositories are not random collections of audiovisual speeches, but databases containing texts classified according to various criteria, such as topic and complexity. A notable example is the Speech Repository,[2] a database of speeches founded and managed by the interpreting service of the European Commission. It contains selected speeches in all languages of interest for the European institutions collected or recorded by professional interpreters and stored in a searchable on-line database. Besides recordings of public conferences, press conferences and parliamentary debates, it contains speeches given specifically for training purposes by EU interpreters and university trainers. Speeches are categorized according to language, level (from basic to very advanced), intended use (consecutive or simultaneous interpretation), domain and type (press conference, debate, pedagogical material), etc.

Besides the wealth of real-life and pedagogical material, the added value of such a platform lies in its categorization mechanisms, vetted by experienced trainers, that enormously simplify the selection of the right speech for the right task. Similar services, even if smaller in scale, are offered by the user-generated content of Speechpool[3] or by initiatives inside university programmes, such as the *Leeds EN IT Interpreting Practice Sessions*.[4] Other projects often mentioned in the literature, such as *Marius* of the University of Granada or the IRIS database (Carabelli 1997), seem to have been discontinued.

Digital communities to peer review interpreter renditions have brought a dynamic turn to the traditional service offered by classical speech repositories. This kind of web-based platforms is in line with the deliberate practice approach (see, for example, Tiselius 2013). Instead of speeches for training, users can find other community members who are available to listen to the rendition and give feedback. This is the model proposed for example by InterpretimeBank,[5] a private platform designed by professional interpreters.

Authoring programmes

The second area comprises interpreting-oriented authoring programmes. The goal of such programmes is to enable interpreter trainers to create various types of exercises to provide trainees with tools to optimize the use of the available resources. Such CAIT applications incorporate utilities to increase interaction between the computer and users and to situate learning in more realistic contexts. Only a few tools have been developed in the past, and the majority of them have been discontinued; among others *Interprit* (Merlini 1996), *Interpretations* (Sandrelli 2003) and its successor *Black Box* (Sandrelli & Hawkins 2006).

Despite the fact that such tools proved useful, the limited interest by the training community caused the discontinuation of many projects. However, in the last decade, the exponential growth of the amount of on-line material that can be used for training purposes has revamped the interest in the creation of ad-hoc training materials by means of general purpose tools, such as video player or video editing software. In the case of radio interviews, for example, audio software such as Audacity[6] has been proposed to be used to remove interpreters' utterances from the broadcast, leaving silence gaps for students to interpret, thus replicating a live situation (Castillo 2019). Simplifying the incorporation of the oral dimension and, in the case of video-recording, of features such as non-verbal language and the setting where the communicative event takes place, this kind of CAIT can broaden the spectrum of speeches used in training (de Manuel Jerez 2006). More recently, a renewed interest in this area has brought companies to develop new products targeting the domain of interpreting training, such as the InterpreterQ Media Player.[7] This tool brings together video editing and performance evaluation features and can be used by both teachers as by trainees.

Virtual learning environments

The third approach to CAIT is based on *virtual learning environments* and seeks to exploit the opportunities offered by computer-mediated communication tools, making the teaching and learning of interpreting more immersive (e.g. Braun & Slater 2014; Kerremans & Stengers 2017; Şahin 2013). This approach is grounded in the widely held belief that learning must be situated in real-world contexts to maximize its effectiveness. The use of multimedia-based CAIT has been claimed to situate learning in a more "realistic" learning environment

(Sandrelli & Jerez 2007), reducing the distance between professional interpreting contexts and the slightly artificial dimension of classroom training. In particular, immersive interfaces, such as those developed for flight simulators, may contribute to enhancing the sense of authenticity (e.g. Sandrelli & Hawkins 2006). A commercial example is the Virtual Interpreting System *VIS – 5.6*.[8]

Remote teaching platforms

Remote teaching platforms or distant learning solutions are applications designed to facilitate student learning in a context where their physical presence is not possible or requested. Their importance has grown exponentially after the outbreak of the Covid pandemic 2020 when most training institutions around the world needed to switch to online teaching. At the time of writing, only one application has been designed specifically for the context of conference interpreting. InTrain[9] is a free and open-source web-based platform for remote simultaneous interpreter training developed at the University of Bologna. Because of the lack of specific applications, institutions and trainers involved in remote teaching of simultaneous and consecutive interpreting generally resort to general video conferencing applications, such as Zoom,[10] hubs for team work, such as Microsoft Teams,[11] which includes chats, video conferencing, file share, etc., online application for language teaching, such as Sanako Connect,[12] or remote interpreting platforms, such as Voiceboxer,[13] Interprefy[14] or KUDO.[15] Depending on the grade of synchronicity of the desired approach to remote training, digital learning management systems, such as Moodle[16] are widely used.

Computer-assisted interpreting

Computer-assisted or computer-aided interpreting (CAI) is defined as a form of speech translation where a human interpreter is supported by computer software developed to facilitate some aspects of the interpreting task, from preparation to information access (e.g Fantinuoli 2018). The instruments used to support and augment the interpreter's work are called CAI tools. In this context, CAI tools are all sorts of computer programs and mobile applications specifically designed and developed to assist interpreters in at least one of the different sub-processes of interpretation, for example, knowledge acquisition and management, lexicographic memorization, terminology access, and so forth.

The number of software programs available is limited and comprises three major tools: InterpretBank,[17] Interpreter's Help[18] and Interplex.[19] They greatly differ from each other both in the set of functionalities offered and in their architecture (cloud-based, desktop or mixed). Since tools are evolving at a very fast pace and characteristics and features are due to become obsolete very quickly, only little reference to specific products will be made here. Instead, the basic principles underlying the CAI tools will be introduced.

Computer-assisted event preparation

Interpreters are called to interpret many different specialized topics for which they are not expert or do not have any specific qualification. For this reason, event preparation, both at a linguistic (e.g. terminology and phraseology) and extra-linguistic level (e.g. domain and event-related knowledge), has been described in literature as one of the most important phases of an interpreting assignment (e.g. Díaz-Galaz 2015, Gile 2009). Advance preparation aims at bridging the linguistic and extra-linguistic gap between the conference participants and interpreters (Will 2009). In doing so, it reduces the cognitive load during the interpreting task as it anticipates parts of it in the preparatory phase (Fantinuoli 2017; Stoll 2009), allowing interpreters to manage the process more efficiently and deliver a higher interpretation quality (Díaz-Galaz 2015). Yet, preparation is generally time-consuming and interpreters often experience the feeling of not knowing exactly how to perform this task efficiently (e.g. Moser-Mercer 1992).

To cope with this, computer-assisted approaches to event preparation have been specifically designed to help interpreters rationalize the process. Several functionalities aimed at streamlining preparation have been implemented for CAI tools, especially in the area of terminology acquisition. CAI tools support, for example, the automatic look-up of terminology resources, such as IATE[20] or Wikimedia,[21] during the compilation of new glossaries; the possibility of manually extracting terminology from parallel documents; the automatic extraction of a list of specialized terms from monolingual preparatory documents; flash-card systems to support the memorization of specialized terminology before the event, and so forth.

Only few empirical studies have been conducted in this area so far. Tests have demonstrated that automatic methods are able to collect preparatory resources (Fantinuoli 2012; 2018) and to extract relevant terms (Fantinuoli 2006; Xu 2018; Xu & Sharoff 2014) in line with interpreter's needs. Xu (2018) has empirically demonstrated that the use of the corpus-based interpreter preparation approach (Fantinuoli 2006), which combines automatic terminology extraction with concordancer features, may improve the interpreter's performance.

Nowadays interpreters, both freelancers and staff members, have to cope with the need to process an ever-increasing number of documents and information in less time. In order to deal successfully with this challenge, tools have been developed that streamline the management of terminology and event-related documents. The number and complexity of available functionalities may vary, but generally they comprise at least a terminology database and a documents management feature.

Different to translators' databases, terminology solutions for interpreters have a simpler terminology structure, do not require any particular set-up, and organize data in glossaries (ordered by subject or event). They may integrate a translation suggestion feature to be used while compiling the glossary, and managing functions, such as duplicate detection, spelling check, etc. As introduced in the previous section, the available solutions for terminology management can be cloud-based, with the clear advantage of not requiring any software installation on the local computer, or classical desktop tools, with the advantage of a higher level of data protection.

Document management features of CAI tools vary enormously depending on whether the tool is intended for freelancers, working alone or in small teams, or for language service providers or international institutions. Language service providers, in fact, have particular needs in terms of workflows and require advanced features to organize and make available documents to a large number of interpreters. Freelancers, on the other hand, seem to be less in need of advanced document management solutions. Both of them, however, derive numerous advantages from processing preparatory texts for the creation of new glossaries. In this context, features such as the possibility of extracting terminology automatically or semi-automatically from the texts seem to be considered a useful feature for most interpreters (e.g. Goldsmith 2020).

Automatic terminology extraction from a small set of documents is available, at the time of writing, only in a monolingual context. Fully automatic bilingual terminology extraction from small parallel corpora (the preparatory texts available to the interpreter) seems difficult to achieve. For this reason, CAI tools resort to manual automatic extraction from parallel texts, in some cases combining it with hybrid approaches, for example, the use of web terminology repositories as a source of translation suggestions. The use of modern machine learning approaches, however, may prove useful in future to improve bilingual terminology extraction from short texts.

Information retrieval in the booth

All commercial CAI tools are equipped with fast querying functionalities to support the search for specialized terminology in the booth. This functionality is designed as a back-up strategy when other interpreting strategies, such as paraphrasing or the use of a synonym, are not viable and could possibly lead to miscommunication or a general degradation of the performance. Interpreters may look up a term while interpreting, while helping the boothmate or simply during the pauses, perhaps to find the translation of a recurring term used in the previous speech.

Looking up a term in the booth while interpreting has been criticized as being “unnatural” (e.g. Tripepi Winteringham 2010), especially because of the time limits characteristic of simultaneous interpretation and the risk of cognitive overload (Gile 2009), or simply because it may distract the interpreter. New research, however, seems to indicate that, in certain circumstances, looking up specialized terms may contribute to improve performance (e.g. Biagini 2016; Prandi 2015, 2017). Yet, the best conditions in terms of speech delivery rate (words per minute) or terminology density, to name just a few, that would allow interpreters to easily use a CAI tool in the booth have not been researched so far, and much is left to the interpreters’ experimenting in their daily working life.

While CAI tools have been designed ergonomically in order to reduce the cognitive effort needed to start a search and to read the results, the integration of automatic speech recognition to automate the look-up process may further increase the usability of such solutions (see below).

Support in the consecutive modality

Outside the interpreting booth, computers have entered the scene only recently. In the context of consecutive interpreting, digital tools, and in particular tablets, have been used instead of paper and pen to perform note-taking (e.g. Goldsmith 2017; see also Ahrens & Orlando, Chapter 3, in this volume) and for other activities, such as access to the Internet, look-up of terminological databases, reading of documents, and so forth.

The rationale behind the use of tablets is straightforward: they do not take up much space and are therefore ideal for mobile use, they are silent, easy to use and have a long battery life. At the time of writing, no application has been specifically created for the interpreters’ note-taking activity. For this reason, note-taking is usually performed with general purpose software, generally using stylus pens as input device.

Mainstream tablets (iOS and Android) are not full-blown computers. Although the features offered by the use of a tablet for consecutive interpreting outstrip the functionalities offered by traditional pen-and-paper methods (e.g. Goldsmith 2018), the limitations of the hardware and the available software should not be ignored. In fact, such limitations reduce the spectrum of other interpreting-related applications usually available on normal desktops. Beyond note-taking, the success of tablets in the context of interpreting relies primarily in the use of general features, for example, to find term definitions on the Internet, look up translations on on-line databases or to find contextual information about the event, to name just a few. Some glossary management systems are available on Android and iOS tablets, too. However, they only offer some basic functionalities. Advanced features, such as automatic terminology extraction, are available only in the full-fledged computer versions or on tablets running desktop OS.

The advantages and disadvantages of tablets have not been empirically studied so far. The ability to take notes digitally may offer a number of advantages compared to traditional pen and paper, yet it may require a different approach to interpreter preparation and performance. For this reason, beyond the introspective analyses of practitioners conducted by Goldsmith (2018), this area of study still remains underexplored and urgently needs to be put on the agenda of academic research.

Simultaneous consecutive interpreting

Another area of possible use of digital devices is the so-called simultaneous consecutive interpreting. Studies have been conducted on the possible development of a hybrid mode of interpreting that mixes both conventional modalities of consecutive and simultaneous interpreting (e.g. Ferrari 2001; Orlando 2014). The proposed modality is based on recording a speech that would normally be rendered in consecutive mode, playing it back on headphones and rendering it simultaneously. In this way the original message in the interpreter's notes and long-term memory is replaced by a digital recording. The playback can be altered if necessary, for example it can be slowed down for particularly difficult passages. Related research suggests that this approach resulted in overall better interpreting performance, for example, in "more fluid delivery, closer source-target correspondence, and fewer prosodic deviations" (Hamidi & Pöchhacker 2007: 276) or in a higher level of performance Orlando (2014). The fact that this hybrid modality has not been taken up by the community may be related to the fact that while this modality is objectively more complete and precise, the rendition seems too artificial. Pauses, hesitations, in-line corrections, redundant elements, etc. of the original speech are generally maintained in the rendition (Orlando 2014). This may cause the rendition to be imperfect too, and this in a way that appears unnatural compared to classical consecutive. Further reasons that have been given are related to lack of confidence in the new technology and in the ability to use it in any situation, especially under stress (Gillies 2019).

In the context of consecutive training, digital pens (see Ahrens & Orlando, Chapter 3, in this volume) have been proposed to support evaluating the progressive acquisition of note-taking systems and skills. A digital pen is a device that offers audio and visual feedback, memory for handwriting capture, and audio recording. While previous research on note-taking focused solely on the final product of the notes, thanks to this technology, trainers have the possibility of capturing simultaneously the video of the notes being taken and the audio of the speech, and therefore of providing better advice and remedial strategies to their students (Orlando 2010; 2015). In this context, Chen (2017: 4) shows how pen distance, duration and speed can be recorded; these recordings "not only tell us what interpreters' note-taking choices are, but also … how interpreters carry out those choices".

Machine learning in the interpreter workstation

Interpreting-related technologies, such as CAI tools, may be augmented by means of integrating machine learning techniques. Machine learning (ML) is an area of artificial intelligence that designs and develops systems with the ability to automatically learn and improve from data and experience without being explicitly programmed. ML algorithms have been proved to achieve unprecedented results in many areas, such as vision systems, machine translation, etc., with an increasing number of applications already in many domains of daily life. In relation to interpreting, there are several areas where ML has started to enter the stage and where it may play a central role in the years to come.

One ML-based technology that has recently attracted the interest of the profession is Automatic Speech Recognition (ASR). ASR has been regarded as a technology "with considerable potential for changing the way interpreting is practiced" (Pöchhacker 2016: 188). For example, it could change the way consecutive interpreting is traditionally performed, outperforming alternative technology-based approaches proposed in the past, such as the digital pen (see above). With ASR, the consecutive interpreter may use the transcription of the spoken word to sight-translate the speech segment, with possible advantages in terms of precision and completeness. In simultaneous interpreting, it could offer real-time suggestions for well-known problem triggers, in particular, terminology, numbers, and named entities (e.g. Fantinuoli 2017). Different to classic CAI tools that require a manual input to get a translation for a given terminological unit, an ASR-enhanced CAI tool[22] is able to automate this process, with obvious advantages at the level of human-machine interaction.

State-of-the-art automatic speech recognition engines leverage the power of machine learning techniques, and in particular of deep neural networks, to achieve unprecedented levels of precision and to offer a speaker-independent transcription of continuous speech (Yu & Deng 2015). As far as the problem triggers mentioned above are concerned, recent studies have demonstrated that transcriptions already achieve—at least for standard speeches—very promising levels of precision for numbers and for terminology, while, for named entities, the results are still not satisfactory (Brüsewitz 2019). The increasing availability of massive computing power and data to train the language models, however, is eliminating many of the constraints of ASR. The integration of more layers of elaboration beyond the extraction of acoustic features, the identification of entries in the lexicon and a check against syntactic constraints, may further increase quality, especially in borderline situations (strong accents, non-native language features, speech-impaired people, out-of-lexicon, etc.).

The first product-oriented studies seem to indicate that ASR-enhanced CAI tools increase the interpreter's performance in the rendition of problem triggers. For example, empirical experiments conducted both with mock-up systems and with real-life prototypes have provided evidence for an improvement of the overall accuracy on numbers (up to 30 per cent) and for a reduction in the number of errors (Canali 2019; Defrancq & Fantinuoli 2020; Desmet et al. 2018; Prandi 2015) Although results are certainly encouraging, the studies have also stressed possible drawbacks, such as the risk of relying too much on suggestions, a sense of disorientation, etc. Further experiments are needed to extend the evaluation from the rendition of the triggers to the whole textual dimension of the rendition (Prandi 2017), to study the effect of providing specific training for interpreters (e.g. Defrancq & Fantinuoli 2020) and to measure the performance increase in relation to the level of interpreting proficiency (students vs. professionals) (e.g. Desmet et al. 2018).

One of the main limitations of ASR-enhanced CAI tools is that they show suggestions in a non-selective way. In the case of terminology, for example, they show all terminological units saved in a database that match the transcription. This has several disadvantages. First of all, the interpreter may tend to concentrate too much on the potential abundance of (unfiltered) suggestions, with the possible consequence of being distracted, may experience a cognitive overload, and ultimately decrease the overall quality of the rendition. Second, this approach requires the interpreter to prepare a glossary beforehand. If a term is not contained in the glossary, no translation will be displayed.

With this in mind, Vogler et al. (2019) have proposed using ML to anticipate the textual units that may cause difficulties for the interpreter and limit the number of suggestions only to these cases. The proposed approach is based on the comparison of a parallel corpus of translated and interpreted speeches in order to automatically identify the text features that led to a terminological issue (e.g. omission) in the interpreted rendition. This approach is based on an ML-augmented corpus-based analysis and represents one of the first attempts at process-oriented research in technology-mediated interpreting (see Mellinger (2019) on the necessity to integrate product-oriented with process-oriented research in the context of interpreting and technology).

An important aspect in the context of ASR-enhanced CAI tools, especially if ML-augmented, is latency, i.e. the time delay of suggestions related to the related utterance. If server-based real-time transcription, and the extraction of numbers from it, seem to be performed within the typical ear-voice-span (EVS) of simultaneous interpreting (e.g. Plevoets & Defrancq 2018), ML-enhanced extraction of terminology may increase the latency and make suggestions unusable for the interpreter. Terminology extraction may require transformations in order to match values from a database (e.g. lemmatization), and ML-enhanced identification of issues in real-time text may require too much context after the unit of interest. In this context, empirical studies should be performed in order to find out the maximum latency that an ASR-supported CAI tool should have in order for interpreters to profit from its suggestions.

ASR is not the only area where machine learning is entering the interpreter's workstation. Another area of development is information management. The so-called 'information overload' which is typical of our digital age may be reduced by means of intelligent algorithms able to systematize and personalize the abundance of information available online. They will filter and prepare information in a semi or fully automatic way, learning from users' past experience, and presenting to the interpreter ready-to-digest glossaries and preparatory information.

Machine interpreting

Machine interpreting (MI), also called automatic speech-to-text translation, speech-to-speech translation or spoken language translation, is an emerging area of natural language processing that aims at building machines that are able to translate, both in consecutive and simultaneous mode, spoken texts from one language to another. Different to all other interpreting-related technologies which are supportive of human-centric interpreting, MI aims at reducing language barriers by means of oral language translation performed in an unsupervised manner, making the human interpreter, at least in some areas and settings, de facto obsolete.

Machine interpreting has seen exceptional performance improvements over the past few years. The discipline started from the rather artificial problem of translating oral utterances recorded "under controlled conditions, with restricted vocabularies, strong domain limitations and the necessity of a constrained speaking style" (Paulik & Waibel 2009: 455). Nowadays, however, research as well as experimental and commercial applications are moving to the more ambitious task of translating real-life spoken language, without any particular constraints. The impressive advances made in this area can mostly be attributed to the use of modern machine learning algorithms, especially neural networks, that dominate the latest developments and the availability of big data (at least for some languages). Notwithstanding such remarkable progress, at the time of writing, no single application has found wide adoption, even in the more forgiving recreational segments. Some of the reasons for this will be introduced in the remainder of this section.

There are at least two approaches to machine interpreting: the cascading and the end-to-end. The former is based on a cascading system where the process of oral translation is broken down into some well-defined sub-processes that can be modelled in computer programs.

The classic approach to tackle this task consists in training a cascade of three separate components: automatic speech recognition (ASR) and machine translation (MT), for speech-to-text systems, as wells as text-to-speech synthesis (TTS) for speech-to-speech systems. This approach is used, for example, in the Google Translate App. These components are generally applied one after the other in a process called cascade translation, where the output of a process becomes the input of the next one. Dividing the task into such a cascade of systems has some obvious advantages: it builds on available technologies, it is transparent as far as the sequential nature of the tasks is concerned and, last but not least, it is extensible: new components can be added to the basic sequence introduced above. However, cascading systems suffer from several shortcomings, such as error propagation, the use of translation models designed for the written language that, consequently, do not model phenomena that occur in spoken language, such as hesitations, or the absence of reliable punctuation produced by ASR, which causes problems to the MT, and so forth. As a consequence, a noisy transcription provided by the ASR (e.g. because of typical ASR errors, such as wrong disambiguation of homophones, or performance errors, i.e. disfluencies) may impair the successive MT process (e.g. Ruiz et al. 2017).

The end-to-end approach applies similar machine learning techniques used for MT or ASR to bilingual speech data, i.e. to original speech and the translated speech in order to create models that are able to directly translate speech from one language into speech in another language, without relying on an intermediate text representation, i.e. the transcription (e.g. Berard et al. 2016; Di Gangi et al. 2018; Jia et al. 2019). Since systems do not divide the task into separate steps, they may provide a few advantages over the cascading system described above, including faster speed, avoiding compounding errors between recognition and translation, better handling of proper names, and so forth. As such a technique is still in its infancy, the output quality has only recently reached the one offered by conventional cascade systems.

Despite the quality improvements reported recently, machine interpreting still has to cope with many challenges of real-life multilingual communication in order to reach a quality level similar to that of professional interpreters. Among others, MI still operates primarily at a linguistic level. Typical aspects of verbal communication, such as inference, interpretation of prosodic features, correction of imperfect speech, pragmatics, to name just a few, are missing from all approaches. As far as the cascading systems are concerned, interpreting systems may be improved by increasing the precision of the three basic components (ASR, MT, STT), by adopting computational models that have been trained on real-life spoken communication, and by adding new layers of elaboration, for example, speaker diarization (the identification of an individual person based on characteristics found in the unique voice qualities), emotion recognition from linguistic and paralinguistic properties of speech, better source speech segmentation, to name just a few.

Other technologies that may add a true understanding of language, situation, communication goals, etc. are still far to come. However, even if there were not significant advances in such areas in the next decade, the exploitation of existing and emerging technologies may lead nonetheless to the use of MI in real-life scenarios. Depending on context and users' expectations, even systems that do not reach such a level of complexity may be good enough for some form of multilingual communication. If this proves to be true, a mixed scenario where both machines and humans will deliver interpretation may become a reality.

The intrinsic characteristics of some conference interpreting settings, such as the nature of monological texts (formality, repetitiveness, availability of training material) and the lower level of interaction in the communication process (reduced presence of changing verbal behaviour, turn-taking, asymmetries) may make machine interpreting especially suitable to be deployed, at least to some extent, in conference settings. General consequences on the labour market, for example, on the demand for human interpretation in lower segments of the market (in terms of demand for quality, prestige, etc.) are tentatively described in Fantinuoli (2019).

Conclusion

This chapter has presented a general overview of the technologies that have entered the profession in recent years. Although such technologies were considered in the past as separate and independent entities by most scholars, practitioners and developers, chances are they will start to combine any time soon, and become subcomponents of an overall changing interpreting ecosystem. So, for example, remote simultaneous interpreting platforms may integrate computer-assisted tools to make specialized terminology accessible to the entire team, commissioning platforms will integrate RSI web applications, and so forth.

We are still in a long transitional phase of this process of technologization. New technologies are entering the stage as a consequence of macro-economic changes, or because of unexpected situations, like the global lockdown following the COVID-19 pandemic of 2020. These events are imposing adaptations in many areas of the profession, from the need to intensify remote teaching to the spread of new forms of service provision.

Technology is one of the main forces of change. Although the pace and scale of technological change are not clearly defined, professionals are recommended to keep themselves informed and updated, to approach these changes in a critical but open manner and to contribute, where possible, to shaping the future of their profession. Academics are advised to intensify the study of the technology influence on the interpretation process and on technology-mediated multilingual communication. They are key to inform both practitioners as well users.

Notes

1 This kind of attitude, which is common to all professions, has been also defined as “technological myopia” (Susskind & Susskind 2017: 44), i.e. the “tendency to underestimate the potential of tomorrow’s applications by evaluating them in terms of today’s technologies”.
2 webgate.ec.europa.eu/sr/
3 www.speechpool.net
4 leedsenit.wordpress.com/
5 www.interpretimebank.net
6 www.audacityteam.org
7 www.televic-education.com/en/interpreterq-media-player
8 www.melissi.co.uk/virtual-interpreting-system
9 https://intrain.ditlab.it/
10 www.zoom.us
11 www.teams.microsoft.com
12 www.sanako.com
13 www.voiceboxer.com/
14 www.interprefy.com/
15 www.kudoway.com
16 www.moodle.org
17 www.interpretbank.com
18 www.interpretershelp.com
19 www.fourwillows.com/interplex.html
20 www.iate.eu

21 www.wikimedia.org
22 The first prototype of an ASR-enhanced CAI tool was developed in the context of the InterpretBank's project and is available at www.interpretbank.com/ASR.

References

AIIC 2018. AIIC Position on distance interpreting: Technical report. Available at: https://aiic.org/document/4837/AIIC_position_on_TFDI_05.03.18.pdf (accessed 16 October 2020).

Berard, Alexandre, Pietquin, Olivier, Servan, Christophe & Besacier, Laurent 2016. Listen and translate: A proof of concept for end-to-end speech-to-text translation. Available at: http://arxiv.org/abs/1612.01744 (accessed 30 April 2020).

Biagini, Giulio 2016. Printed glossary and electronic glossary in simultaneous interpretation: A comparative study. Unpublished Master's thesis, Università degli studi di Trieste.

Blasco Mayor, María J. & Jiménez Ivars, Amparo 2007. E-learning for interpreting. *Babel* 53 (4), 292–302.

Braun, Sabine & Slater, Catherine 2014. Populating a 3D virtual learning environment for interpreting students with bilingual dialogues to support situated learning in an institutional context. *The Interpreter and Translator Trainer* 8 (3), 469–485.

Brüsewitz, Nora 2019. Simultandolmetschen 4.0: Ist automatische Spracherkennung der nächste Schritt? In W. Baur & F. Mayer (eds) *Proceedings of the Conference Übersetzen und Dolmetschen 4.0.—Neue Wege im digitalen Zeitalter*. Berlin: BDÜ Fachverlag, 323–333.

Canali, Sara 2019. Technologie und Zahlen beim Simultandolmetschen: utilizzo del riconoscimento vocale come supporto durante l'interpretazione simultanea dei numeri. Unpublished Master's thesis, Università degli studi internazionali di Roma.

Carabelli, Angela 1997. IRIS Interpreters' Resource Information System. Una banca dati interattiva per la formazione di interpreti e traduttori. Unpublished PhD dissertation, Università degli Studi di Trieste.

Castillo, Pedro 2019. Using NICTs for media interpreting training: Bringing interpreter-mediated TV news and radio interviews to the classroom. *Revista Tradumàtica. Tecnologies de la Traducció* 17, 108–127.

Chen, Sijia 2017. Note-taking in consecutive interpreting: New data from pen recording. *The International Journal for Translation & Interpreting Research* 9 (1), 4–23.

Defrancq, Bart & Fantinuoli, Claudio 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.

de Manuel Jerez, Jesús 2006. La incorporación de la realidad profesional a la formación de intérpretes de conferencias mediante las nuevas tecnologías y la investigación-acción. Unpublished PhD thesis, Universidad de Granada.

Desmet, Bart, Vandierendonck, Mieke & Defrancq, Bart 2018. Simultaneous interpretation of numbers and the impact of technological support. In C. Fantinuoli (ed.) *Interpreting and Technology*. Berlin: Language Science Press: 13–27.

Deysel, Elizabeth & Lesch, Harold 2018. Experimenting with computer-assisted interpreter training tools for the development of self-assessment skills: National Parliament of RSA. In C. Fantinuoli (ed.) *Interpreting and Technology*. Berlin: Language Science Press: 61–90.

Díaz-Galaz, Stephanie 2015. La influencia del conocimiento previo en la interpretación simultánea de discursos especializados: Un estudio empírico. Unpublished PhD thesis, Universidad de Granada.

Di Gangi, Mattia Antonino, Dessì, Roberto, Cattoni, Roldano, Negri, Matteo & Turchi, Marco 2018. Fine-tuning on clean data for end-to-end speech translation. Available at: https://arxiv.org/abs/1810.07652 (accessed 30 April 2020).

European Commission 2019. Interpreting platforms. Consolidated test results and analysis. Technical report. Available at: https://ec.europa.eu/education/knowledge-centre-interpretation/sites/kci/files/interpreting_platforms_-_consolidated_test_results_and_analysis_-_def.pdf (accessed 30 April 2020).

Fantinuoli, Claudio 2006. Specialized corpora from the web for simultaneous interpreters. In M. Baroni & S. Bernardini (eds) *Wacky! Working Papers on the Web as Corpus*. Bologna: GEDIT: 173–190.

Fantinuoli, Claudio 2012. Interpretbank—design and implementation of a terminology and knowledge management software for conference interpreters. PhD dissertation, University of Mainz.

Fantinuoli, Claudio 2017. Computer-assisted preparation in conference interpreting. *Translation & Interpreting* 9 (2), 24–37.

Fantinuoli, Claudio 2018. Computer-assisted interpreting: Challenges and future perspectives. In I. Durán Muñoz & G. Corpas Pastor (eds) *Trends in E-tools and Resources for Translators and Interpreters*. Leiden: Brill: 153–174.

Fantinuoli, Claudio 2019. The technological turn in interpreting: The challenges that lie ahead. In W. Baur & F. Mayer (eds) *Proceedings of the Conference Übersetzen und Dolmetschen 4.0.—Neue Wege im digitalen Zeitalter*. Berlin: BDÜ Fachverlag: 334–354.

Fantinuoli, Claudio & Prandi, Bianca 2018. Teaching information and communication technologies: A proposal for the interpreting classroom. *Trans-Kom* 11 (2), 162–182.

Ferrari, Michele 2001. *SCIC News: Technical Report 26, European Commission—Directorate General for Interpretation*. Brussels: SCIC.

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

Gillies, Andrew 2019. *Consecutive Interpreting: A Short Course*. Abingdon: Routledge.

Goldsmith, Joshua 2017. A comparative user evaluation of tablets and tools for consecutive interpreters. In J. Esteves-Ferreira, J. Macan, R. Mitkov & O.-M. Stefanov (eds) *Proceedings of the 39th Conference Translating and the Computer*. Geneva: Tradulex: 40–50.

Goldsmith, Joshua 2018. Tablet interpreting: Consecutive interpreting 2.0. *Translation and Interpreting Studies* 13(3), 342–365.

Goldsmith, Joshua 2020. Terminology extraction tools for interpreters. In B. Ahrens, M. Beaton-Thome, M. Krein-Kühle, R. Krüger, L. Link & U. Wienen (eds) *Interdependenzen und Innovationen in Translation und Fachkommunikation / Interdependence and Innovation in Translation, Interpreting and Specialised Communication*. Berlin: Frank & Timme: 279–304.

Gorm Hansen, Inge & Shlesinger, Miriam M. 2007. The silver lining: Technology and self-study in the interpreting classroom. *Interpreting* 9 (1), 95–118.

Hamidi, Miriam & Pöchhacker, Franz 2007. Simultaneous consecutive interpreting: A new technique put to the test. *Meta: Journal des traducteurs* 52 (2), 276–289.

Jia, Ye, Weiss, Ron J., Biadsy, Fadi, Macherey, Wolfgang, Johnson, Melvin, Chen, Zhifeng, & Wu, Yonghui 2019. Direct speech-to-speech translation with a sequence-to-sequence model. Available at: https://arxiv.org/abs/1904.06037 (accessed 30 April 2020).

Kerremans, Koen & Stengers, Helene 2017. Using online and/or mobile virtual communication tools in interpreter and translator training: Pedagogical advantages and drawbacks. In J. Esteves-Ferreira, J. Macan, R. Mitkov & O.-M. Stefanov (eds) *Proceedings of the 39th Conference Translating and the Computer*. Geneva. Tradulex: 55–58.

Mellinger, Christopher 2019. Computer-assisted interpreting technologies and interpreter cognition: A product and process-oriented perspective. *Revista Tradumàtica. Tecnologies de la Traducció* 17, 33–44.

Merlini, Raffaella 1996. Interprit—Consecutive interpretation module. *The Interpreters' Newsletter* 7, 31–41.

Moser-Mercer, Barbara 1992. Terminology documentation in conference interpretation. *Terminologie et traduction* 2/3, 285–303.

Orlando, Marc 2010. Digital pen technology and consecutive interpreting: Another dimension in notetaking training and assessment. *The Interpreters' Newsletter* 15, 71–86.

Orlando, Marc 2014. A study on the amenability of digital pen technology in a hybrid mode of interpreting: Consec-simul with notes. *Translation & Interpreting* 6 (2), 9–54.

Orlando, Marc 2015. Implementing digital pen technology in the consecutive interpreting classroom. In D. Andres & M. Behr (eds) *To Know How to Suggest … Approaches to Teaching Conference Interpreting*. Berlin: Frank & Timme: 171–200.

Paulik, Matthias & Waibel, Alex 2009. Automatic translation from parallel speech: Simultaneous interpretation as MT training data. In *Proceedings of the IEEE Workshop on Automatic Speech Recognition & Understanding*, 496–501.

Plevoets, Koen & Defrancq, Bart 2018. The cognitive load of interpreters in the European Parliament: A corpus-based study of predictors for the disfluency uh(m). *Interpreting* 20 (1), 1–28.

Pöchhacker, Franz 2016. *Introducing Interpreting Studies*. 2nd edn. London: Routledge.

Prandi, Bianca 2015. The use of CAI tools in interpreters' training: A pilot study. In J. Esteves-Ferreira, J. Macan, R. Mitkov & O.-M. Stefanov (eds) *Translating and the Computer 37*. Geneva: Tradulex, 48–57.

Prandi, Bianca 2017. Designing a multimethod study on the use of CAI tools during simultaneous interpreting. In J. Esteves-Ferreira, J. Macan, R. Mitkov & O.-M. Stefanov (eds) *Proceedings of the 39th Conference Translating and the Computer*. Geneva: Tradulex: 76–88.

Pym, Anthony 2011. What technology does to translating. *Translation* & *Interpreting* 3 (1), 1–9.

Riccardi, Alessandra 2000. Die Rolle des Dolmetschens in der globalisierten Gesellschaft. In S. Kalina, S. Buhl & H. Gerzymisch-Arbogast (eds) *Dolmetschen: Theorie—Praxis—Didaktik—mit ausgewählten Beiträgen der Saarbrücker Symposien*. St Ingbert: Röhrig, 75–87.

Rosa, Helmut 2015. *Social Acceleration: A New Theory of Modernity*. New York: Columbia University Press.

Ruiz, Nicholas, Di Gangi, Mattia Antonino, Bertoldi, Nicola & Federico, Marcello 2017. Assessing the tolerance of neural machine translation systems against speech recognition errors. In *Proceedings of the Interspeech 2017*, 2635–2639.

Şahin, Mehmet 2013. Virtual worlds in interpreter training. *The Interpreter and Translator Trainer* 7 (1), 91–106.

Sandrelli, Annalisa 2002. Computers in the training of interpreters: Curriculum design issues. In G. Garzone, M. Viezzi & P. Mead (eds) *Perspectives on Interpreting*. Bologna: CLUEB: 189–204.

Sandrelli, Annalisa 2003. New technologies in interpreter training: CAIT. In H. Gerzymisch-Arbogast, E. Hajicová, P. Sgall, Z. Jettmarová, A. Rothkegel & D. Rothfuß-Bastian (eds) *Textologie und Translation*. Tübingen: Gunter Narr Verlag, 261–293.

Sandrelli, Annalisa & Hawkins, Jim 2006. Computer assisted interpreter training (CAIT): What is the way forward? In *Proceedings of Accessible Technologies in Translation and Interpreting Conference*.

Sandrelli, Annalisa & Jerez, Jesús de Manuel 2007. The impact of information and communication technology on interpreter training. *The Interpreter and Translator Trainer* 1 (2), 269–303.

Stengers, Helene, Kerremans, Koen & Segers, Jeroen 2018. Integrating virtual conference tools in interpreter and translator training. *Trans-Kom* 11 (2), 220–237.

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

Susskind, Richard & Susskind, Daniel 2017. *The Future of the Professions: How Technology Will Transform the Work of Human Experts*. Oxford: Oxford University Press.

Tiselius, Elisabet 2013. Expertise without deliberate practice? The case of simultaneous interpreters. *The Interpreters' Newsletter* 18, 1–15.

Tripepi Winteringham, Sarah 2010. The usefulness of ICTs in interpreting practice. *The Interpreters' Newsletter* 15, 87–99.

Vogler, Nikolai, Stewart, Craig & Neubig, Graham 2019. Lost in interpretation: Predicting untranslated terminology in simultaneous interpretation. Available at: www.aclweb.org/anthology/N19-1010.pdf (accessed 30 April 2020).

Vygotsky, Lev Semyonovich 1978. *Mind in Society: The Development of Higher Psychological Processes*. Cambridge, MA: Harvard University Press.

Will, Martin 2009. *Dolmetschorientierte Terminologiearbeit. Modell und Methode*. Tübingen: Gunter Narr Verlag.

Xu, Ran 2018. Corpus-based terminological preparation for simultaneous interpreting. *Interpreting* 20 (1), 29–58.

Xu, Ran & Sharoff, Serge 2014. Evaluating term extraction methods for interpreters. In *Proceedings of the 4th International Workshop on Computational Terminology (Computerm)*, 86–93.

Yu, Dong & Deng, Li 2015. *Automatic Speech Recognition: A Deep Learning Approach*. London: Springer.