Università degli Studi di Trieste
Dipartimento di Scienze Giuridiche, del Linguaggio,
dell'Interpretazione e della Traduzione

# The Interpreters' Newsletter

## *Changes in the Interpreting Landscape – New Developments and Old Modalities*

EUT

Special Issue n. 29 bis/2024

The *Interpreters' Newsletter* of the Dipartimento di Scienze Giuridiche, del Linguaggio, dell'Interpretazione e della Traduzione, Sezione di Studi in Lingue Moderne per Interpreti e Traduttori (SSLMIT) is an international journal promoting the dissemination and discussion of research in the field of Interpreting Studies.

ISSN 1591-4127 (print)
ISSN 2421-714X (online)

Euro 30,00

*The Interpreters' Newsletter*
Dipartimento di Scienze Giuridiche, del Linguaggio, dell'Interpretazione e della Traduzione
Sezione di Studi in Lingue Moderne per Interpreti e Traduttori (SSLMIT)
Università degli Studi di Trieste





ISSN: 1591-4127 (print)
ISSN: 2421-714X (online)

EUT Edizioni Università di Trieste
via Weiss, 21 – 34128 Trieste
eut@units.it
http://eut.units.it
https://www.facebook.com/EUTEdizioniUniversitaTrieste

# The Interpreters' Newsletter

## Special Issue 2024, 29 bis

### Changes in the Interpreting Landscape – New Developments and Old Modalities

**EUT** EDIZIONI UNIVERSITÀ DI TRIESTE

# Contents

# Editorial

Research on the impact of technology on conference interpreting has seen a growing interest at all scientific levels, from monographs (e.g., Frittella 2023; Prandi 2023) to edited volumes featuring contributions from multiple authors (e.g., Corpas Pastor/Defrancq 2023; Ji/Oakes, 2019; Fantinuoli 2018), as well as individual articles in collections focused on translation technologies, papers published in international journals and PHD and Master's theses. The latest issue of the journal *Interpreting*, published in October 2024, is a clear example of this trend, offering a selection of current research on interpreting and technology.

To keep up with the latest developments in interpreting studies, this special issue of *The Interpreters' Newsletter* includes four articles that explore how RSI platforms, hybrid forms of interpreting such as speech-to-text interpreting, and CAI tools equipped with automatic speech recognition and AI are shaping a new landscape for the profession. The impact of CAI tools and AI on speech processing and delivery in simultaneous and consecutive interpreting is still difficult to predict given the many variables that need to be taken into account in experimental studies, ranging from the subject matter to the type of meeting, the age and experience of the interpreters, the different types of speakers and speaking styles, among others. Much research remains to be done to get a picture of possible developments and outcomes.

While the fundamental elements of both conference interpreting modalities remain the same, the addition of new tools to the interpreting process paves the way for the augmented interpreter and augmented interpretation. Change always requires a degree of adaptation and resilience, and exploiting AI and CAI tools to the advantage of the interpreter can be challenging both for the interpreter, given the limits of their cognitive resources, and for the profession, in terms of working conditions.

The first two articles in this Special Issue are descriptive and discuss recent developments in remote simultaneous and speech-to-text interpreting.

Clare Donovan and Cecilia Manzoni, authors of *Remote interpreting: a game-changer for interpreters*?, provide a comprehensive review of the literature, as well as statements and position papers from professional associations and organisations on the subject. The positive and negative implications for conference interpreters and the profession of using remote simultaneous interpreting are discussed. At present, the interpreting market seems to have reached a new balance between remote and on-site interpreting. It is therefore now possible to take stock of the situation, consider possible future developments, and reflect on ways to help the profession develop resilience and adapt to ongoing technological advances. The final section of the paper is devoted to interpreter training and the ongoing debate about whether or not to include a seminar or course on RSI in interpreter training and when to introduce it.

Alessia Valente and Judith Platter describe how speech-to-text interpreting has developed and become established as a service for people with hearing impairments. The first part is a literature review of its main characteristics and how it works. Speech-to-Text interpreting involves the simultaneous transcription of a spoken text by an interpreter and is therefore also defined as written interpreting in Scandinavian, English and German-speaking countries. In Italy, it is mainly used for live subtitling during live broadcasts or film festivals and is less known for overcoming hearing barriers for students in schools and universities. The article reviews and compares the diffusion of speech-to-text interpreting and regulatory frameworks in Austria and Italy, looking at what has been achieved in one country and what still needs to be improved in the other. The more favourable situation in Austria could serve as a benchmark for possible improvements in Italy.

The next two articles are experimental studies, the first carried out in presence, the second online.

Automatic speech recognition (ASR) in consecutive interpreting (CI) is analysed by Michele Restuccia using Sight-Terp, a CAI tool developed for CI and equipped with both ASR and Machine Translation. The aim of using ASR tools to support CI is to provide an accurate transcription of numbers, names and figurative language, which can often be misinterpreted. The study compares the results of consecutive interpretations based on handwritten notes with those obtained using Sight-Terp. Data triangulation was used to observe the same phenomenon from different viewpoints. The results indicate that the ASR tool was used to transcribe the source language speech in the first stage of consecutive interpreting, while its output (the transcript) was used by the interpreters in the second stage of CI, either as a source text for a sight translation or as a backup text for the consecutive notes. The advantages and disadvantages of this approach are discussed, contributing to a better understanding of how effective the use of ASR can be in consecutive interpreting.

The online experimental study by Martina Behr, Markus Martini and Pierre Sachs investigated the relationship between working memory (WM) and simultaneous interpreting (SI) by varying the cognitive load in the first case and the source text difficulty in the SI task. The WM and SI tasks were administered online to 20 students on Master's degree programmes in conference interpreting from seven universities in Germany. The WM task had two cognitive load manipulation conditions, low and high. The interpreting task consisted of simultaneously interpreting a four-minute general

speech from English into German. The source text was manipulated by defining two levels of difficulty, alternating between easy and difficult paragraphs. The chosen procedure provided a more detailed picture of the specific relationships between WM and SI. Contrary to expectations, low and high WM load manipulations were similarly related to difficult ST passages, but not correlated with easy ST passages.

I hope that this selection of papers will provide an overview of some of the latest developments regarding the role of technology in different settings and help stimulate further interest in this area of research of Interpreting Studies.

Alessandra Riccardi

References

Corpas Pastor G. / Defrancq B. (eds) (2023) *Interpreting Technologies – Current and Future Trends*, John Benjamins, Amsterdam/Philadelphia.

Fantinuoli C. (ed.) *Interpreting and technology*, Berlin, Language Science Press.

Frittella F.M. (2023) *Usability research for interpreter-centred technology. The case study of SmarTerp*, Berlin, Language Science Press.

Ji M. / Oakes (eds) (2019) *Advances in Empirical Translation Studies: Developing Translation Resources and Technologies.* Cambridge University Press.

Prandi B. (2023) *Computer-assisted simultaneous interpreting. A cognitive-experimental study on terminology*, Berlin, Language Science Press.

**The Interpreters' Newsletter Special Issue 2024 (29 bis), 1-29**
DOI: 10.13137/2421-714X/36763 ISSN 1591-4127 (print) ISSN 2421-714X (online)
https://www.openstarts.units.it/dspace/handle/10077/2119

# Remote interpreting: a game-changer for interpreters?

Clare Donovan
ISIT and University of Lomé
Cecilia Manzoni
Conference interpreter

Abstract

*Remote interpreting has been studied and tested since the 1970s, but it has become more widespread since the Covid-19 pandemic. It is one of the key recent developments in the field of interpreting – along with CAI tools and artificial intelligence. It has the potential to bring about major changes in the way interpreters work, interact, and define their profession. While it is a cost-effective and flexible mode that offers opportunities to interpreters and clients, it is important to be mindful of the consequences for the profession. The aim of this study is descriptive, to review the literature on the subject to illustrate the impact that remote interpreting may have on various aspects of the profession. The focus is on the potential effects of remote interpreting on the interpreter as an individual, addressing auditory health and psychological health; and on the interpreter as a professional figure, analysing aspects such as identity and the process of professionalisation, teamwork between members of the profession, and training and research issues. The goal is to take stock of the situation currently, to consider possible future developments, and to reflect on ways to help the profession develop resilience and adapt to the technological advances underway.*

Keywords

Remote interpreting, RSI, SIDPs, hearing health, psychological health, professionalisation, teamwork, remote teaching.

Since its inception, and for most of its development, the practice of interpreting has presented the distinctive characteristic of the co-presence in the same physical location of at least three actors in communicative context; two interlocutors and an intermediary, the interpreter. Interpreting is intrinsically linked to communication developments by its very nature. This is how, starting from the late 20th century, with the breakthroughs in information and communication technologies, this conventional co-presence paradigm was irretrievably challenged by the advent of remote interpretation in its different forms.

Precise definitions of "remote interpretation" have varied. Sometimes remote has been used interchangeably with "distance interpretation", as in the following quote: "Remote interpreting (RI), which is also known as distance interpreting, means that the interpreter is not in the same room as the speaker and/or the audience" (FIT 2019)[1].

Most authors now differentiate between the broader concept of distance interpreting (whenever a speaker is not on the same site as the interpreter) and the more specific concept of remote interpreting: for instance, remote interpreting is described by Constable as "a situation in which interpreters are no longer present in the meeting room, but work from a screen using earphones, without direct view of the meeting room or the speaker" (European Parliament Interpretation Directorate, 2001: 12 in Constable 2015)[2]. Indeed, as early as 2003 Moser described remote interpreting as "any form of simultaneous interpreting where the interpreter works away from the meeting room either through a video-conferencing set-up or through a cabled arrangement" (Moser-Mercer 2003: 1). In 2015 Braun defined remote interpreting as having access "to an interpreter in another room, building, town, city or county" whilst the "primary participants" are "together at one site" (Braun 2015: 352). Of course, new developments especially since Covid mean that participants may also be off-site.

In the following, we will focus on remote interpreting, as distinct from the broader distance interpreting, as removal from the meeting room has more far-reaching effects for the interpreter than interpreting a few speakers who are on a video call.

## 1.1 Benefits and Disruptions of Remote Interpreting

Although various forms of remote interpreting have been tested and studied with interest over the years, this practice did not experience its heyday until the outbreak of the Covid-19 pandemic in 2020. During lockdown, remote interpreting proved to be a lifesaver for interpreting and for the many interpreters forced to work from home.

1 FIT (International Federation of Translators) (2019) "Discussion Paper on Remote Interpreting", <https://library.fit-ift.org/public/Publications/positionpapers/PDP_201906_Remote_Interpreting_ENFR.pdf>. Accessed October 3, 2024.

2 Constable A. (2015) "Distance Interpreting: A Nuremberg Moment for our Time. AIIC 2015 Assembly Day 3: Debate on Remote 01.18.2015", <https://aiic.ch/wp-content/uploads/2020/05/di-a-nuremberg-moment-for-our-time-andrew-constable-01182015.pdf>. Accessed October 3, 2024.

Without remote interpreting they would not have been able to pursue their profession. Even now, after the end of the restrictions, this very multifaceted mode offers many advantages for organisers. First and foremost, flexibility: it can be set up at short notice, and there is often no need to hire technical support staff, especially for meetings of a less formal nature. There are also considerable cost advantages for organisers: it is no longer necessary to rent booths and equipment, or to reimburse travelling expenses for interpreters.

Apart from these logistical aspects, it has been suggested that remote interpretation promotes the global expansion of the interpretation market, offering new opportunities to clients and professionals, especially for typically disadvantaged professional groups and sectors (Ünlü 2021: 332). The greater ease of organising meetings with interpretation could facilitate the matching of supply and demand. Many associations and organisations can thus expand their audience and capacity for action. Prominent guests who would otherwise be unable to attend can be invited to events. It could even be said that remote interpretation leads to a democratisation of interpretation services, which are now considered even in smaller meetings or where previously this would not have been possible for cost or logistical reasons[3]. It reduces the environmental impact due to less travel. And it has been argued that remote interpreting could enhance the quality of life for professionals by allowing them to maintain a better work-life balance (Tejada Delgado 2019: 91). Admittedly, many of these potential advantages are foregrounded by RSI platform providers with a vested interest in remote.[4] However, interpreters themselves have also identified some benefits. In a survey of some 900 interpreters, 26% stated a preference for some home working and 64% of respondents were happy to continue with RSI assignments (Buján/Collard 2023: 203). This was of course during the pandemic when other options were much reduced. But similar findings occur in a more recent study in which some interpreters interviewed identified advantages, such as less travel, more convenience and easier work-life balance (Salaets/Brône 2023: 198). This relates mainly to interpreting from home, although hubs could also reduce the need for travel to faraway conference locations.

Despite these potential advantages, remote interpreting has brought about major disruptions for interpreters and is continuing to do so, along with other recent developments that Fantinuoli (2018: 3) describes as the "technological turn" of interpreting: for instance, CAI tools, machine interpreting and recent innovations in artificial intelligence.

These are changes that have the potential to profoundly affect every dimension of the role and persona of the interpreter, thus impacting not only the interpreter as a professional figure, but also as an individual. Distance interpreting where some speakers are on another site does not profoundly change the way interpreters work, although sound quality for those speakers is often poorer. But remote interpreting substantially alters the way interpreters work: the place where they are located, the technical support they receive, the equipment and software employed, the way they interact with other stakeholders, the way their input and output are transmitted. All of this has a

3 The 2023 Nimdzi Interpreting Index (2023): <https://www.nimdzi.com/interpreting-index-top-interpreting-companies/#remote-vs-onsite-interpreting>. Accessed October 21, 2024.

4 See for instance Interprefy: <https://www.interprefy.com/resources/blog/benefits-of-remote-simultaneous-interpretation>. Accessed October 10, 2024.

multifaceted impact on the individual interpreter, and affects every aspect of the interpreting profession, shaping not only how the outside world perceives interpreters, but also how they perceive themselves.

### 1.2 Scope of the study

The aim of this study is to explore, through a review of the literature on the subject, the way various aspects of the interpreting profession have been affected, or may potentially be affected, by the use of the remote mode. The study will begin with the effects on the interpreter as an individual – the consequences on auditory and psychological health. Then we will consider the consequences on the interpreter as a professional figure, analysing the effects of remote interpreting on aspects related to professionalisation that, it could be argued, have made this profession what it is today. This analysis will include professional identity, cohesion, and teamwork. Finally, we will consider training issues in academic education, which is fundamental to the transmission of knowledge, skills and professional values. The main focus of this study is conference interpreting; however, some brief remarks will be made on interpreting for public services.

The two authors bring to the study different experiences of interpreting and training, one being a recent master's graduate, the other a trainer and interpreter of many years. This dual perspective should enrich the discussion. The study gives an overview of issues, but it is not intended to be prescriptive or all-encompassing. But it is hoped that it will contribute to creating greater awareness about the possible impact of the latest technological developments on interpreters and on interpreting. Greater knowledge and understanding will hopefully help negotiate change and direct the profession along the best possible pathway for the future.

## 2. The effects on the interpreter's health

As highlighted by the 2022 AIIC Staff Interpreters' Committee resolution, a safe and healthy working environment is one of the fundamental rights enshrined in the Declaration on Fundamental Principles and Rights at Work adopted by the International Labour Organisation in 1998. The objectives identified in it are constantly evolving, as expressed in the preface integrated in 2022: "Occupational safety and health is a moving target. While some improvements take place, new occupational risks emerge due to technological innovation or organizational change" (ILO 1998/2022: 1)[5]. It is therefore essential in every professional area to monitor the latest developments and ensure that they do not affect the well-being and health of professionals. The concept of health is here understood in the broadest sense, not only identifying health from a purely physiological point of view but also psychological.

5 ILO (International Labour Organisation) (1998/2022) "ILO Declaration on Fundamental Principles and Rights at Work and its Follow-up", <https://www.ilo.org/wcmsp5/groups/public/---ed_norm/---declaration/documents/normativeinstrument/wcms_716594.pdf>. Accessed October 26, 2024.

## 2.1 Hearing health

One of the focal points when it comes to safety in the field of interpreting is hearing health, as healthy hearing is essential for the successful pursuit of this profession. For an interpreter, monitoring one's hearing is as crucial as taking care of one's voice. Fears about the practice of simultaneous interpreting damaging hearing are not new. Some studies suggest that, generally speaking, there is a relationship between frequent use of headphones or headsets and a deterioration of hearing over time, possibly leading to so-called noise induced hearing loss and other conditions (Dehankar/Gaurkar 2022: 1; Gupta *et al*. 2022: 1-2). One of the many risks is injuries or acoustic shock:

> […] following an abrupt, intense and unanticipated acoustic stimulus, usually delivered by a telephone handset or headset, some individuals report a symptom cluster that includes otalgia, altered hearing, aural fullness, imbalance, tinnitus, dislike or even fear of loud noises, and anxiety and/or depression. Symptoms start shortly after the triggering acoustic incident and can be short-lived or can last for a considerable time (McFerran/Baguley 2007: 301).

Moreover, many interpreters listen with one ear and leave the other uncovered to better control their output. This, and the need to hear simultaneously their own voice, may prompt them to turn up the volume above recommended thresholds to hear more clearly (Hynes 2021: 170).

Interpretation entails therefore by its very nature risks to hearing. This situation seems likely to be exacerbated by the emergence of audio input, often encountered in remote interpreting, that is not qualitatively adequate because it is artificially altered, poor in terms of frequencies and transmitted via insufficiently wideband connections (Seeber 2022)[6]. Often, to remedy the poor quality of the signal and to be able to decipher what is being said – while at the same time producing an output with one's own voice and monitoring this output – interpreters are forced to increase the volume for extended periods.

## 2.2 Interpreters' response to RSI-related hearing risks

A lot of anxiety has been expressed amongst interpreters themselves about RSI-degraded sound as well as the risk of the aggravation of the consequences of dangerous phenomena such as acoustic shock, which is more likely to occur when the middle ear is already hypersensitive due to prolonged exposure to harmful sound (see statement at AIIC Assembly 2022)[7]. Interpreters as a professional body have been active in studying RSI-related hearing risks. AIIC in particular has issued numerous statements,

6 Seeber K. G. (2022) "When less is not more: Sound quality in remote interpreting", <https://untoday.org/when-less-is-not-more-sound-quality-in-remote-interpreting/>. Accessed October 24, 2024.

7 AIIC (Association Internationale des Interprètes de Conférence) Assembly (2022) "Declaration on Auditory Health", <https://aiic.org/document/10157/Declaration-auditory-health_janv22.pdf>. Accessed October 22, 2024.

and interpreters have taken an active part in the drafting of ISO standards on the topic.

Inadequate sound often occurred during the period of the pandemic crisis, with interpreters having to work with distorted sounds transmitted via the internet (Buján/Collard 2023: 145); and it continues to occur today in many meetings. As a result, interpreters have expressed fears of hearing disorders multiplying since 2020, a phenomenon borne out by various reports and investigations. For example, those carried out in Canada by AIIC and the Canadian Association of Professional Employees (CAPE) among interpreters in the Canadian Parliament. Symptoms reported by interpreters there range from hearing disorders such as tinnitus, hyperacusis and ear pain to more widespread discomfort such as insomnia, nausea, headache, and physical and mental fatigue (CAPE 2022: 5)[8]. Other possible consequences are listed in the Statement on Hearing Health issued by the AIIC Assembly in 2022: partial hearing loss, dizziness, acute migraines, eye discomfort, and chronic sleep disorders (AIIC Assembly 2022: 1)[9].

The CAPE report indicates several cases of symptoms so persistent and severe as to require much more sick leave than usual. There was also a significant increase in the number of reports on health and safety incidents submitted by interpreters. Several professionals have also initiated proceedings to claim compensation (CAPE 2022: 6). This led the AIIC Assembly to issue a strongly worded statement that "in many cases, this hearing damage is irreversible and, in some cases, has already led to permanent disability, prematurely ending the careers of the affected interpreters" (AIIC Assembly 2022: 1)[10]. This statement reflects the degree of concern within the profession.

A similar situation to the Canadian Parliament occurred in the European Parliament, where in June 2022 the interpreters launched a protest that lasted until mid-October due to what were felt to be deteriorating working conditions. More than one third of the professionals working in the Parliament claimed to suffer from symptoms similar to those described by CAPE, due to the strain of interpreting remotely-connected speakers from public environments with loud background noise, or with poor connections and inadequate microphones (De la Baume 2022)[11]. Fearing permanent hearing damage, the interpreters resorted to a boycott of remote interventions and online or hybrid events (VKD 2022)[12]. The boycott ended following an interim agree-

8 CAPE (Canadian Association of Professional Employees) (2022) "Recommended Actions to Solve Interpreters' Health and Safety Concerns", <https://www.acep-cape.ca/sites/default/files/2022-04/CAPE%27s%20submission%20to%20the%20BOIE.pdf>. Accessed October 24, 2024.

9 AIIC (Association Internationale des Interprètes de Conférence) Assembly (2022) "Declaration on Auditory Health", <https://aiic.org/document/10157/Declaration-auditory-health_janv22.pdf>. Accessed October 22, 2024.

10 AIIC (Association Internationale des Interprètes de Conférence) Assembly (2022) "Declaration on Auditory Health", <https://aiic.org/document/10157/Declaration-auditory-health_janv22.pdf>.

11 De la Baume M. (2022) "Streik, grève, strike: European Parliament interpreters (virtually) walk off the job", <https://www.politico.eu/article/strike-european-parliament-interpreters-walk-off-the-job/>. Accessed October 24, 2024.

12 VKD (Verband der Konferenzdolmetscher im BDÜ e.V.) (in English, Association of Conference Interpreters in the BDÜ e.V.) (2022) "Zur aktuellen Situation der angestellten

ment with the parliament. Reactions of concern have also spread among UN interpreters: in 2021, 60 per cent of interpreters in the Geneva office complained of some of the symptoms listed above after a remote interpreting session via a platform. Consequently, the professionals appealed to the institution for a total return to the face-to-face mode (Brady/Pickles 2022)[13].

In 2022 the International Organisation for Standardisation published the standard ISO 24019:2022 concerning SIDP platforms (ISO 2022), which clarifies the responsibilities of service providers and remotely connected speakers in achieving good sound quality. AIIC representatives took an active part in the ISO standard negotiations and pressed for audio criteria that would protect the interpreters' hearing. The AIIC Staff Interpreters' Committee also issued a resolution in September 2022, urging interpreters to report any symptoms and employers to take such reports seriously and discuss appropriate working conditions. The resolution stresses the importance of raising awareness about hearing risks involved in remote interpreting, and calls for a deeper scientific understanding of the effects of sound likely to damage hearing health. The EU institutions have proposed several initiatives as part of an awareness-raising campaign (News Hound 2022)[14]. They provide all participants of remote meetings with guidance on good practice and compliant equipment, informing them that the interpretation of remote speeches can be interrupted at any time if the sound quality is not adequate (DGI 2022)[15].

### 2.3 Awareness of auditory risks amongst other stakeholders

Awareness of auditory health problems related to remote interpretation has grown on the part of institutions and associations, but uptake of recommendations by platform operators and participants connecting remotely is patchy. Relevant ISO standards and best practice often run the risk of being disregarded. Many remote meetings, on the private market, but also on the institutional market, are held on generic platforms, such as Zoom or Webex, which are not covered by the ISO requirements. Zoom, in particular, is widely used. Chia-Ming Fan in a survey of remote interpreting in Taiwan notes it is the most widely used platform (2022: 174), as does Şengel (2022) in a survey of the market in Turkey, or again Donovan referring to experience of the Paris market (2023: 37). Indeed, the Nimdzi Atlas of translation-related technologies noted

und akkreditierten Dolmetscherinnen und Dolmetscher beim Europäischen Parlament" (in English, "On the current situation of employed and accredited interpreters at the European Parliament"), <https://bdue.de/fileadmin/files/PDF/Positionspapiere/BDUe_PP_EP_und_RSI.pdf>. Accessed October 24, 2024.

13 Brady A. / Pickles M. (2022) "Why remote interpretation doesn't work for interpreters", <https://untoday.org/why-remote-interpretation-doesnt-work/>. Accessed October 26, 2024.

14 See the BeHeard Campaign: News Hound (Internal newsletter of the European Parliament) (2022), "Remote speaking: get it right!", <http://www.linc.europarl.europa.eu/articles/remote-speaking-get-it-right.pdf>. Accessed October 26, 2024.

15 European Commission Directorate-General for Interpretation (DGI) (2023), "Code of conduct for remote speakers in multilingual meetings and conferences", <https://commission.europa.eu/system/files/2023-03/code-of-conduct-for-remote-speakers-dg-scic_en.pdf>. Accessed October 26, 2024.

that in 2021 Zoom was "de facto the largest RSI platform judged by the number of meetings" (Akhulova *et al*. 2021)[16]. On the other hand, the European Commission, the largest employer of interpreters, holds remote meetings with interpretation on Interactio which is subject to the ISO standard.

In sum, much remains to be done to ensure the hearing safety of interpreters working remotely, both in conference interpreting and, even more so, in less regulated areas such as community interpreting.

## 3. Psychological health: stress and fatigue

Interpreting is traditionally considered a high-stress activity, so much so that the ability to withstand tension over a relatively long period of time is considered one of the fundamental requirements for being able to undertake this activity at all, as well as an indicator of competence (Jiménez Ivars /Pinazo Calatayud 2001: 105). One factor of emotional agitation during interpreting is fear related to public speaking due to exposure to the judgement of others and ensuing feelings of insecurity and the undermining of self-esteem. As explained by Riccardi *et al*. (1998: 96): "The level of stress depends on the perception of the consequences of a failure to fulfil a request". Negative performance in interpreting is very visible and can have serious consequences professionally. Loss of face has often been cited in the research as a source of stress for interpreters (Monacelli 2009). Setton and Dawrant describe how stress is induced by "high stakes" and "public exposure" (Setton/Dawrant 2016: 321).

An interpreter, however, must not only speak in public, but also perform a series of complex cognitive processes, from listening to and analysing the source speech, memorising it, to producing and monitoring the output. This complex set of tasks has been described by Gile who concludes from his empirical studies that interpreters work close to processing capacity saturation (Gile 2017). All these components of the interpreting activity can be impacted by external factors, e.g. time constraints, pace and accent of the speaker, availability of the written text and access to preparation material, density of information, presence of new terminology and numerical data in the text, technical support, relationship with one's booth colleagues, the need to provide or use relays, and much more (Moser-Mercer 2005a: 75).

Many variables are outside the interpreters' control, thus generating emotional tension, which may impact their ability to effectively carry out the cognitive operations necessary for interpretation, and, consequently, impair the quality of performance (Riccardi *et al*. 1998: 96). At the best of times, the high cognitive load intrinsic to interpreting also causes progressive fatigue over time, which risks leading to a decrease in performance quality. So much so that Grosjean[17] (2011 in Seeber/Fox 2021: 503) states that interpreting is one of the most cognitively complex linguistic activities of which the human brain is capable.

16 Akhulkova Y. / Hickey S. / Agulló García B. (2021) "Nimdzi language technology atlas: The definitive guide to the language technology landscape. Nimdzi", <https://www.nimdzi.com/language-technology-atlasZoom>. Accessed October 25, 2024.

17 Grosjean F. (2011) "Those Incredible Interpreters", <https://www.psychologytoday.com/intl/blog/life-bilingual/201109/those-incredible-interpreters>. Accessed on October 25, 2024.

It can therefore be argued that interpreting is, by definition, a stressful, emotionally taxing, and cognitively demanding activity. And, as with auditory health risks, there are indications that these characteristics may be amplified by the remote mode. Perceptions of cognitive overload occurring more quickly, feelings of increased stress and tension, a sense of alienation and lack of control, a lower level of satisfaction with one's performance, difficulties in concentrating: these are all aspects that have been cited regularly by interpreters since the first experiments in remote interpretation carried out by the United Nations in the 1970s (Moser-Mercer 2005b: 77-80; Mouzourakis 2006: 52; Ziegler/Gigliobianco 2018: 123-125). They are the subject of research to this day. Alongside these perceptions, symptoms of a psychosomatic nature such as exhaustion, fatigue, nausea, eye strain and physical discomfort are also often reported during remote interpreting (Corpas Pastor/Gaber 2020: 67).

Among the first research studies to deal with these aspects in detail were the ETI-ITU study (reported by Moser-Mercer 2003) conducted in 1999, and the European Parliament's DG LINC study from 2004 (reported by Roziner/Shlesinger in 2010). Both studies aimed to compare the stress levels and performance quality of in-person and remote interpreting assignments. These aims were pursued by means of both objective measurements and investigations of subjective factors indicative of the emotional reaction of the interpreters. The results were convergent, confirming that the interpreters' perception was of higher stress levels, greater fatigue, and poorer performance quality. Nevertheless, the findings on objective parameters (e.g. cortisol levels, external evaluations of interpreting quality), while partly confirming the subjective considerations, did not show such a marked difference between in-presence and remote modes as to be considered statistically relevant (Braun/Taylor 2012: 43).

It is therefore still unclear to what extent (or if indeed at all) remote interpreting affects the interpreter's cognitive load and the quality of his or her performance. Moreover, these are aspects that are extremely difficult to assess with systematic and objective criteria. What is however regularly detected is the interpreters' subjective perception of discomfort and fatigue related to this modality.

Recently, various studies have been conducted to analyse the impact of specific aspects of remote interpreting, for instance the impact on stress and performance of lack of training and familiarity with the new SIDP platforms; or of the need to deal with technical and connection problems, or poor sound quality (Murgu 2021: 69). Interestingly, the same problems – such as early fatigue, feelings of stress and lower satisfaction with one's performance – seem to be relevant to all modes and domains in which interpreting services are employed remotely: not only simultaneous interpreting in conference settings, but also in consecutive and dialogic interpreting in public services. This observation is corroborated by several studies, such as Braun and Taylor's 2012 study in the legal sphere and Corpas Pastor and Gaber's 2020 survey covering the entire sphere of public services.

### 3.1 The sense of alienation

Various factors have been hypothesised to explain why interpreters report greater stress, fatigue and discomfort when interpreting remotely. One early hypothesis attributes these feelings to the sense of alienation due to the physical separation from other

participants in the communicative context (Moser-Mercer 2005a: 79). Researchers hypothesize that this sense of alienation is caused by a combination of individual perceptions and attitudes on one hand; and objective environmental factors related specifically to the remote mode on the other (Moser-Mercer 2005a: 79). For example, the lack of "immersiveness" of the remote interpretation work environment. During a face-to-face assignment, interpreters have access to a wide array of stimuli. This is not replicated with the same intensity during a video conference. Insufficiently rich visual information may, in particular, contribute to a sense of alienation, by depriving the interpreter of fundamental information about the situational context (Moser-Mercer 2005b: 730; Mouzourakis 2006: 54).

A perceived lack of control generally, related to the remote mode, may exacerbate the sense of alienation (Moser-Mercer 2005a: 94). Control can be defined as the ability to predict and react promptly to problematic situations; an ability that allows the individual to have better stress management and feel more positive, thus helping the interpreter to respond promptly to new stimuli in a virtuous circle (Murgu 2021: 70). In this hypothesis, unforeseen situations beyond the individual's control trigger a reaction of emotional tension. Certainly, many factors in a remote interpreting situation cannot be influenced by the interpreter. First and foremost, especially if not assisted by specialised technical staff, interpreters depend on the correct functioning of the technology, such as the Internet connection, interpreting software, or hardware such as headphones and microphones.

The degree of mastery of the SIDP platform, if one is used, also seems to be relevant to a greater or lesser sense of control. A study conducted in 2021 with UNINT students in Rome showed the importance for interpreters to have training with the specific interpretation platform before the assignment. The students found this essential for perceived good performance (Saina 2021: 10). Although, there is a move towards standardisation, each platform has its own specificities, and they change continually to meet different needs (Saina 2021: 10).

Another aspect likely to contribute to the feeling of alienation is the loss of immediacy in interaction with the other actors in the communicative situation. In an in-presence situation, it is relatively easy for an interpreter to point out any technical problems and to interact with event organisers. In a remote interpreting scenario, however, "the immediacy of control, the mode of control, and the modifiability of physical environments [...] are often compromised" (Moser-Mercer 2005a: 79), as it is more difficult to intervene in the actions of those who do not share the same physical space. These observations also apply to the interaction between booth colleagues, which is more cumbersome in the fully remote mode: this certainly makes the process of turn-taking and mutual assistance less smooth (Donovan 2023: 32). Chia-Ming Fan also records in interviews of interpreters working in home-based remote a lesser sense of "belonging" and worse communication with boothmates or colleagues (2022: 184).

### 3.2 Strain on cognitive load and interpreting performance

Interpreters also report that remote leads to an increased cognitive load. Just as with the sense of alienation, this is difficult to substantiate, being based for now mainly

on reports of subjective perceptions by interpreters (Seeber/Fox 2021: 503). Findings regarding performance quality are mixed. In her 2003 study, Moser-Mercer noted an early decline in the performance quality of interpreters in remote mode compared to in-presence mode, and formulated the need for shorter shifts than the conventional 30 minutes.

However, in their article summarising the findings of the 2004 European Parliament study, Roziner and Shlesinger (2010: 242) noted instead that the quality of performance measured with objective parameters was not significantly affected by the remote mode, contrary to the participants' individual perception. In a more recent survey, the results regarding perceived performance were not very clear-cut (Salaets/Brône 2023). The complexity of this issue is compounded because systematically evaluating the quality of an interpreting performance, whether by listeners or in self-evaluation, has proven difficult, owing to lack of consensus about the ranking of the numerous criteria involved. And working from home may produce different performance results than interpreting in a remote booth with colleagues and a technician (the latter being the conditions of the experiment described by Roziner and Shlesinger).

### 3.3 On-line Fatigue

One of the causes of the early fatigue could be the need to pursue multiple activities in parallel with interpreting (Mahyub Rayaa/Martin 2022: 24). Working from home is particularly problematic, as interpreters must troubleshoot technical and connection problems; manage a parallel channel to communicate with their colleagues; keep an eye on different screens in some configurations; keep on hand materials shared by the speaker, and consult their glossaries and preparation documents at the same time. This would explain rapid cognitive overload, and the reported difficulty in concentration for many interpreters. Moreover, the home environment is not as isolated and soundproofed as that of an interpreting booth, and any external sound can easily become a source of distraction (Mahyub Rayaa/Martin 2022: 25).

More intense cognitive load, however, is probably not the only reason why interpreters report tiring faster when interpreting remotely. Increased fatigue when online does not seem to be confined to interpreters. "Zoom fatigue" is a widely used phrase to describe feelings of exhaustion for all participants after long virtual meetings. Reports on the phenomenon date back to the surge in remote meetings during covid (Lee 2020; Ratan *et al*. 2022) and have been confirmed by objective physiological measurements since (Riedl *et al*. 2023). The fatigue is attributed to poor auditory input, stress due to intrusive camera display, temptation of multi-tasking, and difficulty to isolate from distractions, most of which also concern interpreters. As Mahyub Rayaa and Martin (2022: 25) state, "It is reasonable to assume that remote interpreting entails the same screen fatigue that many non-interpreters, suddenly forced to lead their lives online during the pandemic lockout period, talk about".

Nowadays, RSI in the fully remote mode, is not as common as during the pandemic; that is, where each interpreter works from their home office, separated from each other and the communication setting. However, remote from hubs or booths without a direct view of the meeting is still widespread. Professional associations, especially

AIIC, have issued recommendations for good practice to offset the additional fatigue and stress entailed, as seen above. These include shorter shifts and increased manning strength; preferably working from a distance interpreting hub[18] with colleagues and with the assistance of specialised technical staff; using platforms designed specifically for interpreting (AIIC TFDI 2020)[19].

## 4. The effects on professionalisation and status

Although interpreting is a longstanding activity, it has only recently become professionalised, with the emergence of dedicated academic training, standards, and professional associations. This process has not always been linear; in some ways it can be considered still ongoing today (Gentile 2015: 9). The degree of professionalisation still varies across the world, and even within the same country (Ünlü 2021: 318); it also varies according to the field – public service, court, sign language or conference interpreting.

It was the emergence of simultaneous interpreting particularly at the Nuremberg Trials and then in the United Nations, which led to the birth of conference interpreting, as recognised today. Simultaneous interpreting marks the inclusion of a first technological interface (Fantinuoli 2018: 2), and the emergence of a more precise definition of the tasks and role of an interpreter. Previously, interpreters were regarded as all-round officials who performed various functions in addition to their role as language mediators (Baigorri-Jalón 2014). During the post-war period the responsibilities of this professional figure gradually became better defined, and later laid down in codes of professional ethics (Horváth/Tryuk 2022: 291).

Thus, interpreting progressed slowly towards greater professionalisation and recognition. Fundamental elements in this process include the founding of the first interpreter training institutes in the 1940s and 1950s, and the creation of an international professional association, AIIC, in 1953, which marked the first step towards the creation of a structured professional group with clearly defined working conditions.

Ünlü (2021: 318) argues that «today, the profession of interpreting has become a full-fledged occupation, especially in terms of organization», as it presents all the essential elements of a profession with professional goals, as formulated by Furmanek (2012). One example is the notion of identity, a concept inextricably linked to status and prestige, and to professional cohesion. In other words, identity defines how the outside world perceives interpreters and how interpreters perceive themselves. Another objective is standardisation, achieved through cooperation involving professional associations and standardisation bodies, to create common standards and good practices. A third element cited notably by Boéri is the establishment of an academic

18 An interpretation hub is a studio equipped with the equipment (booths, consoles, etc.) and technical support staff necessary for simultaneous interpretation, from which interpreters can perform remote assignments together with their colleagues.

19 AIIC TFDI (AIIC Taskforce on Distance Interpreting) (2020), "AIIC Covid-19 Distance Interpreting Recommendations for Institutions and DI Hubs", <https://aiic.org/document/4839/AIIC%20Recommendations%20for%20Institutions_27.03.2020.pdf>. Accessed on October 24, 2024.

framework, with all that entails: "a body of theoretical and practical knowledge" and "formal training programmes" (Boéri 2015: 30). These characteristics are present in the interpreting profession today and make it what it is. In the following, an analysis will be conducted on how these aspects have been or may be affected by the emergence of remote interpreting, as further evidence of the potentially wide-ranging effects of this mode of interpreting.

### 4.1 Professional identity

In sociological studies on interpreting, it has often been pointed out that a continuous process of negotiating their own identity is intrinsic to the nature of the interpreter. The interpreter's role may be more problematic than norms would suggest. In a survey by Pöchhacker and Zwischenberger of 700 conference interpreters the most frequent definition of interpreters' role given is "facilitator/enabler of communication", followed by "mediator/intermediary" but even the most quoted definition only accounts for one fifth of replies[20]. Gentile notes also that interpreters are often perceived as intermediaries and bridges (Gentile 2015: 20). The ideal of interpreter invisibility – a neutral mediator or bridge – is queried in Diriker's case study in which she observes more active intervention than prescribed by socio-professional norms (Diriker, 2004).

The introduction of technology into the practice of interpreting has always brought major changes in professional identity. This is certainly true for the first revolution identified by Fantinuoli (2018: 2), i.e. the displacement of consecutive interpreting by simultaneous. In many ways this is not dissimilar to what he describes as the third, ongoing revolution: namely, the "technological turn" that includes the spread of distance interpreting.

The first "turn" – the switch from consecutive to simultaneous – entailed the removal of interpreters from the podiums and conference rooms where they were previously seated amongst the diplomats and other participants, and their relegation to the separate space of the booths, physically apart from the communicative situation, connected to speaker and listeners through technical apparatus. This change was initially met with scepticism and even outright hostility by interpreters themselves (Baigorri-Jalón 1999: 37; Fantinuoli 2018: 2). Elsewhere, Baigorri-Jalón describes the pejorative discourse used by consecutive interpreters to discredit simultaneous, describing the practitioners as "monkeys" or "parrots" (Baigorri-Jalón 2014: 254). Consecutive interpreters questioned the feasibility of simultaneous, but they were also anxious about the perceived status of their role; they feared that the transmission of their output via technological equipment and the lack of interpersonal interaction with the recipients would mean they were considered as nothing more than a machine (Jacob 1962; Gentile 2015: 14).

The feeling of being displaced from the heart of the meeting, the lack of appreciation or feedback about the importance of their work, and the consequent sense of

20 Pöchhacker F. / Zwischenberger C (2010) «Survey on quality and role: conference interpreters' expectations and self-perceptions", <https://aiic.org/document/9646/Survey%20on%20quality%20and%20role%20conference%20interpreters%E2%80%99%20expectations%20and%20self-perceptions.pdf>. Accessed on October 25, 2024.

a devalued status was a source of anxiety for interpreters during this period of early simultaneous. Moreover, objectively their working conditions had deteriorated or at least changed radically. Baigorri Jalon describes how in the early days UN interpreters were expected to work long shifts in overcrowded spaces and unsuitable working environments (Baigorri-Jalón 2004: 114). Eventually, the UN interpreters went on strike in 1974. They subsequently obtained improved conditions that are still in place today (Baigorri-Jalón 2004: 115). This exemplifies how conference interpreters have responded to restore conditions conducive to quality and health as their work environment changes.

It is easy to draw an analogy with the recent developments in remote interpreting. In a way, remote interpreting can be seen as a further step in the process of technologisation of the profession, further widening the gap between interpreters and conference participants. The distance is no longer that separating the booth from the meeting room, but that between places that may be in two different cities, countries or even continents. It has always been difficult to associate interpreters with a specific workplace, but with remote interpreting workplace location can potentially be anywhere. The first UN simultaneous interpreters expressed concerns that are echoed today by interpreters worried about the introduction of remote interpretation, notably through AIIC resolutions as seen above, as well as in research-driven surveys. The fears relate to deteriorating working conditions, a sense of alienation, lack of control, and absence of feedback (Salaets/Brône 2023: 198-204, Buján/Collard 2023: 147-148). In 2018 Fantinuoli (2018: 3) spoke of remote interpreting, along with the other recent developments, as a breakthrough that:

> […] not only could […] lead to a transformation of the interpreting ecosystem in all its complexity, but it is reasonable to assume that it may have a significant impact on many socio-economic aspects related to the profession, from the way it is perceived by the general public to the status and working conditions of interpreters.

Since then, the pace of technological development has accelerated, in part confirming these observations. The longer-term effects are not yet clear, as so much is in a state of flux. However, the risk of remote interpretation leading to depersonalisation is part of a broader trend related to technological developments more generally (Gentile 2015). Fantinuoli (2018: 7) points out how the apparent homogenisation of a product, caused by the perceived depersonalisation of its providers, leads buyers to choose the cheapest option, potentially triggering a mechanism of deterioration of the service's quality. An analysis of the translation market by Dwyer (2010) shows how such a downward spiral can come about in the absence of protective mechanisms, such as accreditation. Ultimately, this pattern could reverse the process of professionalisation of interpreting.

In the case of remote interpreting, professionals can find themselves isolated at the end of the service provision chain, which amplifies the distance between professional and client by putting much of the decision-making power into the hands of the platform operator. This is a trend described as worrisome by many participants in the study by Buján and Collard (2023: 146). Many interpreting platforms provide interpretation as part of a video conference package, and there is no communication between the actual organisers and the interpreters. This is the case with regard to

sound quality: as pointed out by Caniato (2021)[21], SIDP platform operators decide how much bandwidth to dedicate to audio transmission, so if they opt to reduce costs this can lead to a deterioration in working conditions and hence service quality. In this scenario, the interpreter's ability to intervene is very limited.

Moreover, when working from home in particular, interpreters are very isolated. This exacerbates their invisibility both metaphorically and physically. Usually, interpreters are not visible at all to participants during video conferences, as their cameras are off and their names do not appear in the list of participants. In this sense as well, we are witnessing an accentuation of the tendencies that emerged during the shift from consecutive to simultaneous interpreting. The reduced visibility deplored by the first simultaneous interpreters behind the glass windows of their booths has turned into the absolute invisibility of interpreters behind their computer screens. The risk of a weakening of the interpersonal relationship with listeners is the reinforcement of or the return to the idea of the interpreter as a mere tool that automatically transfers a message from one language to another. This is what Giustini (2022: 206) refers to as a 'disembodiment' of interpreters. The risk is that their physical invisibility could lead to an invisibility of their role as workers and as individuals.

Another possible impact of remote interpretation might seem in contradiction with the phenomenon just highlighted, but is in fact just another consequence of a more blurred definition of their professional role. This is the potential broadening of their responsibilities. With the shift of their activity online, interpreters often find themselves taking on new tasks, in particular dealing with technical issues such as equipment, and connection problems.

On some markets interpreters are not only considered experts in interlingual communication, but also in organising online events. According to a survey administered in 2020 by the Turkish association of conference interpreters (TKTD), about 42% of professionals guide clients in choosing the most suitable platform and technical solution when preparing a remote assignment (Ünlü 2021: 325). Especially when clients have little experience in the logistics of remote events, interpreters may have to act as consultants in this area. Contrary to fears about interpreter invisibility this development, currently restricted to some markets, might, if it were to become more prevalent, give interpreters more influence in decision-making. It could also provide an opportunity to raise the client's awareness about the specificities of remote interpretation and point to the best solutions.

All this means that interpreters will need resilience, adaptability, and openness to lifelong learning. An expansion or redefinition of one's skill set is required to survive a career shock, such as that triggered by the pandemic (see Ünlü 2021: 321). Interpreters have had to acquire a new set of digital skills that now may prove to give added value to their careers. It remains to be seen what the long-term consequences of these developments will be for the status and identity of the interpreter; what is certain is that professionals and professional associations will need determination, commitment, and unity to navigate the inevitable changes ahead.

21 Caniato A. (2021) "RSI Sound Myth Buster: Ten Misconceptions that Result in RSI Sounding Terrible", <https://www.lourdesderioja.com/2022/10/12/good-sound-please/>. Accessed on October 24, 2024.

## 4.2. Professional cohesion: teamwork

As Ünlü (2021: 326) puts it, "whether conducted on distance interpreting hubs or in the home, it is important to bear in mind that simultaneous interpretation is task-shared, collaborative work". So much so that it is perceived as a unified service by the recipients and the clients, who rarely differentiate between the different team members (Donovan 2023: 33). The entire team of interpreters works together towards a common goal: to produce good quality interpretation.

The importance of optimal cooperation between colleagues is clear. A 2016 study (Rangponsumrit/Rattanabutr: 68) showed that most interpreters expect colleagues who are not actively engaged at the microphone to remain in the booth and be ready to offer their help if needed. The "AIIC Covid-19 Distance Interpreting Recommendations for Institutions and DI Hubs" published in 2020[22] also lists some of the tasks with which interpreters routinely help each other, e.g.: jotting down numbers, acronyms and proper names; writing down and researching technical or unusual terms; searching for documents and keeping track when a speaker reads a written text previously provided to the interpreters; monitoring colleagues' output and input to ensure that the relay system is working properly; requesting assistance in case of technical problems; replacing colleagues in case of unexpected events. Precisely to foster all these dynamics and enable effective cooperation within the booths and the équipe in general, AIIC strongly recommends that even during a remote assignment "all interpreters be in the same room or space. Individual language teams/booths are required to be co-located in the same place in order to avoid added cognitive load […]" (AIIC TFDI 2020: 8)[23]. All interpreters interviewed in a recent survey stated that teamwork and cooperation were more difficult in a fully remote setting (Salaets/Brône 2023: 201).

In addition to these practical aspects, however, there are many other reasons why time spent in the booth together with one's colleagues is essential for interpretation. Interaction between members of the same community is one of the founding aspects – along with professional associations and training – of the process of professionalisation of every work activity. It allows common ideas, values, practices, and norms to be developed and passed on internally (Donovan 2023: 26). Thanks to the conversations that take place within the booths, therefore, what is considered acceptable (or not acceptable) in the practice of interpreting is reaffirmed, thus also contributing to the consolidation of a full-fledged professional identity. Not only that, but it is also within the booths, and not only in the universities, that knowledge underlying professional cohesion – e.g. strategies, solutions, and norms of behaviour – is transmitted to new generations of interpreters.

One of the reasons the birth of simultaneous interpreting provided such an impetus to the professionalisation of interpreting is precisely the creation of a common space – which was previously lacking – exclusively dedicated to the professionals in this field, a space where advice and good practices are now exchanged on a regular basis (Bai-

22 AIIC TFDI (AIIC Taskforce on Distance Interpreting) (2020), "AIIC Covid-19 Distance Interpreting Recommendations for Institutions and DI Hubs", <https://aiic.org/document/4839/AIIC%20Recommendations%20for%20Institutions_27.03.2020.pdf>. Accessed October 24, 2024.

23 Ibidem.

gorri-Jalón *et al*. 2022). Moreover, this dynamic is established not only in the booths but also in the adjacent spaces shared with the rest of the team (for instance, corridors, rest areas). In addition, other important opportunities for socialisation are represented by trips to other cities for conferences or multi-day events, during which interaction spreads throughout all shared moments and generates a sense of belonging within the team (Taylor Bouladon 2007). Similarly, when interpreters regularly work together, for example in international organisations, it gives rise to cohesive professional communities with shared narratives and practices (Duflou 2016: 16).

A fully remote mode inevitably makes interaction between colleagues more complex and less spontaneous. Donovan's 2023 study compares interaction and cooperation between interpreters of the same booth at face-to-face and at remote events. It concludes that in on-site meetings "The booth can also be seen as a space of exclusive socialization in which interpreters cultivate their own analysis of the event in a parallel meta-discourse to the actual meeting" (Donovan 2023: 45).

However, the study by Donovan also indicates that interpreter exchanges at remote events are shorter and drier. They tend to focus on practical matters such as turn-taking or schedule changes. In general, they leave less room for (or in some cases are completely lacking in) offers of mutual help, interactions of a social and personal nature, face-saving techniques to justify sub-standard performance and exchanges of solutions and knowledge regarding meeting content – all dynamics typical of face-to-face meetings. It also seems that interpreters more rarely listen to their colleagues' deliveries, probably because it is more difficult to get their attention and to provide help in case of difficulty. Almost entirely absent, moreover, are those conversations that are vital for the reinforcement of common knowledge and thus for professional cohesion, in which conventions, norms and current trends in the field are discussed. Thus, in fully remote mode the booth – reduced to a shadowy virtual space - is no longer a privileged place for the creation of common narratives and identities, and for fostering a spirit of collegiality. However, this study only focuses on interpreter interaction for isolated home working. The same is unlikely to hold true for interpreting from hubs or from adjacent areas to the meeting.

Nor can it be said that the remote mode totally prevents cooperation between interpreters. In some ways the development of communication technologies has made the exchange of information much easier and more immediate than in the past. For instance, it is very easy to send files and presentations, or to create shared folders and glossaries to which everyone can contribute. Moreover, since the start of the pandemic, many remote interpretation platforms have also improved their offerings with better tools to facilitate interaction between colleagues. These include features to facilitate turn-taking, such as a timer indicating how much time has passed since the start of each shift. To some extent, this simplifies work management as it automatically signals when it is time to hand over the microphone. Other functionalities allow the interpreter on break to signal to their colleague they are ready to take over. It is often possible to monitor the colleague's output by varying the volume levels of both channels – the original and the interpretation (Giustini 2022: 208). These features were absent from most platforms even a few years ago. Another function now commonly found in remote interpreting platforms is a group chat for interpreters to communicate with each other.

Nevertheless, as seen above, most organisers – including some institutions – still opt for generic video conferencing platforms which do not have these features, so

interpreters need to establish a parallel communication channel, often via WhatsApp messages. As already noted, introducing an additional tool to manage alongside the platform interface entails greater cognitive load and stress. It should also be noted that even when a group chat is available, interpreters often still prefer a private channel, probably due to reluctance about their conversations being monitored by technicians or event organisers. The feeling of being observed by others would certainly risk making interactions less spontaneous (Giustini 2022: 209). Another option is to establish communication between boothmates via a private video call on another platform or device. This option would make the interaction richer and provide visual input as well, so interpreters could read the colleague's non-verbal language and pick up on signs of distress and offer help. Nevertheless, this option is in practice rarely used, perhaps to avoid a further element of complexity.

The extreme consequence of these difficulties in interacting with colleagues is, once again, a sense of alienation due to a perceived estrangement from the rest of the team. In fact, in the 2020 TKTD survey, when asked to name the five most negative characteristics of remote interpreting, the most common response was the inability to cooperate in the usual way with other team members (Ünlü 2021: 326).

Currently there is not enough data available to establish with certainty which trends and developments of remote interpreting uptake will be consolidated in the long term. The observations made so far, however, seem to indicate that many interpreters think that the consequences of a massive spread of remote interpretation could be deleterious on several fronts. Firstly, with regard to the quality of the interpretation service, which is highly dependent on cooperation and mutual help within booths and teams. As Donovan (2023: 45) states:

> […] if these findings are confirmed on the basis of further research, the conclusion may well be that growing recourse to interpreting in a fully remote mode could […] lead to a more limited understanding by interpreters of meeting content and pragmatics, and therefore to a more superficial interpretation, as full remote provides little opportunity to share background knowledge.

Secondly, professional cohesion could also be weakened by a lack of interaction and exchange amongst interpreters, as well as by the interruption of the chain of direct knowledge transmission between old and new generations of interpreters.

In the current post-pandemic market situation, a complete substitution of the in-presence mode in favour of the fully remote mode is, of course, unlikely. As is, therefore, the threat of a complete disappearance of physical interaction spaces for interpreting professionals. Nevertheless, the considerations made so far underline the importance of preserving interpreter collaboration. Therefore, whenever opportunities for professional cohesion and cooperation are at risk of being curtailed, it is crucial that interpreters and their associations be alive to the need to create new spaces to replace them, for instance, congresses and conferences on profession-related topics, blogs and platforms on which to share experiences and knowledge. And this does seem to be happening, with a marked increase in webinars, online discussions, and conferences by and for interpreters, often using the same online tools as for remote interpreting.

## 5. Training issues

The academic sphere, both research and teaching, has not remained unscathed by the passage of the Covid-19 pandemic and the rise of remote interpretation. The sudden changes in the interpretation market forced researchers to focus more on remote interpreting; firstly, to gain a deeper understanding of these developments, and secondly to provide proven data to define and justify new working conditions and training methods adapted to remote (Mahyub Rayaa/Martin 2022: 23). Although researchers' interest in remote interpretation had originated well before the pandemic crisis, from 2020 onwards the number of studies and theses on the subject grew exponentially.

### 5.1 Training at a distance

As for training, remote teaching was not new. Glendon college in Canada has run blended CI training for some time, with a first year entirely online, followed by a second year on site. Various online further training modules are provided by experienced interpreters in various settings. Geneva University has successfully offered an online training for trainers' course for many years. Another more recent example is the one-year CI Masters course set up in Lomé University in Togo in 2023. To benefit from trainers based in Europe and North America, without entailing huge expense, about two thirds of the classes are given in remote mode, supplemented by onsite classes with younger, local colleagues. However, these examples are based on deliberate decisions made in specific circumstances, often involving further training, unlike the abrupt and constrained switch to remote training brought about by the pandemic restrictions.

Academic training in interpretation had to be held completely online for the duration of the acute phase of the pandemic, due to the various lockdowns that forced numerous areas of the world to suspend all social contacts for several months. Within a few weeks, interpretation teachers, many of whom had previously been sceptical or downright hostile towards distance training, had to adapt their teaching methods and materials to the new remote mode. Moreover, they often had to use generic videoconferencing platforms that lacked crucial features such as simultaneous interpreting options. Students had to adjust to a new learning environment without the participative and communication dimensions so typical of conventional interpreting classes. In fact, academic curricula in interpretation were among those most impacted by distance learning, precisely because of their intensely interactive nature and the importance of non-verbal communication cues (Han *et al*. 2022: 2).

Distance teaching of interpreting had to deal with the same issues that characterise remote interpreting itself: connection problems for both lecturers and students, poor audio quality, difficulties in picking up non-verbal signals, impaired concentration, etc. Not only that, the lack of platforms specifically designed for teaching interpretation remotely also proved problematic, particularly in the case of simultaneous interpreting. Traditional videoconferences were adequate for teaching consecutive and dialogic interpreting, at least for smaller groups, but simultaneous was trickier. In some cases, lecturers opted to combine lectures with exercises in asynchronous mode

by having students record their delivery, send it to them and then provide feedback later (Ünlü 2021: 331).

The spread of distance learning led to the emergence of many studies on the effectiveness and downsides of this training modality in general, and more specifically in the field of interpreting (e.g. Ünlü 2021; Kyong-Jo/Hyang-Ok 2022; Han *et al*. 2022; Ho/Zou 2022). These studies chose to explore various angles of training: for example, experiments with specific platforms, comparing synchronous and asynchronous modes; the level of satisfaction among students and teachers; student progress; the feeling of involvement and the "immersiveness" in the remote learning environment; the impact on the ability to concentrate. The general trends observed seem to indicate that distance learning in interpreting does not lead to any specific learning deficits. Students' skills have progressed consistently. Students taught in remote mode during lockdowns at training institutions such as ESIT and ISIT in Paris have similar pass rates at final exams, and similar early career pathways, as other cohorts (observations by Donovan)[24].

Researchers have assessed remotely-trained student skills in terms of, for example, completeness, accuracy, speed of delivery, use of terminology, coherence, and cohesion, and not observed major discrepancies (Ünlü 2021: 330). It thus appears that the quality of teaching and learning for both consecutive and simultaneous mode is not necessarily impacted negatively by remote training in a significant way. Indeed, some lecturers claim distance learning can lead to a more relaxed environment, helping students to manage the stress and anxiety related to classroom performance (Hodáková 2021: 69).

In some situations, remote, or at least blended, training may even be the best option. The Lomé Master's programme mentioned above recorded a two-thirds pass rate at the end of the first year. It would have been impossible to teach the Masters in the conventional way, owing to scarcity of trainers locally. As it is, the students have benefitted from wide training expertise from a range of trainers based in different markets and working for various institutions.

Nevertheless, these apparently positive outcomes mostly relate to exceptional circumstances during the pandemic with both trainers and students able to invest considerable time and energy in classes due to the specific and unusual context of lockdown restrictions. Remote training takes its toll in terms of fatigue and stress. The studies on the distance learning of interpreting have found repercussions in terms of motivation and psychological engagement on the part of both students and lecturers (Kyong-Jo/Hyang-Ok 2022: 107). Participants in some cases expressed feelings of isolation and alienation and complained about the lack of face-to-face contact with colleagues and teachers. Furthermore, the frequent occurrence of technical problems and, in some cases, the lack of sufficient skills to solve them, led to episodes of stress and fatigue. Similarly, students reported difficulties in concentrating, which often led them to become distracted and focus on other activities during online lessons. The lower level of participation and interactivity exacerbate this tendency.

These findings are confirmed by a survey conducted among the 15-student cohort enrolled in the one-year Master's at Lomé University. Respondents were glad to have

24 Unpublished statistical review of student and graduate outcomes 2018 to 2022.

had online classes with experienced trainers from different backgrounds. However, they identified several issues of concern – class size, student concentration, the risk of multi-tasking, poor connectivity, and fatigue[25]. Their feedback confirms the other reports of online interaction (see above). Firstly, remote classes were felt to be more tiring than onsite and therefore should not be too long. Secondly, it was harder to concentrate on the class (less sense of involvement), and harder to resist the temptation to multitask. Thirdly, class size was felt to be a significant factor, with larger groups exacerbating distraction and feelings of non-involvement. Solutions would logically be to have shorter sessions, to ask students to keep their cameras on, connectivity permitting (this is a requirement for many French online courses), and to have smaller student groups, or organize classes with two trainers and divide students up into breakout sessions.

The use of distance learning in interpreting, therefore, although not necessarily detrimental to the learning curve, may not be ideal in the long term, especially not as an exclusive teaching method. There are also some specificities in the practice of interpreting that might be easier to learn and exercise in a face-to-face situation; such as the use of booth equipment (consoles, headphones, microphones, amplification system); "etiquette" in the booth, i.e. the rules of good conduct to be observed when interacting and cooperating with colleagues; the correct presentation of a consecutive performance in terms of posture, gestures, maintaining eye contact, etc. It seems therefore inadvisable to consider distance learning as a complete substitute for in-person teaching. Moser-Mercer points out the importance of psychological factors when interpreting remotely: "Human factors then emerge as one of the most important issues to be explored in remote interpreting" (2003: 1). These factors relate to motivation, processing capacity and social isolation (Moser-Mercer 2003: 1). The same factors are likely to intervene in remote training. Still, remote teaching can be an inexpensive and quickly-organised addition to face-to-face lessons or be used in specific cases.

### 5.2 Training for remote interpreting

Another issue of dispute within the academic world is the possibility of including remote interpreting as a formal teaching subject within university curricula. Already six years ago, Ziegler and Gigliobianco (2018: 137) wrote:

> It goes without saying that fellow (conference) interpreters need to be prepared for remote simultaneous interpreting during their training, as this modality is experiencing a growing demand in different interpreting specializations, including tele- and video-conference interpreting.

This need was recognised even before the pandemic crisis, so now that demand has grown, the topic is certainly even more relevant. There are basically two schools of thought in this regard: those who believe that specific training modules should be devoted to the subject given the technical, cognitive, and communicative specificities;

25 Survey Donovan July 2023, article to be published.

and those who believe that the conventional approach to the teaching of interpretation already provides sufficient exposure to and familiarity with the remote mode, and that more time should be devoted to core and transversal skills that can help students in any circumstance. This second line is the one that was advocated by Seeber and Fox (2021: 504), who argue that most university curricula already use pre-recorded audio and video material for their exercises – whether these are their own materials or databases available on the web, such as the European Commission's Speech Repository. Such source texts thus create a working environment comparable to distance interpreting (remote speakers) or even interpreting from a hub, the remote option recommended by professional associations.

Seeber and Fox (2021: 505) also observe that, if the main difference between video material used in class and an actual remote interpreting situation is a better quality audio-visual signal in the former case, then rather than accepting this state of affairs and training students to work in situations that may be detrimental to their well-being, it would be preferable to invest more resources in ensuring compliance with the various recommendations and standards for remote interpreting in the work environment. This is a valid, but perhaps somewhat idealistic stance, especially for interpreters on the private market.

Similar opinions were also expressed by some participants in a study carried out in 2022 by Mahyub Rayaa and Martin. Other comments from those who believed that specific training for remote interpreting is not necessary – or not essential – are based on the view that working from an interpreting hub is comparable to a face-to-face situation (this is not a view shared by many interpreters who continue to express a preference for onsite interpreting, as evidenced regularly in surveys, such as that of Buján and Collard quoted above.) Some participants emphasised that despite the change in format, the basic techniques remain the same. Even the equipment does not necessarily have to be different: for instance, when working from a hub and, also when working via SIDP platforms, there is always the possibility of connecting a hard console[26] to the device (Seeber and Fox 2021: 504).

There are other more pragmatic reasons that may cause reluctance on the part of universities to integrate specific remote interpreting modules into their curricula – for instance, the difficulty in building a training programme that can keep up with fast-changing technical set-ups. Also, the sheer range of technical features that vary greatly depending on many factors (e.g. platform, hub, or home office interpreting, consecutive or simultaneous interpreting, etc.). Moreover, the theoretical foundations of this subject rest on a research base that is still being consolidated and to which new contributions are continuously being added. It could be argued that it is not necessarily the responsibility of universities to provide specific knowledge in this field, but rather that of bodies dedicated to professional development; for example, professional associations and agencies. Indeed, the companies providing SIDP platforms themselves already offer onboarding workshops. Finally, it may be difficult to find space to integrate modules on the latest developments and technologies in interpreting

26 Hard consoles refer to the classic interpretation consoles that are installed in physical booths and used for simultaneous interpretation; as opposed to soft consoles, i.e. digital consoles integrated into remote interpretation platforms that replicate the features of physical consoles.

into curricula that already suffer from the pressure of cutbacks in teaching hours and financial support.

On the other hand, there does seem to be a consensus amongst most trainers that specific training in remote interpreting is important. For instance, Ünlü (2021: 329):

> Providing awareness of the implications, fostering knowledge of technical equipment and international standards, disseminating information about codes of conduct, and client communication are, among many others, core dimensions that need to be conveyed to future practitioners of the profession.

The consensus is reflected in the EMCI (European Masters in Conference Interpreting) consortium's, Quality Assurance Standards[27] which state students must regularly be exposed to the new technologies used in the working environment and that lectures must include technology-mediated work scenarios, such as remote and video conference interpreting. According to this view, it is important for universities to take on the task of preparing aspiring professionals for what some consider to be the 'future of the profession' (Mahyub Rayaa/Martin 2022: 31). This not only prepares them to deal competitively with actual market trends and demands, but also gives them the tools to make informed choices. Thus, their future decisions are less likely to undermine professional standards and they can advocate for appropriate working conditions. University courses constitute an important basis for training and for the transmission of identity and values vital to the process of professionalisation. Moreover, the advent of new technologies in the interpreting market (in addition to distance interpreting, developments related to artificial intelligence, CAI tools, etc.) increasingly raises new ethical questions that young professionals must be prepared and able to address (Horváth/Tryuk 2022: 298).

A grounding in remote interpreting can therefore help train professionals who are aware, responsible, and attentive to developments in the practice of interpreting. But not only that; it could be argued that such training can also be crucial in providing practical skills. 76% of the participants in Mahyub Rayaa and Martin's study (2022: 31) stated that they considered training in distance interpreting necessary before undertaking a remote assignment. Remote interpreting, they claim, presents specificities that novice interpreters need to learn. Examples cited were coordination with colleagues, recognising and managing auditory and cognitive fatigue, and troubleshooting technical problems. In some responses it was also argued that training can help reduce the stress related to lack of control in remote and the feeling of being reliant on technology. These findings are in line with those of Saina (2021). In addition, Mahyub Rayaa and Martin's study also revealed that some respondents find training helps them to take a more pro-active role in dealings with clients. They are better able to recommend optimal solutions, thus strengthening the position of the interpreter within the decision-making and organisational processes. This brings us back to the possibility of an expansion of the interpreter's role with new technology, as discussed above (section 4.1).

27 EMCI (European Masters in Conference Interpreting) (2012) "EMCI Quality Assurance Standards", <https://www.emcinterpreting.org/quality-assurance-standards/>. Accessed January 20, 2024.

The field of conference interpreting is not the only one to be affected by the impact of remote interpreting. Dialogue and consecutive interpreting in public services are also concerned, but with specific challenges (Amato/Spinolo 2018: 7-8). Not least, the interaction between the participants in the communicative situation may be hampered by their unfamiliarity with or fear of the remote mode. This means the interpreter, as a communication expert, may need to take on a leading role to ensure a successful outcome of the exchange. Given that many interpreter training institutions prepare students for both conference interpreting and public service markets specific training seems all the more necessary.

In summary, traditional curricula already provided some familiarity with the handling of situations which resemble one type of remote interpreting at least – i.e. simultaneous interpreting from a hub – through the use of video input. However, the introduction of a dedicated didactic component that emphasises the idiosyncrasies for all kinds of remote is now offered by many courses. It is of value for the academic and professional preparation of an interpreter. If not in the form of actual university courses, which may be difficult to fit in the curricula, at least in the form of additional lectures and seminars. Obviously, it would be difficult to illustrate in detail the characteristics of the various remote options, and probably not very fruitful, given the speed of change in this field. However, an overview of existing platform types, together with the presentation of specific examples, would already give students some tools to navigate this field independently, and help them make informed choices in terms of equipment and methods, and decisions conducive their own well-being and health.

## Conclusion

This overview of the impact of remote interpreting on interpreters inevitably leaves many questions unanswered. The extent of remote, its modalities, and the associated working conditions are all in a state of flux at the time of writing. Remote has not displaced onsite interpretation, contrary to fears during the pandemic, but it is likely to remain part of the interpreting mix for the foreseeable future.

Remote interpreting is not a single entity. The term encompasses a range of situations – working with a team from a hub, working alone from home, interpreting a remote meeting in an institution's booths. If distance interpreting more broadly is considered, this can be interpreting an occasional remote speaker in an otherwise standard onsite configuration. The consequences for the interpreter and for the interpreting output will be very different each time.

Much of the criticism and alarm expressed within the profession about remote interpreting relates primarily to a fully remote mode, with interpreters working in isolation from home. However, with the lifting of health-related restrictions this modality is less prevalent. Remote and hybrid meetings are frequent, but interpreters are often on-site or in hubs, and can at least work together as a team.

Several lessons emerge clearly from the experience over the past few years. Firstly, out-of-hand rejection of remote interpreting is not an option. Interpreters need to be pro-active, engage with platform providers, with organisers, and with research. Professional associations have a major role to play in investigating and documenting

the consequences of remote, and in defending interpreting quality, and interpreters' health and well-being.

Remote interpreting has been around for many years, steadily becoming more prevalent, but it surged during the 2020-2022 Covid restrictions. Despite its many drawbacks, it is thanks to remote that meetings and interpretation could continue through the lockdowns. The two-year surge has subsided since, as clients, both private and institutional, have come to realize the limits of distant communication. Nonetheless, remote interpreting and remote participation are likely to remain a fixture of conferences, and hence of the interpreting market, both for whole events or for certain groups of participants.

As seen above, remote interpreting has numerous consequences for interpreters, and marks a major shift in their working environment. The consequences concern their auditory and psychological health, professional cohesion, or status, and, as a result, also training. Generally, remote does seem intrinsically more tiring and stressful than onsite participation. Some of the negative health-related consequences have been at least partially addressed by recent AIIC recommendations and agreements, and by new ISO standards, although not all platforms are compliant and generic platforms are not covered. Other consequences now seem less acute than initially feared, as remote becomes just one form of professional activity, alongside onsite interpreting. Many forms of remote still allow interpreters to work together physically in teams when they work from hubs or in proximity to the event, and even to have contact with meeting participants, when in adjacent areas.

Regarding the impact on training, the emerging consensus seems to be that students must be exposed to remote interpreting conditions, ideally in dedicated training modules or at the very least through exposure to video recordings. But remote can also be used to teach, and as such – at least in blended training – may provide new opportunities to extend quality training to geographies and demographics that would otherwise be excluded, and thus ultimately to enhance and standardize quality and ethical practices.

The remote interpreting landscape is still changing fast, and predictions are difficult in this complex environment. But on balance the situation seems less alarming than a few years back. Remote interpreting has opened up a more diverse professional landscape for clients and interpreters alike. It also should be said that individual interpreters and the profession as a whole have responded, not only with flexibility, but also with resolve and unity in addressing the new issues to which remote gives rise.

## References


Amato A. / Spinolo N. (2018) "Introduction" in A. Amato / N. Spinolo / M. J. González Rodríguez (eds) *Handbook of Remote Interpreting*, 7-10, <http://amsacta.unibo.it/id/eprint/5955/1/HANDBOOK_SHIFT.pdf>.

Baigorri-Jalón J. (1999) "Conference Interpreting: From Modern Times to Space Technology", *Interpreting* 4/1, 29-40.

Baigorri-Jalón J. (2004) *Interpreters at the United Nations. A History*, Salamanca, Ediciones Universidad de Salamanca.
Baigorri-Jalón J. (2014) *From Paris to Nuremberg. The birth of conference interpreting,* Amsterdam, John Benjamins Publishing.
Baigorri-Jalón J. / Fernández-Sánchez M. M. / Payàs G. (2022) "Historical developments in conference interpreting: An overview", in M. Albl-Mikasa / E. Tiselius (eds) *The Routledge Handbook of Conference Interpreting*, Oxon, Routledge, 9-18.
Boéri J. (2015) "Key internal players in the development of the interpreting profession" in H. Mikkelson / R. Jourdenais (eds) *The Routledge Handbook of Interpreting*, London & New York, Routledge, 29-44.
Braun S. (2015) "Remote Interpreting", in H. Mikkelson / R. Jourdenais (eds) *The Routledge Handbook of Interpreting*, Oxon, Routledge, 352-367.
Braun S. / Taylor J. L. (2011) "Video-mediated interpreting: An overview of current practice and research" in S. Braun / J. L. Taylor (eds) *Videoconference and remote interpreting in criminal proceedings*, Guilford, University of Surrey, 25–57.
Buján M. / Collard C. (2023) "Remote simultaneous interpreting and COVID-19: conference interpreters' perspective", in K. Liu / A.F.K. Cheung (eds) *Translation and Interpreting in the Age of COVID-19*, Berlin, Springer, 133-150.
Chia-Ming Fan D. (2022) "Remote Simultaneous Interpreting: Experiences and Opinions of Conference Interpreters in Taiwan", *Compilation and Translation Review* 15/2, 159-198.
Corpas Pastor G. / Gaber M. (2020) "Remote interpreting in public service settings: technology, perceptions and practice", *SKASE Journal of Translation and Interpretation*, 13/2, 58-78.
Dehankar S. S. / Gaurkar S. S. (2022) "Impact on Hearing Due to Prolonged Use of Audio Devices: A Literature Review", *Cureus* 14/11, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9747083/>.
Diriker E. (2004) *De-Re-Contextualizing Conference Interpreting: Interpreters in the Ivory Tower?* Amsterdam, John Benjamins.
Donovan C. (2023) "The Consequences of Fully Remote Interpretion on Interpreter Interaction and Cooperation. A threat to professional cohesion?", *INContext* 3/1, 24-48, <https://incontextjournal.org/index.php/incontext/article/view/59>.
Duflou V. (2016) *Be(com)ing a Conference Interpreter. An ethnography of EU interpreters as a professional community*, Amsterdam, John Benjamins Publishing Company.
Dwyer T. (2010) "Traducteurs, Associations Professionnelles et Marché: Approches Empiriques", Hermes 56, 77-82.
Fantinuoli C. (2018) "Interpreting and technology: The upcoming technological turn" in C. Fantinuoli (ed) *Interpreting and technology*, Berlin, Language Science Press, 1-12.
Furmanek O. (2012) "Professionalization of Interpreters", in C.A. Chapelle (ed) *The Encyclopedia of Applied Linguistics*, New Jersey, Wiley Blackwell.

Gentile P. (2015) *The interpreter's professional status: a sociological investigation into the interpreting profession*, PhD thesis, SSLMIT, University of Trieste.
Gile D. (2017) "Testing the Effort Model's Tightrope Hypothesis in Simultaneous Interpreting – A Contribution", *Hermes* 23, 153-172.
Giustini D. (2022) "COVID-19 and the Configuration of Materiality in Remote Interpreting: Is Technology Biting Back?", in: L. Kanglong / A. K. F. Cheung (eds) *Translation and Interpreting in the Age of COVID-19*, Singapore, Springer, 197-213.
Gupta A. / Bakshi S. S. / Kakkar R. (2022) "Epidemiology and Risk Factors for Hearing Damage Among Adults Using Headphones via Mobile Applications", *Cureus* 14/5, <https://pmc.ncbi.nlm.nih.gov/articles/PMC9245586/>.
Han L. / Wang Y. / Li Y. (2022) "Student Perceptions of Online Interpreting Teaching and Learning via the Zoom Platform", *The Electronic Journal for English as a Second Language*, 26/1, 1-19.
Ho C. / Zou Y. (2022) "Teaching Interpreting in the Time of COVID: Exploring the Feasibility of Using Gather" in L. Kanglong / A. K. F. Cheung (eds) *Translation and Interpreting in the Age of COVID-19*, Singapore, Springer, 311-330.
Hodáková S. (2021) "Mental Health and Well-being of Students in the Context of Remote Training of Future Translators and Interpreters", *Bridge: Trends and Traditions in Translation and Interpreting Studies*, 2/2, 63-76.
Horváth I. / Tryuk M. (2022) "Ethics and codes of ethics in conference interpreting" in M. Albl-Mikasa / E. Tiselius (eds) *The Routledge Handbook of Conference Interpreting*, Oxon, Routledge, 290-304.
Hynes R. (2021) "Assessing the Risk Factors for Hearing Problems amongst Simultaneous Conference Interpreters", in M. Phelan (ed) *Translation Ireland*, Cork, Lettertec Publishing, 163-188.
Jacob H. (1962) *Kind meiner Zeit. Lebenserinnerungen* (in English, "Child of my time. Memoir"), Köln, Kiepenheuer und Witsch.
Jiménez Ivars A. / Pinazo Calatayud D. (2001) "«I failed because I got very nervous». Anxiety and performance in interpreter trainees: an empirical study", *The Interpreters' Newsletter* 11, 105-118 <https://www.openstarts.units.it/entities/publication/10aad7b8-94c8-4c50-bf32-2a3ead92f2bd/details>.
Kyong-Jo W. / Hyang-Ok L. (2022) "So Far Away and Yet So Close: A study of remote interpretation in the classroom", *INContext*, 2/2, 92-111.
Lee J. (2020) "A Neuropsychological Exploration of Zoom Fatigue", *Psychiatric Times*, <https://www.psychiatrictimes.com/view/psychological-exploration-zoom-fatigue>.
Mahyub Rayaa B. / Martin A. (2022) "Remote Simultaneous Interpreting: perceptions, practices and developments", *The Interpreters' Newsletter* 27, 21-42, <https://www.openstarts.units.it/server/api/core/bitstreams/3a8098c9-3e4d-4a4c-8e98-ee5497039f3a/content>.
McFerran D. J. / Baguley D. M. (2007) "Acoustic shock", *The Journal of Laryngology & Otology* 121, 301-305.
Monacelli C. (2009) *Self-preservation in Simultaneous Interpreting: Surviving the Role*, Amsterdam, John Benjamins.

Moser-Mercer, B. (2003) Remote interpreting: Assessment of human factors and performance parameters. Joint project, International Telecommunication Union (ITU). Ecole de traduction et interprétation, University of Geneva (ETI).

Moser-Mercer B. (2005a) "Remote interpreting. The crucial role of presence", *Bulletin suisse de linguistique appliquée*, 81/2, 73-97.

Moser-Mercer B. (2005b) "Remote Interpreting: Issues of Multi-Sensory Integration in a Multilingual Task", *Meta*, 50/2, 727-738.

Mouzourakis P. (2006) "Remote Interpreting: A technical perspective on recent experiments" in *Interpreting* 8/1, 45 – 66.

Murgu D. (2021) "Approaching Stress and Performance in RSI: Proposal for Action to Take Back Control" in R. Mitkov / V. Sosoni / J. C. Giguere / E. Murgolo / E. Deysel (eds) *TRanslation and Interpreting Technology Online – Proceedings of the Conference*, 68-75, <https://aclanthology.org/2021.triton-1.pdf>.

Rangponsumrit N. / Rattanabutr P. (2016) "Teamwork in Simultaneous Interpretation", *Journal of Translation and Interpretation Thailand*, 1/2, 61-87.

Ratan R. / Miller D. B. / Bailenson J. N. (2022) "Facial Appearance Dissatisfaction Explains Differences in Zoom Fatigue", *Cyberpsychology, Behavior, and Social Networking*, 25/2, 124-129, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8864415/>.

Riccardi A. / Marinuzzi G. / Zecchin S. (1998) "Interpretation and stress", *The interpreters' newsletter* 8, 93-106, <https://www.openstarts.units.it/entities/publication/91f664ca-a98b-4874-88da-73eb665a4a4e/details>.

Riedl R. / Kostoglou K. / Wriessnegger S. C. / Müller-Putz G. (2023) "Videoconference fatigue from a neurophysiological perspective: experimental evidence based on electroencephalography (EEG) and electrocardiography (ECG)", *Scientifica Reports*, 13/18371.

Roziner I. / Shlesinger M. (2010) "Much ado about something remote: Stress and performance in remote interpreting", *International journal of research and practice in interpreting*, 12/2, 214-247.

Saina F. (2021) "Remote Interpreting: Platform Testing in a University Setting" in R. Mitkov / V. Sosoni / J. C. Giguere / E. Murgolo / E. Deysel (eds) *TRanslation and Interpreting Technology Online – Proceedings of the Conference*, 57-67, <https://aclanthology.org/2021.triton-1.7/>.

Salaets H. / Brône (2023) "Working at a Distance from Everybody: Challenges (and some advantages) in working with video-based interpreting platforms". *The Interpreters' Newsletter*, 2023 (28), 189-209.

Seeber K. G. / Fox B. (2021) "Distance conference interpreting", in M. Albl-Mikasa / E. Tiselius (eds) *The Routledge Handbook of Conference Interpreting*, Oxon, Routledge, 491-507.

Şengel Z. (2022) "Zooming in: Interpreters perspective towards remote simultaneous interpreting (RSI) ergonomics", *Çeviribilim ve Uygulamaları Dergisi*, 33, 169-190.

Setton R. / Dawrant A. (2016) *Conference Interpreting: A Trainer's Guide*, Amsterdam-Philadelphia, John Benjamins Publishing Company.

Taylor-Bouladon V. (2007) *Conference Interpreting. Principles and Practice*, Charleston, BookSurge Publishing.

Tejada Delgado A. (2019) “Is the public sector interpreting market ready for digital transformation?”, *Revista Tradumàtica. Tecnologies de la Traducció* 17, 88-93.

Ünlü C. (2021) “Remote Simultaneous Interpretation in Pre and Post-COVID-19: An Overview of the Profession Revisited from Sectoral, Ethical, and Pedagogical Perspectives” in T. Didem / K. Mesut (eds) *Interdisciplinary Debates on Discourse, Meaning and Translation*, Ankara, Anı Yayıncılık, 311-337.

Ziegler K. / Gigliobianco S. (2018) “Present? Remote? Remotely present! New technological approaches to remote simultaneous conference interpreting”, in C. Fantinuoli (ed) *Interpreting and technology*, Berlin, Language Science Press, 119-139.

**The Interpreters' Newsletter Special Issue 2024 (29 bis), 31-58**
DOI: 10.13137/2421-714X/36764 ISSN 1591-4127 (print) ISSN 2421-714X (online)
https://www.openstarts.units.it/dspace/handle/10077/2119

# I've never heard about it...! Speech-to-text interpreting in Italy and Austria: a comparative analysis

Alessia Valente
Conference Interpreter
Judith Platter
University of Vienna*

Abstract

*Speech-to-text interpreting (STTI) is increasingly gaining traction as an effective service to ensure access to communication for people encountering a hearing barrier. Significant developments in offering this service and training professionals have been made in the Scandinavian and English-speaking countries for more than three decades; the German-speaking countries, Germany and Austria, following in the early and late 2000s. In other countries, such as Italy, however, this service still remains relatively unknown and therefore is not used to its full potential. This paper proposes a comparative analysis of the professionalisation status in terms of awareness, use and spread of STTI in Italy and Austria, with the aim of highlighting the gaps in the Italian context and providing a possible model to address them. To explore this topic, the paper also briefly addresses a closely related issue, namely accessible communication, which is often neglected in Italy. Furthermore, a theoretical overview of the main features and functioning of STTI is provided. This article emphasises the need to make progress in Italy with regard to accessible communication and, in particular, the need to recognise STTI as a service capable of significantly improving the quality of life for a substantial segment of Italian society.*

* Although this article is a result of joint discussion, Valente is the author of Sections 1, 2 and 4; Platter of Section 3.

Keywords

Speech-to-text interpreting, hearing barriers, accessibility, accessible communication, post-lingual hearing barriers.

## Introduction

Accessible communication refers to the provision of services that enable those in need of specific support to participate in specific communicative acts. Ensuring accessibility in communication is crucial. As Aristotle asserted as early as the 4th century BC in his work "Politics" (1999), humans are social animals, naturally inclined to live in communities and establish social bonds. From this, it can be deduced that interaction with others is one of the essential characteristics of being human. Therefore, depriving a person of access to communication means denying them access to a fundamental dimension of their life.

The value of accessibility in communication is not, however, always adequately recognised. Indeed, typically, when accessibility is mentioned in general terms, it is immediately associated with overcoming physical or architectural barriers, such as a lack of ramps at the entrance of a building. "This approach however is only partially correct" (Perego 2020: 21), as the word encompasses a broader range of barriers. The United Nations Convention on the Rights of Persons with Disabilities (UNCRPD) of 2006 emphasises this aspect, providing in Article 9 a list of areas that are included in the concept of accessibility:

> (a) Buildings, roads, transportation and other indoor and outdoor facilities, including schools, housing, medical facilities and workplaces;
> (b) Information, communications and other services, including electronic services and emergency services (UN 2006: 9).

Another relevant aspect related to accessibility is the common perception that it concerns only the inclusion of people with disabilities. This is a limited view that does not reflect the true scope of the concept because, on the contrary, accessibility can affect anyone. This is particularly evident in the context of accessible communication. A text[1] can become a barrier when it has characteristics that prevent fair access to any group of users. Among communicative barriers (cf. Schubert 2016; Rink 2020), one can identify not only those inherently linked to disabilities, such as sensory or intellectual barriers, but also those stemming from factors that are external to acknowledged disabilities. For instance, a text written in a language not understood by users creates a language barrier, which can be overcome through translation or interpretation. Similarly, if a text includes references to specific cultural aspects, it becomes inaccessible to those from a different culture who lack the necessary cultural prerequisites to understand those references. For these reasons, it is evident that accessible communication concerns not only specific groups of people but is a matter that can involve anyone at various points and means of communication in life. In this

1 In this paper, the word "text" refers to both an oral text and a written text.

context, accessible communication advocates a more neutral denomination, replacing terms like disability with the more common and identifiable term of "barrier"[2].

This article focuses on an accessible communication service primarily aimed at people encountering hearing barriers in the sense of hearing loss in different degrees but which, as will be seen, can benefit many other user categories. Currently, ensuring inclusion for people with hearing barriers largely involves intersemiotic sign language interpreting (SLI) (cf. Grbić 2023), which is recognised and protected by legislation in numerous European countries – for interpreted communication in this case, the service user has to understand sign language.[3]

However, in recent years, following advancements in technology, audiovisual translation and artificial intelligence, there has been the emergence of speech-to-text interpreting (STTI), which involves a simultaneous written transcription of a spoken text in various contexts and text formats – for interpreted communication in this case, the service user[4] has to be able to read written texts. This form of interpretation is increasingly spreading across various European countries such as Sweden, Finland, Austria and Germany. These countries have robust research activity in this field and have had laws mandating the provision of STTI services for some years. Additionally, there are numerous associations promoting the use of STTI and several universities and training centres offering courses to learn how to provide this service (cf. Valente 2024). The situation is different in Italy, where initial research indicates that STTI is still relatively unknown, and communication accessibility is often overshadowed by the more established focus on physical accessibility.

The aim of this paper is to highlight the shortcomings within Italy in the field of STTI and, more broadly, in accessible communication. To achieve this, a comparative analysis is presented of the current state of STTI dissemination in Italy and in Austria, a neighbouring country where the situation is significantly more favourable and could serve as a benchmark. Furthermore, with the goal of showcasing the versatility and utility of STTI, a brief theoretical overview of its main characteristics and functioning will be provided.

2 In this paper, we therefore use the term "barrier" in its comprehensive and inclusive approach, except in those cases in which we refer to laws and regulations still citing the term "disability" or "impairment".

3 For example, Austria recognised its own sign language in its Constitution in 2005, Germany in 2002 through the Disabled Equality Act (Behindertengleichstellungsgesetz, our translation) and Italy in 2021 with the Act on Urgent Measures for Companies and Workers due to COVID-19 of May 21, 2021, n.69 (Conversione in legge, con modificazioni, del decreto-legge 22 marzo 2021, n. 41, recante misure urgenti in materia di sostegno alle imprese e agli operatori economici, di lavoro, salute e servizi territoriali, connesse all'emergenza da COVID-19, our translation).

4 We wish to point out that in this paper with the term "user" we refer to primary users of STTI (people with a hearing barrier) and secondary users of STTI (people with other communicative barriers) alike. The first group, in some publications, is referred to as "recipients" (Eichmeyer 2017).

## 1. STTI: a theoretical overview

### 1.1 Definition and general features

STTI, also called captioning, written interpreting, or speech-to-text reporting (Nofftz 2014), is a type of interpretation that began to spread in the 1970s in the United States and in Scandinavian countries, particularly in Sweden, Finland, and Norway (Witzel 2020). This service is primarily aimed at people encountering hearing barriers and involves the simultaneous transcription of a spoken text by an interpreter in an interaction where one of the interactants has a hearing barrier. The transcription includes all verbal, paraverbal, and non-verbal features necessary for the comprehension of a communicative act. Consequently, not only are the spoken words reported, but punctuation marks, text paragraphing, indications of the speaker's intonation and voice volume, speech rhythm, and other non-verbal sounds such as laughter and sighs are also included. Additionally, further information is generally provided, such as the speaker's name, the use of dialects, and the use of specific body language induced by auditory stimuli – for example, if the audience, upon hearing a noise from outside, looks out of the window (Witzel 2020)[5].

Therefore, the final product is a written text with certain characteristics of orality. To describe the relationship between written and spoken language in this process, Tiittula refers to it as "orality clothed in writing"[6] (Tiittula 2006 translated in Eichmeyer 2017: 6).

In the consulted literature, this service is categorised as a form of interpretation (cf. Tiittula 2006; Nofftz 2014; Platter 2015), as it fulfils the primary role of interpretation, which is to provide access to the content of oral communication. (Platter 2019). Moreover, it shares many features with simultaneous interpretation, such as the division of attention between listening and reproducing and the limited ability to correct the final product (cf. Gerzymisch-Arbogast 2013). Unlike simultaneous interpretation, however, STTI is mostly performed intralinguistically[7] and is considered a diasemiotic type of interpretation, as there is a change in the semiotic channel from oral to written. Furthermore, it often involves unidirectional work, where the interpreter transcribes only what is said to the user. This occurs because, in most situations, users can express themselves orally without any difficulty (see 1.2).

STTI is also described as a form of translation hybrid (*Translationshybrid*), where the oral input is simultaneously transformed into a written output. This means it can be classified both as interpretation, due to the oral nature of the source text, and as translation, given the written nature of the target text (Tirinnanzi 2016: 137).

In a broader sense, STTI is amongst the services provided within the scope of accessible communication, alongside easy language, sign language, subtitles for the

5 In most cases, these paraverbal and non-verbal indications are placed in parentheses. For example, (laughs) is written at the end of a sentence if the speaker laughs, or (ironic) to indicate that the sentence is characterised by a sense of irony (Tirinnanzi 2016; Platter 2019).

6 The original expression in German is *Mündlichkeit im schriftlichen Gewand* (Tiittula 2006: 481).

7 Although STTI is mostly requested at an intralinguistic level, in recent years interlinguistic STTI has also been gaining ground, particularly at international events (Eichmeyer 2017: 6).

deaf and hard-of-hearing and audio description for people with visual disabilities (Maaß 2020).

### 1.2 Target audience

STTI primarily targets people encountering some kind of hearing barrier. Specifically, the focus is on people who have developed a post-lingual deafness or hearing impairment, starting from the age of five, for example, due to an accident, illness or old age[8]. These are people who have lost their hearing partially or completely but have had the opportunity to develop the necessary language acquisition skills, enabling them to communicate orally without major difficulty. Unlike those with pre-lingual hearing impairments, such as those who are born deaf or hard-of-hearing or who become so within the first five years of life, this group often does not belong to the signing deaf community as they have not learned sign language and communicate through spoken language and lip-reading (Tirinnanzi 2016: 138-139). For these reasons, it can be inferred that the primary challenge for this category of people is not expressing themselves, but rather, hearing what is communicated to them. To address this difficulty, cochlear implants or hearing aids can be used, although they do not always ensure adequate comprehension of speech (Tirinnanzi 2016: 139). In this context, STTI emerges as a critical tool to provide people with post-lingual hearing barriers with full access to oral communicative situations.

It is interesting to note that among STTI users, people with pre-lingual hearing barriers can also be included. Although, as previously mentioned, most of these people can communicate using sign language, significant preferences have been found for communication through STTI services. For example, in a study conducted by Stinson *et al*. (2017), it emerged that according to some students with pre-lingual hearing barriers in American high schools and universities, STTI services were more beneficial compared to handwritten note-taking and SLI. These students found that STTI better supported their concentration and understanding during classes, as the projected words remained on the screen for about a minute, allowing them more time to assimilate the information. Another important aspect is that students could obtain a written text of the lecture content[9]. This way, they did not need to take notes independently and could revisit what was said later (cf. Stinson *et al*. 2017). Moreover, thanks to STTI, students had the opportunity to focus on just two actions: reading the transcribed text and observing the speaker. In the case of sign language, there were three actions to focus on: observing the interpreter, taking independent notes and observing the speaker. In the context of this group of users, it is important to consider that they have often had limited exposure to verbal language and may therefore experience difficulties in reading. To address this situation, the solution may involve transcribing the

8 German Society of People with Hearing Disabilities (2004) <https://www.deutsche-gesellschaft.de/fokus/einige-informationen-zum-thema-hoerschaedigung>

9 In some countries, such as Austria, in schools and universities the final written text is provided to students free of charge. However, there are cases, for instance in professional or legal contexts, where the availability of the text is considered a separate service from the interpretation itself and therefore entails an additional cost (cf. Eichmeyer *et al*. 2017).

spoken text using plain language or easy language, a practice that has begun to spread recently (cf. Eichmeyer 2017; Platter 2019).

The use of plain language or easy language makes STTI a service accessible also to people with cognitive barriers in terms of intellect and general understanding (Platter 2019; Witzel 2020). While this group may encounter greater difficulty in reading compared to listening to an oral discourse, providing a written text in a simplified language could facilitate their understanding, especially in situations where the speaker uses complex language, speaks rapidly, or has an accent not readily comprehensible.

As previously mentioned, STTI can also be useful to those encountering communication barriers unrelated to an acknowledged disability. A clear example is the instance of people with limited or no understanding of a language used in a speech, who may find it easier to follow the oral text through a written transcription. In the former case, a simplified language transcription could be employed, akin to what has been proposed for people with cognitive barriers. In the latter case, one could opt for interlinguistic STTI, a practice that involves the simultaneous transcription and translation from one language to another (see footnote 8).

Users of STTI can access the text through its display on a specific device. The mode of text display may vary depending on the number of people relying on this service and their specific needs (Eichmeyer 2017; Platter 2019). If the group of users is large, the speech can be projected as continuous text on a screen installed in the room. Alternatively, a system like Text on Top[10] can be used to transmit the text as subtitles or surtitles on a screen or personal device. Another option is to use an online platform, such as Text on Tap, which users can access to view the text through their computer, tablet or phone (Eichmeyer 2017; Platter 2019). If the group of users consists of only one or two people, the text can be read directly on the interpreter's device monitor or on external monitors that are connected to the interpreter's computer or tablet.

### 1.3 Fields of application

In addition to the various categories of potential STTI users, there are also many fields in which it can be employed. Indeed, it can be used both in broader contexts such as seminars, fairs, conventions, conferences, festivals or parliamentary sessions, and in more specific situations like medical examinations, business meetings or courtroom hearings (Platter 2019).

The use of STTI services is highly valued in schools and universities as well. Despite being currently less widespread than SLI, it is considered to be more advantageous to students for several reasons. Alongside the findings from the study conducted by Stinson *et al.* in 2017, STTI proves to be more effective than SLI, especially when used for subjects with complex or specific terminology, like science or mathematics (Valente 2024). This is because often there are no corresponding signs and, therefore, it is necessary to elaborate them from scratch. Furthermore, STTI is more advantageous in foreign language learning, where it is important to learn not only individual words but also their pronunciation (*Ibid.*). Indeed, thanks to STTI, it is possible both

10 Text on Top (2015) "What is Text on Top?", <https://text-on-top.com/en/what-is-text-on-top/>

to transcribe words and to provide additional information on correct pronunciation. Another benefit of introducing STTI in school and university settings is the possibility of making the interpreter "invisible". A SLI needs to be clearly visible to the user – otherwise the user would struggle to follow the discourse. This visibility can cause discomfort for both the user, who might feel stigmatised[11] in the classroom, and for other students, who could be bothered by the presence of an external person. In the case of STTI, however, this issue can be addressed. As will be seen, online platforms can be used so that the interpreter can work remotely and does not need to be physically present. Alternatively, if the interpreter is in the classroom, they can work from a less conspicuous location[12] to avoid interfering with the classroom dynamics.

Another significant context where STTI is widely used is in the fields of television and film, where it is specifically known as live subtitling. Live subtitling can be considered a specific application of STTI focused on creating subtitles, for example, during live broadcasts or film festivals and it differs from other uses of STTI in several ways. For instance, professionals providing subtitles in the television and film sectors often lack direct contact with speakers or users, limiting their ability to seek explanations and clarifications or to tailor their services to the specific needs of users, as they primarily work behind the scenes. In contrast, in other applications of STTI such as medical visits or meetings, such interactions are more frequent and extremely important to ensure a tailor-made service (Norberg & Stachl-Peier 2017: 164).

### 1.4 Working techniques

Before the arrival of computers, STTI services were commonly delivered using pen and paper, which allowed for the annotation of approximately 30 words per minute (Nofftz 2014: 3). In this way, only a general account of the communicative situation could be offered, and true access could not be guaranteed. Although technological advancements have led to the development of more efficient resources, this practice remains useful in contexts where short communicative contents need to be transcribed or where technical issues prevent the use of more modern equipment (*Ibid*.).

Currently, the main working methods used for STTI include the conventional QWERTY keyboard, the Velotype keyboard[13], the stenotype keyboard and speech recognition and they vary according to the language for which they are used as well as the contexts where STTI is needed and provided first hand.

The conventional method adopted by professionals involves using the computer's QWERTY keyboard in combination with text processing software, such as MS Word. In addition, specific programmes can be used[14], in which abbreviations of long words and recurring phrases as well as autocorrection options for recurring errors can be

11 This could stem from feeling "different" from others and from feeling responsible for the presence of an outsider in the classroom or lecture hall.

12 For example, they can work in a corner at the back of the classroom or lecture hall.

13 Velotype (2024) <https://www.velotype.com/en/homepage-eng/>

14 These autocorrection options are either included within the text processing programme itself or run on the operating system (e.g. Atext (2024) <https://www.trankynam.com/atext/?ref=chiaracokieng.com>, AutoHotkey (2024) <https://autohotkey.com/download/>,

stored. These abbreviations automatically turn into their full and/or correct versions once typed (Tiittula 2018: 199-200).

The conventional keyboard has advantages, such as the ability to use it efficiently after a short training period and the integration of free software (Nofftz 2014: 13). Furthermore, users often appreciate the fact that by typing the word or abbreviations turning into words or phrases, the reading of the text is facilitated, as the letters and words appear in a linear way. There are, however, also significant disadvantages. Indeed, with the use of the QWERTY keyboard, one can reach a maximum typing speed of about 90 words per minute (Platter 2019: 124), which is much slower than an average speaking speed, which ranges from 129 to 222 words per minute (*Ibid*.). Furthermore, the physical and mental effort required from interpreters to type as quickly as possible in order to report as many words as possible from the speaker is considerable and can undermine the overall quality of the work (Witzel 2020: 315).

To increase typing speed, the Velotype keyboard can be used (Velotype 2024). Designed in the Netherlands in the 1980s, it is particularly widespread in Sweden, the Netherlands and France (cf. Nofftz 2014). The Velotype is a chords keyboard, meaning that "chords" of keys are pressed simultaneously to write whole syllables or words (Witzel 2020). With this type of keyboard, a typing speed of 200 words per minute can be achieved (Platter 2019: 123). However, this tool is not very common or in use for languages other than Dutch because using it effectively often requires over a year of training (Norberg *et al*. 2015: 39). Additionally, purchasing the Velotype keyboard involves a significant financial cost, making it less accessible compared to the conventional computer keyboard.

Another keyboard commonly used in English-speaking countries (cf. Nofftz 2014) is the stenotype keyboard, which is connected to a computer and specialised software. This technique is also known as communication access real-time translation or computer-compatible steno machines (Wagner 2005: 4; Nofftz 2014: 4). It is a chords keyboard where phonetic abbreviations of words are typed using specific key combinations. These abbreviations are then transmitted to the computer and converted into written text by the software. With this method, the performance of the Velotype keyboard is exceeded, as it enables the attainment of a typing speed of over 200 words per minute. However, the use of this keyboard is not widespread either, as acquiring the necessary skills requires at least three years of training, and the cost of the complete equipment can reach up to €10,000 (Wagner 2005: 8).

The most innovative method that has emerged in the field of STTI is speech recognition in combination with respeaking. Initially, this method was mainly used to produce subtitles for the deaf and hard-of-hearing during live television programmes but it is now also common in other contexts.

In this case, the interpreter repeats the speaker's words almost simultaneously through shadowing and adds other relevant information for understanding the speech through specific vocal commands, including punctuation and paragraph breaks. These words are then converted into written text by specialised software that can be downloaded directly to one's personal computer and exploits speech recognition technology.

PhraseExpress (2024) <https://www.phraseexpress.com/de/doc/textbausteine-verwalten>, TextExpander (2024) <https://textexpander.com>

Below is presented an example of how a text should be interpreted while using respeaking:

Good morning [comma][15], welcome to this event [period].
[new paragraph]
Today we are going to discuss a very important topic [period].

In most cases, speaker-dependent software[16] is used. This means that it is possible for the interpreter to train the software to recognise the specific characteristics of their voice, such as intonation and pronunciation, in order to optimise the speech recognition process. An example of such software is LilySpeech[17], which provides a free service and allows users to transcribe what is dictated to their computer across various applications, including Notepad, Microsoft Word, and Google Chrome. Among the most well-known and widespread software is Dragon NaturallySpeaking[18], which offers a service for payment.

The use of respeaking enables attainment of a typing speed of 160-190 words per minute (Platter 2019: 125); in contrast to the keyboard-based techniques, the output in words and sentences is not as linear, as the software writes bunches of words or sentences. This, for certain users, may result in a major reading effort; however, the costs associated with this technique are lower compared to those of the Velotype keyboard and the stenotype keyboard. Indeed, this method has been increasingly adopted in recent years, particularly for conferences and seminars (Tirinnanzi 2016).

## 1.5 Work settings

With regards to STTI work settings, depending on where the STTI user and STT interpreters are located, three kinds can be distinguished: in-presence setting, online setting and semi-presence setting (cf. Eichmeyer 2017; Platter 2019); while within an in-presence setting, the STTI user and the STT interpreters share the same physical space, in online settings either the STT interpreters or even the STTI user are using a remote device and share a digital space. In semi-presence settings, one STT interpreter shares the same physical space with the STTI user; the other one joins in the digital space.

First, it is important to emphasise that for assignments longer than 60 minutes, interpreters work in pairs, just as in traditional simultaneous interpreting, and typically switch every 15 minutes (Eichmeyer 2017: 8). They work together on the same text using specialised programmes[19]. This allows them to help each other if needed. For instance, the interpreter who is not actively interpreting at that very moment can support

15 The commands in square brackets are the vocal instructions that the interpreter must give to the speech recognition software.
16 Speaker-dependent software differs from speaker-independent software, which cannot be trained and responds equally to any type of voice.
17 LilySpeech (2024) <https://lilyspeech.com/>
18 Nuance (2024) <https://www.nuance.com/dragon.html>
19 E.g. Text on Top or Gobby. Gobby is a free software programme that allows several people to work together on a document and provides them with the ability to communicate through

their colleague by correcting any errors in the written text (Witzel 2020: 318). This form of collaboration to improve the quality of the interpreted live text is called co-editing.

Typically, if there are many people who need STTI, interpreters work in a booth or at a table separate from the audience, whereas if the service is intended for one or two people, as in the case of a medical appointment, the interpreters can sit directly next to them (Platter 2019: 126-127). In these examples, interpreters operate from a fixed position, meaning their work is stationary. There are, however, some contexts, such as tourist or business visits, where interpreters are required to move around and thus work in a mobile setting.

According to Witzel (2020), the on-site setting offers the advantage of establishing direct contact with the user or users and the speaker, which improves the working environment. However, this approach can be associated with some issues, such as poor audio quality due to background noise and significant travel or accommodation expenses for the interpreters. Moreover, interpreters specialised in STTI are still few in number and unevenly distributed across countries, at least when it comes to Austria and Germany. As a result, it is often difficult to find two interpreters available to work in a specific geographic area (Eichmeyer 2017: 10).

Even before the boost in remote interpreting activities during and after the COVID pandemic in 2020, STTI in German-speaking countries adopted remote working assignments, where interpreters connect online via a specialised platform and follow the event remotely – in the majority of cases before the pandemic, access was limited to audio input of the original text only. Several companies providing STTI services therefore developed platforms that allow interpreters to listen to the speaker, transcribe the speech and communicate with the users via chat if necessary[20]. This setting represents a significant advance as it eliminates travel and accommodation expenses for interpreters and makes it easier to find more professionals specialised in STTI, thus overcoming geographical barriers; however, direct contact with the speaker and the clients is lost. Additionally, if there are problems with the internet connection, interpreters may experience poor audio or video quality of the speaker. This can vitiate the final output, as crucial linguistic and non-linguistic references necessary for understanding the speech might be missed (Tirinnanzi 2016: 142).

The third and final work setting is the semi-presence setting, which represents a middle ground between the previously described settings. In this case, there are two interpreters: one present at the event location and the other connected remotely. According to Eichmeyer (2017), working in a semi-presence setting allows the combination of the benefits of both on-site and online work. On the one hand, the on-site interpreter can directly address any technical problems and has direct contact with the speaker and users. On the other hand, the remote interpreter helps reduce travel and

a dedicated chat (https://gobby.github.io/). Text on Top is a software developed in the Netherlands that STT interpreters use frequently.

20 Examples are the German companies VerbaVoice which developed the VerbaVoice System platform <https://www.verbavoice.de/>, as well as Kombia <https://www.kombia.de/> which provides these services via its solution Komline. In Switzerland, Swisstxt has a comparable solution <https://swisstxt.ch/services/accessibility>All these platform solutions may only be used in assignments handled exclusively by the company itself; their use by freelance STT interpreters is not allowed which might result in a possible drawback for the latter.

accommodation costs. In addition, finding an interpreter willing to work in presence is simpler, as only one interpreter needs to be available.

## 2. STTI in Italy

### 2.1 Terminology

In Italy, STTI is a service that is still relatively unknown. The lack of attention to this service is already evident from the way it is named in Italian. Indeed, it is commonly referred to as *sottotitolazione in tempo reale*, which is the equivalent of the English term "live subtitling"[21]. This designation can be considered restrictive and inadequate, since it refers specifically to the application of STTI within the television and film setting (see 1.3). The use of the term *sottotitolazione in tempo reale* does not fully reflect the wide range of contexts in which the service can be employed. Moreover, it wrongly suggests that the written text is displayed exclusively through surtitles or subtitles, when this mode of display is just one of the available options.

A possible solution to address this terminological shortcoming would be to introduce a new term in Italian that conveys the essence of the service provided. First and foremost, it would be appropriate to include a direct reference to interpreting, as STTI is essentially a type of interpretation. Furthermore, it would be necessary to emphasise the transition from spoken to written language – for example, English does this with the addition of "speech-to-text" or "written", while German, Dutch and Swedish use their respective prefixes[22] *Schrift-*, *schrijf-* and *skriv-*. Based on these considerations, a suitable expression for Italian could be *interpretazione scritta* (cf. Valente 2024), which would be the equivalent of the English "written interpreting", the German *Schriftdolmetschen*, the Dutch *schrijftolk*, or the Swedish *skrivtolkning*.

Such a formulation would, on the one hand, be specific enough to express the transition from spoken (*interpretazione*) to written language (*scritta*), and, on the other, generic enough to encompass any field of application of the service. Additionally, promoting the adoption of a clear, concise and comprehensive term could be crucial for clarifying this highly versatile service and increasing awareness of its importance within the Italian context.

In this way, one could make use of both the umbrella term *interpretazione scritta* and the term *sottotitolazione in tempo reale* to specifically indicate the service's application in the television and film sectors.

### 2.2 Regulatory basis

To understand Italy's potential for fostering STTI on all levels, one must also consider its legislative situation regarding accessible communication and STTI as a specific

21 This is not surprising, as STTI in Italy is predominantly used in the television and film sectors. As will be discussed, the academic literature on the subject also focuses mainly on this field. This situation is comparable to that of other European countries, such as Spain (cf. Romero-Fresco 2020).

22 The prefixes refer to writing.

form of it (cf. Valente 2024). Currently, at a national level, there are few laws that specifically refer to accessible communication and there are no laws dedicated to STTI. However, there are some provisions aimed generally at providing accessibility for people with disabilities. For instance, Law No. 104/1992, "The framework law for assistance, social integration, and the rights of disabled persons"[23], emphasises the Italian Republic's commitment to removing disabling conditions that hinder an individual's development and participation in social life. In this law, the primary focus is on physical accessibility. Accessible communication is addressed exclusively in article 25, which provides vague guidelines regarding the installation of decoders and complementary equipment for broadcasting and telephone services. Another relevant law for people with disabilities is Law No. 4/2004 "Provisions to facilitate access to information technology for disabled individuals"[24]. Specifically, it regulates the accessibility of public service information and telematic services.

With regard to the community of people with hearing barriers, the primary users of STTI, there are various regulations that govern their integration in the educational, university, healthcare and employment sectors. A pivotal moment for this community was marked by the enactment of Law No. 145/2018 "State Budget for the Financial Year 2019 and Multi-Year Budget for the Three-Year Period 2019-2021"[25]. In this law, article 1, subsection 456 states that a fund has been established for the inclusion of deaf and hard-of-hearing people, aiming to promote "the implementation of experimental projects for the dissemination of Italian Sign Language (LIS) interpreting services and remote video interpreting, as well as to encourage the use of innovative technologies aimed at removing communication barriers" (Law No. 145/2018: 76; our translation).

Subsequently, in 2021, there was the official recognition of Italian Sign Language (LIS) through article 34 of Decree-Law No. 41 of March 22 "Act on urgent measures for companies and workers due to COVID-19"[26]. In particular, subsection 3 establishes that the Italian Republic recognises Italian Sign Language and the role of the LIS interpreter. Furthermore, it mentions the promotion of subtitling services, without specifying whether the subtitles are produced in real time or for pre-recorded events.

Before 2021, there had been attempts to recognise Italian Sign Language and other services for people with hearing disabilities, such as Bill Proposals Nos. 4207[27] and

23 Legge-quadro per l'assistenza, l'integrazione sociale e i diritti delle persone handicappate, our translation (1992) <https://archivio.pubblica.istruzione.it/news/2006/allegati/legge104_92.pdf>

24 Disposizioni per favorire e semplificare l'accesso degli utenti e, in particolare, delle persone con disabilità agli strumenti informatici, our translation (2004) <https://www.senato.it/documenti/repository/eventi/dicembre2004/fscommand/Elenco%20leggi/004.pdf>

25 Bilancio di previsione dello Stato per l'anno finanziario 2019 e bilancio pluriennale per il triennio 2019-2021, our translation (2018) <https://www.miur.gov.it/documents/20182/3395588/Legge+30+dicembre+2018%2C+n.145.pdf/24c521b9-4655-3aca-2d54-dec0eee3d6fc?t=1592820944368>

26 Conversione in legge, con modificazioni, del decreto-legge 22 marzo 2021, n. 41, recante misure urgenti in materia di sostegno alle imprese e agli operatori economici, di lavoro, salute e servizi territoriali, connesse all'emergenza da COVID-19, our translation (2021) <https://www.gazzettaufficiale.it/eli/id/2021/05/21/21G00080/sg >

27 Proposta di legge n. 4207, our translation (2011) <https://leg16.camera.it/126?tab=&leg=16&idDocumento=4207&sede=&tipo=>

4679[28]. In both cases, however, although the Senate had approved the proposals, the legislative process was interrupted as the Chamber of Deputies failed to reach a conclusion on the matter before the end of the parliamentary term. As a result, neither text could proceed further in the legislative process. Bill Proposal No. 4679 is particularly interesting in the context of this paper because its article 6 explicitly mentions the use of stenotyping and respeaking as services intended to facilitate access for students with hearing barriers to university and post-university education. However, in the research conducted, this turns out to be the only document specifically addressing such services, and it never reached the status of law.

This disregard for the communicative dimension in Italian legislation contrasts with the principles outlined in both the United Nations Convention on the Rights of Persons with Disabilities (UNCRPD), ratified by Italy in 2009[29], and the European Accessibility Act (EAA)[30], transposed into Italian legislation in 2022. Indeed, these two documents emphasise the importance of considering accessibility in all its facets and forms, from the construction of ramps at building entrances to the provision of technological and communicative services. It is therefore evident that there is a need to develop future national regulations or update existing laws, making them more precise and suitable for accessibility in all its aspects.

### 2.3 Research and academic status

There are still several gaps in academic literature on STTI in Italy. Nevertheless, there are significant contributions that provide a solid starting point for further studies. Most notably, existing research primarily focuses on respeaking techniques or the application of speech-to-text services in the television and film sectors (cf. Eugeni 2006; 2008; Trapani 2012). An exception is the study by Pirelli (2006), which offers a brief insight into the use of respeaking and stenotyping in some Italian universities. An attempt to provide a more complete view of STTI and its advantages was made by Valente (2024) in her master's degree thesis. Her study provides an overview of the potential and applications of this service, as well as its implications across various sectors, thereby helping to bridge the current gaps in the Italian academic literature.

The focus on respeaking is also reflected in the educational curricula of Italian universities. At the Interpreting and Translation Department in Forlì[31], the master's degree programme in interpreting includes an elective course titled "Advanced Meth-

28 Proposta di legge n. 4679, our translation (2017) <https://documenti.camera.it/_dati/leg17/lavori/stampati/pdf/17PDL0054950.pdf>

29 UN (United Nations) (2006) "Optional Protocol to the Convention on the Rights of Persons with Disabilities", <https://treaties.un.org/pages/ViewDetails.aspx?src=TREATY&mtdsg_no=IV-15-a&chapter=4&clang=_en>

30 European Parliament / Council of the European Union (2019) "Directive (EU) 2019/882", <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32019L0882>

31 Interpreting and Translation Department, Forlì Campus (University of Bologna) (2023) "Metodi e tecnologie avanzate per l'interpretazione", https://www.unibo.it/it/studiare/dottorati-master-specializzazioni-e-altra-formazione/insegnamenti/insegnamento/2023/433465

ods and Technologies for Interpreting". This course develops skills in voice recognition, specifically utilising Dragon NaturallySpeaking software. Similarly, the University of International Studies of Rome (UNINT)[32] offers master's interpreting students the opportunity to take a course in both intralingual and interlingual (English-Italian) respeaking. Additionally, the University of Parma[33] offers a first-level master's programme in audiovisual translation that includes modules focused on respeaking techniques. Another noteworthy initiative at the university level is the "Introductory course on intralinguistic subtitling and subtitling for the deaf", a summer course organised in June 2024 by the Civic School Interpreters and Translators Altiero Spinelli[34], which also focused on the production of live subtitles.

The study of other STTI techniques is particularly encouraged only by the Scuola Superiore per Mediatori Linguistici (SSML) in Pisa. Indeed, this institution took part in the European project LiveTextAccess (LTA), which focused on developing a training programme that includes both respeaking and the use of the Velotype keyboard (Hagmann-Schlatterbeck/ZDF 2019).

With regard to STTI as a service offered to university students with communicative barriers, not much information is available. A brief review of the websites of Italian universities revealed that some of them, such as the University of Bologna[35], the University of International Studies of Rome, Rome Tor Vergata[36], the University of Padua[37], the University of Siena[38] and the University of Cagliari[39] provide real-time transcription services. Most of these universities only mention the availability of stenotyping services upon request without detailing the operational modalities. For instance, it remains unclear whether these services involve the presence of an interpreter or just rely on specialised software. In the specific case of the University of Bologna, the reference is solely to the use of Google Docs' voice typing tool, without

32 University of International Studies of Rome (UNINT) (2023) "Respeaking intralinguistico e interlinguistico EN-IT", <https://unint.coursecatalogue.cineca.it/insegnamenti/2023/1942-GRU%20A%20INTE/2018/9999/10025?coorte=2023>

33 University of Parma (2022) "Corso Master universitario di I livello in Traduzione audiovisiva", https://www.unipr.it/sites/default/files/2022-08/Traduzione%20audiovisiva.pdf

34 Civic School Interpreters and Translators Altiero Spinelli (2024) "Corso di introduzione alla sottotitolazione intralinguistica e alla sottotitolazione per sordi", <https://lingue.fondazionemilano.eu/corsi/archivio/summer-school-2024/corso-di-sottotitolazione-intralinguistica-e-alla-sottotitolazione-per-sordi-1>

35 University of Bologna (2023) "Speech to text gratuiti: lo strumento "Digitazione vocale" su Documenti Google (Google Docs)", <https://site.unibo.it/studenti-con-disabilita-e-dsa/it/per-studenti/presentazione-di-strumenti-automatici-per-la-sottotitolazione-e-o-dettatura-vocale/guida-strumento-speech-to-text-google-docs-rivisto1.pdf/@@download/file/Guida%20strumento%20Speech%20to%20text%20-%20Google%20Docs%20(rivisto.pdf>

36 University of Rome Tor Vergata (2024) "Servizio per la disabilità", <http://elettronica.uniroma2.it/index.php?page=caris>

37 University of Padua (2024) "Supporti per la frequenza delle lezioni, per lo studio individuale e per gli esami", https://www.unipd.it/supporti-studiare

38 University of Siena (2021) "Guida ai servizi", <https://www.unisi.it/sites/default/files/allegatiparagrafo/guida_ai_servizi_2021_20luglio_web_light_2.pdf>

39 University of Cagliari (2024) "Ausili disponibili presso il S.I.A.", <https://people.unica.it/disabilita/ausili-disponibili-presso-il-sia/>

clarification of whether an external person is involved in dictating the speech[40] or if the software operates autonomously. A notable example is the University of Cagliari, which has implemented the use of Dragon NaturallySpeaking software to provide a speech-to-text service for students with communicative barriers (cf. Valente 2024).

## 2.4 Associations

In Italy, there are few associations active in the field of STTI. However, there is one association that is particularly involved in this field. It is called onA.I.R.[41] and it stands out because of the great attention it devotes to STTI. Indeed, it is dedicated to organising projects and events aimed at ensuring accessibility for people who may face difficulties in a specific communicative act, such as people with pre- and post-lingual hearing barriers, immigrants, and even Erasmus students. Specifically, onA.I.R. promotes initiatives which aim at disseminating speech recognition, fast typing on QWERTY keyboards and stenotyping. Furthermore, this association forms the Italian delegation of a larger international organisation known as the International Federation for Information and Communication Processing, or more briefly, Intersteno. This international organisation brings together professionals who use specific fast writing methods, such as shorthand, stenotyping, and respeaking, to produce high-quality texts. The onA.I.R. association works closely with Intersteno to organise international competitions as well, including fast typing and audio transcription contests, which are held annually.

An interesting aspect is that the onA.I.R. website provides information on accessibility, subtitling, and respeaking. In this way, the association significantly contributes to the dissemination of knowledge in the sphere of STTI and promotes greater awareness of the need for adequate support for anyone who can benefit from real-time transcription of spoken language.

Alongside onA.I.R.'s efforts, since 2023, the Italian landscape has seen the creation of The Deaf Soul, a community extremely active on social media, whose objective is to promote awareness campaigns about the inclusion of people with hearing barriers in every area of daily life. In addition to disseminating informative material on accessibility, inclusion, and services for hearing barriers, The Deaf Soul organises events where people with hearing barriers meet hearing people, thus fostering dialogue between these two realities (The Deaf Soul 2024)[42].

The birth of such a community represents an excellent starting point for raising greater awareness about the needs of people with hearing disabilities. This increased awareness can, in turn, highlight the crucial importance of services like STTI, representing a significant step towards a more inclusive and accessible society.

40 This aspect is particularly important in the case of Google Docs, as its automatic speech recognition system requires the use of specific voice commands to insert punctuation marks.

41 OnA.I.R. (Associazione Italiana Respeaking) (2022) <https://www.respeakingonair.org/it/>

42 The Deaf Soul (2024) https://www.thedeafsoul.com

## 2.5 Final remarks

From this overview, it is evident that the situation of STTI in Italy has significant gaps. Currently, there is no legislation that specifically mentions this service and, despite some progress, there are still few active associations in the field. Moreover, although respeaking courses have been introduced at some Italian universities and institutions, training opportunities remain limited.

First and foremost, a legislative breakthrough is necessary, with the enactment of laws that recognise STTI as an essential service for facilitating accessible communication, similar to what has been recently done for Italian Sign Language. It is estimated that there are approximately seven million people with hearing loss in Italy and just over one million with severe hearing impairment (Censis 2019). Although the exact number of post-lingual hearing-impaired individuals is not known, it has previously been highlighted that introducing this service could benefit a much larger user base. For example, including also people with cognitive barriers, the Italian socio-economic research institute Censis estimates that in Italy there are approximately 48,000 people with Down's syndrome (Anffas 2017)[43]. Therefore, it can be inferred that the legislative recognition of this service would benefit a considerable segment of Italian society.

It is certainly positive that there are degree level courses dedicated to this field. However, it is important to note that only a few universities and institutions offer these courses, which provide only basic training and focus almost exclusively on respeaking techniques. STTI, on the other hand, encompasses a wide range of working techniques that deserve more in-depth analysis in Italy, such as the use of the conventional QWERTY keyboard. It would be beneficial to introduce courses dedicated to STTI at more universities and institutions, not only for interpreters but translators as well, in order to broaden the training offer. This would be very useful, as such a service may represent a promising career path, especially for those studying interpretation. As previously mentioned, STTI shares several similarities with simultaneous interpreting. Indeed, interpreters, compared to other professional profiles, demonstrate a greater ability to deal with a communicative situation where they must divide their attention between several concurrent tasks and are exposed to a high cognitive load (cf. Szarkowska *et al*. 2018).

For these reasons, the skills acquired in simultaneous interpreting could also be used to train students in the field of STTI. This approach would provide new career opportunities while simultaneously expanding the range of services available for people with hearing barriers and beyond.

43 Anffas (2017) – Associazione Nazionale di Famiglie e Persone con disabilità intellettive e disturbi del neurosviluppo (in English, National Association of Families and Persons with Intellectual Disabilities and Neurodevelopmental Disorders) "Disabilità intellettive, alcuni dati", https://www.anffas.net/it/disabilita-intellettive-e-disturbi-dello-spettro-autistico/alcuni-dati/

## 3. STTI in Austria

### 3.1 Terminology

STTI in Austria and in the German-speaking countries is nowadays largely known as *Schriftdolmetschen* (written interpreting) (see 2.6), after first appearing as *Schreibdolmetschen* (writing interpreting, cf. Noffzt 2014; Norberg *et al*. 2015). In specific contexts, as in settings more connected to audiovisual translation, it is also referred to as *Live-Untertitelung* (live subtitling).

As an interpreting service backed by public funding and, therefore, paid for by the authorities for users with hearing impairment alongside SLI (cf. Grbić 2023), further definitions of the activity itself exist. It is defined as simultaneous presentation in written form of verbal utterances intended for users with hearing impairment to allow them active participation in communication. Furthermore, within existing regulations, it is laid down that specifically trained interpreters must write a text in real-time and that the users should read it live. The interpreter may adapt the text according to the user's need and provide a verbatim, condensed or simplified text. In general, the existing definitions also specify that the service is limited to intralingual STTI activities in German – which is the case for publicly funded activities within a working context (see 3.2), as well as in foreign languages, as is the case for educational settings.

### 3.2 Regulatory basis

The United Nations Convention on the Rights of Persons with Disabilities (UNCRPD) came into force in Austria in 2008. The European Accessibility Act (EAA) dating from 2019 will be transposed into Austrian legislation in 2025 through the "National Act on Accessibility"[44]. Additionally, since 2016 the "National Disability Act"[45], has been legally effective. These regulatory bases aim to foster users' rights to their respective means of communication. Furthermore, they highlight the importance of trained professionals in order to provide requested services and to ensure users' rights to inclusion in the labour market, in education, in jurisdiction and in society in general.

While these international and national regulatory standards do often explicitly mention SLI, they only implicitly highlight STTI as an adequate form of communication for specific users, such as subtitles. On a national level, nevertheless, in Austria there are more specific regulatory bases of various kinds in which the provision of STTI services is defined in a more detailed way when it comes to its use by users with hearing barriers in terms of sensorial impairment. First of all, there is the "Directive on Work and Education for an Accessible Working Environment"[46] of the Ministry for Social Affairs dating from 2022 as it is the Ministry which has dealt with the provision of SLI

44 Barrierefreiheitsgesetz, our translation (2023) <fname_1560223.pdf (parlament.gv.at)>

45 BGStG – Bundesgesetz über die Gleichstellung von Menschen mit Behinderungen, our translation (2005) <https://www.ris.bka.gv.at/eli/bgbl/I/2005/82>

46 BMSGPK (2022) - Bundesministerium Soziales, Gesundheit, Pflege und Konsumentenschutz "Richtlinie Arbeit und Ausbildung für eine barrierefreie Arbeitswelt", our translation, <https://www.sozialministeriumservice.at/Downloads/Richtlinie_Arbeit_und_Ausbildung.pdf>

and STTI alike: Austrian Sign Language was recognised as official and independent language in 2005, while STTI has been actively promoted since 2011. It is, therefore, the Ministry which bears the most experience with STTI and collaborates with the Austrian STT interpreters' association (see 3.4). This is why its directive and regulations specifically on STTI, to ensure inclusion into the labour market, define STTI as intralingual live text provision for users with a hearing barrier[47]. Within this regulation, it is indicated that for the provision of STTI services, professionals have to possess specific training certificates. STTI services may be offered on-site or remotely and are offered in teams of two if a certain assignment is particularly demanding. Last but not least, the charges for STTI activity, travelling time and post-processing of interpreted live texts, are fixed within these regulations. It is also important to highlight that, within the social services regulations, users have the right to opt for a preferred activity based on their individual preferences. Therefore, in some cases they may use STTI, in others SLI.

When it comes to inclusion in education, in secondary school settings a recent regulation, developed in collaboration with the Austrian STT interpreters' association in 2023, came into force at the beginning of the winter term 2023. It is the first regulation for users with hearing barriers in schools where STTI and SLI services are nominated alongside other forms of assistance[48]. As a result of this regulation, users with hearing barriers can present a request to the school's coordinators, who pass on the request to the regional coordinators within the Ministry of Education. For STT interpreters, this regulation establishes a charge per teaching unit covering STTI activity and travelling time, as well as postprocessing for the rough correction of the written live text. Regarding STTI in higher education, the provisions of the "University Act"[49] refer to the UNCRPD as well as the National Disability Act.

STTI services as a publicly funded activity are also covered by the "Federal Fee Act"[50], enacted in 2023. Although STTI is not specifically mentioned, this law provides a basis for remuneration for authorities under the Ministry of Justice and the Ministry of Internal Affairs, as interpreters for court procedures are assigned and paid according to this regulation. Certified court interpreters in Austria are the only professionals bearing a protected title after passing a certification exam held by the Ministry of Justice and the Austrian Association of Sworn and Court-certified Interpreters. To date, the examination scheme covers spoken language interpreting and SLI, but does not yet include STTI.

In general, STTI in Austria, along with translation and interpreting, is a free-lance profession bearing no professional title. However, for the provision of STTI activity for individual users entitled to public funding in the working environment, training

47 BMSGPK (2023) – Bundesministerium Soziales, Gesundheit, Pflege und Konsumentenschutz "Schriftdolmetschung", our translation, <https://www.sozialministeriumservice.at/Downloads/regelungen_-schriftdolmetsch-leistungen.docx>

48 BMBWF (2023) – Bundesministerium Bildung, Wissenschaft und Forschung (in English, Federal Ministry of Education, Science and Research) "Erlass betreffend Unterstützung für Schülerinnen und Schüler mit einer Behinderung in Bildungseinrichtungen des Bundes", our translation, <https://www.bmbwf.gv.at/dam/jcr:28231582-f900-412c-90af->

49 Universitätsgesetz, our translation (2021) <https://www.ris.bka.gv.at/GeltendeFassung.wxe?Abfrage=Bundesnormen&Gesetzesnummer=20002128&FassungVom=2021-09-30>

50 Gebührenanspruchsgesetz, our translation (2023) <https://ris.bka.gv.at/GeltendeFassung.wxe?Abfrage=Bundesnormen&Gesetzesnummer=10002337&FassungVom=2023-02-21>

certificates are necessary. STTI freelancers are required to register their professional activity according to the Austrian Trade, Commerce and Industry Regulation Act as a "language service" at the Chamber of Commerce, under the Ministry of Commerce. Since 2015, this registration has included STTI services (cf. Platter 2019).

Regarding the regulatory basis, it should be mentioned that in Austria, alongside the aforementioned national laws, federal laws provide further frameworks for STTI activities. These laws differ in their revised versions and conciseness, some just providing charges for STTI activities in general, others even underlining the differences between STTI and other forms of accessibility services. This is the case, for example, with the "Tyrolean Act on Participation"[51] dating from 2017, where STTI is defined alongside with SLI, relais interpreting and tactile interpreting:

> 2. STTI: STTI involves interpreting from the spoken source language (German) into the target language (written German) in order to ensure communication between people with hearing impairments and people without hearing impairments. (Tiroler Teilhabegesetz 2017, Tyrolean Act on Participation; our translation)

### 3.3 Research and academic status

In Austria, as part of the German-speaking STTI world, research and scientific literature on STTI has been gaining impetus in recent years. Whilst Norberg *et al*. (2015) still highlighted the fact that, due to its intralingual function and the fact that SLI is much more requested as an interpretation service for users with a hearing barrier, interest in research was limited. Nowadays, the academic point of view on the discipline is differentiated, both in terms of quantity of publications as well as research topics covered. This, in fact, is due to a process of professionalisation of STTI in Austria, which gained impetus as academia recognised the field as research topic and started teaching and researching it. Furthermore, practitioners started acting as researchers.

These developments do, in fact, show particular similarities to the professionalisation of SLI (cf. Roy/Napier 2015: 123)

When it comes to teaching STTI at an academic level, since 2018, students at the University of Graz have been able to enrol in an MA course dedicated to it. In this course, students get to know about speech recognition and conventional keyboard techniques, as well as strategies for co- and post-editing of live texts. At the University of Vienna, STTI is included in a special course within the BA programme dedicated to accessibility, treating STTI alongside easy and plain language as well as audiodescription. A new course exclusively teaching STTI has been included in the new MA programme at the University of Vienna since the winter term of 2022/2023. In this course format, students, besides speech recognition and conventional keyboard techniques, get to know strategies for STTI with a strong practical focus. Furthermore, they get insights into the ethical principles of STTI professionals and acquire knowledge about the regulatory basis of STTI services in German-speaking countries as well as STTI in their working languages.

51 Tiroler Teilhabegesetz, our translation (2017) <https://ris.bka.gv.at/GeltendeFassung.wxe?Abfrage=LrT&Gesetzesnummer=20000709&FassungVom=2022-04-26>

At the postgraduate level, the Centre for Translation Studies, the Austrian STT Interpreters' Association and the Postgraduate Centre of the University of Vienna held three academic training and certification courses for STT interpreters from 2019 until 2022. These academic courses prove the will to foster the aspects of interpretation within STTI activity and, at the same time, they try to combine STTI training with foreign language skills which prove to be suitable for practitioners in order to also offer their services in languages other than German. Besides this, trained interpreters are also able to offer interlingual STTI, as they have gained both interpretation and interlingual skills.

With regards to research activities, the interest of students of translation and interpreting at the Universities of Graz and Vienna in STTI combines with fostered teaching activity and courses in accessibility, and existing courses in SLI at the University of Graz. In recent years, MA theses dedicated to STTI have treated the topic from a technical point of view, analysing inter- and intralingual respeaking and comparing it to fully automated technologies (Hattinger 2013; Dobrovac 2018; Jorda 2021; Sakalidis 2021; Pechhacker 2022). Other theses have focused on STTI's workflows and quality (Neuper 2012; Kuzera 2021), and another via means of eye-tracking (Matzenberger 2023). At PhD level, STTI was analysed in view of its professionalisation by Platter (2015); a further PhD thesis is to focus on the quality of STTI methods (Eichmeyer 2017 and forthcoming).

Publications on STTI in Austria were primarily developed from an international point of view, as Norberg, Stachl-Peier and Tiittula dedicated them to quality and interpretation strategies (Norberg *et al*. 2015; Norberg/Stachl-Peier 2017). Later publications focused on the development of STTI in Austria, comparing it amongst other things to other forms of accessibility (Platter 2019; 2021; 2022). As a major event dedicated to the exchange of science with the public, the Austrian STT Interpreter's Association organised the 6th European Conference of STT interpreters, held in August 2022 under the title "PostPandemicPerspectives"[52].

Research within research groups focusing on STTI is actively being conducted by the Vienna Interpreting Research Group, where a study on collaboration between spoken-language interpreters and STT interpreters was presented at the TIC conference in Bratislava 2023 – a respective publication is forthcoming[53].

With regard to STTI as a service that students with hearing barriers can use within academic training institutions, students attending courses of different kinds in Vienna, the region of Lower Austria and in Styria, can benefit from the services offered by dedicated service points for deaf and hard-of-hearing students[54]. The agency in Vienna is based at the Technical University of Vienna, offering its services for STTI and SLI since 2011. The agency covering the region of Styria started its activity in

52 ÖSDV (2022) – Österreichischer Verband für Schriftdolmetschen – (in English, Austrian Speech-to-Text Interpreters'Association) "ECOS Vienna 2022", <https://www.oesdv.at/ecos-vienna-2022>

53 VIRG (2024) – Vienna Interpreting Research Group <https://viennainterpreting.univie.ac.at/news-events/>

54 GESTU Vienna (2024) – Gehörlos und Schwerhörig Erfolgreich Studieren – (in English, Service Point for Deaf and Hard-of-Hearing Students) <https://www.tuwien.at/studium/studieren-an-der-tuw/gestu/gestu-geschichte4>

2022[55]. Both agencies organise and coordinate interpreters for assignments in which the individual students need STTI at university, this being primarily for courses and conferences. The students are accompanied by the STT interpreters, whether this be with both of them on-site or one of the interpreters working remotely. Furthermore, they get roughly corrected live text versions after the assignment, in order to use the texts for future reference. STT interpreters in specific cases also provide transcriptions for videos and audios that teachers at universities can give their students. In order to obtain STTI assignments, STT interpreters have to be included in the agencies' interpreter pool, therefore, they have to present a training certificate in translation and interpreting or a specific STTI training certificate. Amongst the booking and coordinating of the STTI assignments, the agencies consult students on the various possibilities for assistance and ensure quality control by regularly asking users for feedback as well as holding annual meetings with STT interpreters.

### 3.4 Associations

In Austria, there are stakeholders and associations promoting STTI. Possible users of STTI services are represented by the Association for the Deaf[56] and various associations for the hard-of-hearing at federal level and at national level as well. The Austrian Association of Hard-of-Hearing Associations[57], alongside its German and Slovakian counterparts, introduced the first training courses for STT interpreters held in Vienna in the years 2009-2011 (cf. Platter 2015). Since then, it has held three other courses for training and certifying STT interpreters and manages a STTI agency where users and clients can book STT interpreters[58] for assignments.

On its pathway to professionalisation, STTI in Austria reached a considerable milestone with the founding of the Austrian Speech-to-Text Interpreters' Association *Österreichischer Verband für Schriftdolmetschen - ÖSDV*[59]. After the completion of a further continuing training and certification course for STT interpreters, organised by the BFI Tirol in collaboration with the Tyrolean Association for the Hard-of-Hearing v-OHR-Laut, the certified participants founded the association in January 2019 (cf. Platter 2022). As of summer 2024, it has a membership of 50 residing in Austria and Germany, the majority of them listed as certified STT interpreters on the Association's

55 GESTU Graz (2024) – Gehörlos und Schwerhörig Erfolgreich Studieren – (in English, Service Point for Deaf and Hard-of-Hearing Students) <https://www.tugraz.at/en/studying-and-teaching/studying-at-tu-graz/prospective-students/gestu-graz-service-point-for-deaf-and-hard-of-hearing-students/>

56 ÖGLB (2024) – Österreichischer Gehörlosenbund <https://www.oeglb.at>

57 ÖSB (2024) – Österreichischer Schwerhörigenbund-Dachverband <https://www.oesb-dachverband.at>

58 Calculating (approximately) and considering the fact that in each course around 10 interpreters were trained (which corresponds to about 40 trained STT interpreters), the number of active STT interpreters within this organisation is very low and not sustainable. In 2015, there were 12 STT interpreters listed within this agency (Platter 2015), and in 2024 it is just 3 interpreters, working primarily with the conventional keyboard (ÖSB 2024).

59 ÖSDV (2024) – Österreichischer Verband für Schriftdolmetschen <https://www.oesdv.at>

website. This list contains the first public list of qualified professionals and is made available to the regulators and authorities as a useful tool to search for STT interpreters by languages, STTI techniques used (either speech recognition or conventional QWERTY keyboard), areas of activity and specialisations. In order to be included in this list, the STT interpreters have to be members of the Association. Admission requires a peer-to-peer admission scheme and furthermore, they have to prove continuous professional activity amounting to 30 hours of STTI and continued training amounting to at least two courses per year within a period of three years[60].

For continued training of its members, the Association established a mentoring scheme: within its offers, there is a regular on-site conference held at the United Nations Office in Vienna, where mentorees within a simulated conference setting can provide interlingual and intralingual STTI in English and German. In addition to special workshops dedicated to requested topics which are held on-site or online, the Association regularly holds informal exchange formats (potpourri), where members can share experiences and insights into specific fields of STTI. For its members, the Association provides an international list of scientific, academic and popular science publications on STTI which is constantly updated.

It was the Association that established a code of conduct, which in its most updated version is made available to the public and is used by authorities and agencies alike as part of their quality assurance activities[61]. Within this code of conduct, STTI as an activity is defined as interpretation. The professional behaviour of STT interpreters prior to, during and after assignments is described as well as within the STTI community, in order to prevent unfair competition. The code of ethics aims to highlight the quality of STTI as a holistic approach and to foster role awareness and commitment to professional behaviour in general. The Association itself is actively involved in the majority of training and research activities for and on STTI in Austria (see 3.2), and it is in constant contact and exchange with stakeholders at national and international levels in order to establish professional working conditions for STT interpreters. Together with other associations in interpreting and translation, it is part of the Austrian collaboration network *Translationsplattform*[62] formed by the six associations within the T&I sector in Austria. It is part of the standardisation committees working in T&I and since 2022 has been a member of the International Federation of Translators (FIT)[63]. As such, it is actively involved in the Ethics Working Group on AI and the Standing Committee on Audiovisual Translation. Besides working together with its German counterpart BSD, the Association is also involved in the ECOS conference organisation.

As a major focus within its activities, the Association is fostering public relations in order to share knowledge and insights into STTI for users, authorities and the public alike. The Association's members are encouraged to include their personal website on the list of certified STT interpreters and to use it and their Social Media accounts

60 ÖSDV (2024) – Österreichischer Verband für Schriftdolmetschen "Liste der zertifizierten Schriftdolmetscher:innen", <https://www.oesdv.at/schriftdolmetscherinnen-uebersicht>

61 ÖSDV (2020) – Österreichischer Verband für Schriftdolmetschen "Berufs- und Ehrenordnung", <aff72d_d39dfe72fd0a458bb67c21ca2747bb47.pdf>

62 Translationsplattform (2024) <http://www.translationsplattform.at/>

63 FIT (2024) – International Federation of Translators <https://en.fit-ift.org/>

by linking them to the Association's website. In recent months, the Association itself, besides including news and information on its website, has concentrated its efforts on raising awareness about STTI services in secondary schools and amongst event organisers. Therefore, a mailing campaign later followed by an online 1-hour presentation was launched.

## 3.5 Final remarks

Analysing the professional status of STTI in Austria, it can be stated that STTI is professionalised in various aspects (Tseng 1992; Ju 2009; Platter 2015; Platter 2022). In terms of stakeholders, there are trained professionals, the number of which has increased in recent years since the introduction of the first STTI courses at continuing education institutions in 2011 in Vienna. The offer was widened by a second institution in the western part of Austria including courses on both speech recognition and conventional keyboard techniques, as well as research and translation approaches, which has been particularly fruitful. Furthermore, offering STTI training in various forms at institutions for translation and interpreting at the University of Vienna and the University of Graz, has helped considerably in increasing the interest of trained interpreters and translators in foreign language skills in this particular activity. At present, a shortcoming is that no postgraduate training and certification course is being offered which would be of primary interest to translation and interpreting professionals.

The establishment of the professional association for STTI on a national level has been a further milestone on the pathway to professionalisation. This has enabled interaction with individual users, user groups, clients, training institutions, agencies, language service providers and the market in general to become more efficient. For the professional association, this achievement makes it even more effective when approaching regulators and legal authorities.

Regarding the regulatory basis, several laws at a federal and national level exist, covering users' rights to SLI, as well as to choose STTI as an equivalent interpretation service if the users have a hearing impairment above 70 percent. Nevertheless, a major shortcoming is the fact that within the National Act on Accessibility, STTI is not mentioned at all. Furthermore, many authorities managing requests for STTI services are dealing with them for the first time. That is why major public relations activity from all stakeholders is necessary, a challenge the Austrian Speech-to-Text Interpreters' Association is facing on many levels.

Existing regulations for users with hearing impairment do include intralingual STTI for individual users and the respective rates paid by public institutions, within education, social inclusion jurisdiction and execution, and in addition, many STT interpreters are trained in translation and interpretation. Nonetheless, major consensus and commitment are still needed to further focus on role awareness of STT practitioners as highly professional service providers. This should help users with hearing and other barriers to benefit from more professionalised service providers in terms of quality and quantity.

Regarding primary and secondary user groups, the latest available data for Austria suggest that 6.9 per cent of young people in secondary education up to the age of 15

have a hearing impairment in a quiet ambiance, even if they use an acoustic hearing aid and 14.5 percent have problems with memory or concentration. The respective rates for men and women of working age from 15 until 59 years of age are around 2 per cent for hearing impairment and around 8.5 per cent for concentration difficulties, reaching a peak in the ages between 45 and 59 with 5 and 12.3 per cent respectively. For people above 75, the percentages are 29 and 42 percent, proving the fact that in an ageing society, STTI services are becoming more and more crucial for accessibility (Statistik Austria 2020).

## 4. Discussion and conclusion

The aim of this article was to conduct a comparative analysis of the diffusion of STTI in Italy and Austria, in order to highlight the gaps that characterise the Italian context and to indicate a possible model to address them. Additionally, the aim was to provide an overview of the potential, versatility, and functionality of STTI, to emphasise how this service could significantly improve the quality of life for a substantial segment of Italian society.

As revealed by the comparative analysis, STTI is still relatively unknown in Italy. It was found that there are no specific laws recognising this practice. Furthermore, there are few active associations in the field and, at the university level, the educational offers are largely limited to the teaching of respeaking. The overview of the situation in Italy turned out to be markedly different from that in Austria. Indeed, in this country there is a rich research activity in the field and established laws that mandate the provision of STTI services upon request. Moreover, there are numerous associations dedicated to promoting STTI, as well as many universities and training centres offering comprehensive courses on the service.

To achieve an improvement in STTI provision in Italy, it would be essential to first address communication accessibility. For example, as previously mentioned, a turnaround at the legislative level would be needed, with the enactment of precise laws that recognise the importance of ensuring inclusive communication services and highlight the challenges associated with this aspect of daily life, which often remain invisible.

Once greater awareness of the importance of accessible communication is established, it may be easier to promote individual services. In the specific case of STTI, first of all, a terminological evolution would be necessary. The inadequacy of the currently used term *sottotitolazione in tempo reale* and the need for the introduction of the broader concept of *interpretazione scritta* have been emphasised (see 2.1). This is crucial, as promoting correct terminology is the basis for a better understanding of the nature of the service. Furthermore, considering the framework outlined for Austria, other factors could favour the further future development and establishment of STTI in the Italian context. These could include increased activity by associations of people with hearing barriers to advocate for the service, legislative regulation of the interpreting profession, publication of specific academic research, integration into university curricula and the establishment of an association representing professionals in the field.

This article also aims at providing a contribution that could serve as a basis for further research on STTI in Italy. There are, indeed, several research avenues that could

be explored in the future to enrich the landscape of studies on STTI in Italy. First, it would be interesting to analyse the perception of STTI among its users. Given that this would be a first study of its kind, one could start by referring to a representative sample of the main users of STTI, namely people with post-lingual hearing barriers. Specifically, questionnaires could be developed and administered to these users to assess their knowledge of the topic and gather their opinions on aspects such as the advantages and disadvantages of the service. Such a study would provide valuable insights to tailor STTI to the needs and expectations of its beneficiaries.

Another area for further research in the Italian context could focus on the professional competences required to work in the field of STTI. According to empirical research results, several skills may be necessary, including technical, relational, and behavioural skills (cf. Valente 2024). More in-depth research could expand on this framework, provide a more comprehensive view and help develop adequate training pathways.

Lastly, it would be interesting to further investigate the use of STTI as a service offered to students in Italy. A brief review of Italian university websites revealed that some of them do provide STTI services for students with hearing barriers (see 2.3). An attempt to analyse the diffusion of STTI in Italian universities was made by Pirelli (2006), specifically examining the cases of the University of Padua and the University La Sapienza in Rome. However, this study dates back 18 years and focuses solely on the university environment. For this reason, it would be worthwhile to explore the current situation at the university level and potentially extend the investigation to the scholastic dimension as well.

In conclusion, the scope for action to improve the status of STTI in Italy is certainly wide and there is still much work to be done. It is hoped that this contribution may mark the beginning of a significant shift, encouraging greater attention and concrete commitment towards the implementation and dissemination of this service in Italy, with the aim of ensuring more inclusive and equitable access to communication for all.

## References


Aristotle (1999) *Politics*, translation by Benjamin Jowett, Kitchener, Batoche Books.

Censis (Centre for Social Investment Studies) (2019) "Sentirsi bene – Il valore sociale dell'audioprotesi", https://www.censis.it/sites/default/files/downloads/Studio%20Censis_web.pdf.

Dobrovac S. (2018) *Respeaking in Österreich: neue Berufsmöglichkeit für Dolmetschstudierende?*, unpublished MA Thesis, ITAT Graz, Universität Graz, <https://unipub.uni-graz.at/obvugrhs/content/titleinfo/2945862>.

Eichmeyer D. (2017) "Speech-to-Text Interpreting: Barrier-Free Access to Universities for the Hearing Impaired", in S. J. Jekat / G. Massey (eds) *Barrier-free Communication: Methods and Products. Proceedings of the 1st Swiss Conference on Barrier-Free Communication*, Winterthur, ZHAW digitalcollection, 6-15.

Eichmeyer D. (forthcoming) *WIRA: Model for qualitative assessment of a speech to-text interpreting service, taking into account the user perspective*. Forthcoming PhD Thesis, Universität Wien.

Eichmeyer D. / Kaul M. / Nofftz B. (2017) “Schriftdolmetschen – Handhabe von Mitschriften”, *CoMe III (1)*, 81-103.
Eugeni C. (2006) “Introduzione al rispeakeraggio televisivo”, Special Issue of *inTRAlinea*, <https://www.intralinea.org/specials/article/1683>.
Eugeni C. (2008) “Respeaking the TV for the Deaf: For a Real Special Needs-oriented Subtitling”, *Studies in English Language and Literature* 21, 37–47, <https://ltaproject.eu/wp-content/uploads/2019/06/Eugeni_C.-2008_Respeaking-the-TV-for-the-Deaf-For-a-Real-Special-Needs-Oriented-Subtitling.pdf>.
Gerzymisch-Arbogast H. (2013) „Gutachten zu der Bezeichnung “Schriftdolmetschen“, University of Saarlandes, <https://bsd-ev.org/wp-content/uploads/2019/09/gutachten_gerzymisch.pdf>.
Grbić N. (2023) *Gebärdensprachdolmetschen als Beruf: Professionalisierung als Grenzziehungsarbeit. Eine historische Fallstudie in Österreich*, Bielefeld, transcript Verlag.
Hagmann-Schlatterbeck M. / ZDF Digital Medienproduktion GmbH (2019) “Live-TextAccess“, <https://ltaproject.eu/wp-content/uploads/2019/06/LTA_SDI_Event_The-LTA_Project_28.03.2019.pdf>.
Hattinger C. (2013) *GESTU. Erarbeitung eines Respeaking Trainingsprogramms und Vergleich zur automatischen Spracherkennung, um diese Techniken zur Untertitelerzeugung in E-Learning Plattformen zu verwenden und somit speziell hörbeeinträchtigte Studierende zu unterstützen*, unpublished MA Thesis, Technische Universität Wien, <https://repositum.tuwien.at/handle/20.500.12708/2494>.
Jorda B. (2021) *Intralinguales Respeaking für die Live-Untertitelung: eine Fallstudie zu Problemauslösern und Strategien*, unpublished MA Thesis, Universität Wien.
Ju E.M. (2009) “The professionalization of interpreting in Taiwan: A critical review of Tseng’s model”, *Compilation and Translation Review* 2/2, 105-125.
Kuczera L. (2021) *Qualität beim Schriftdolmetschen mittels Spracherkennung: Modelle zur Qualitätsbewertung für das interlinguale Respeaking und das Simultandolmetschen im Vergleich*, unpublished MA Thesis, Universität Graz, <https://unipub.uni-graz.at/obvugrhs/content/titleinfo/5749722>.
Maaß C. (ed.) (2020) *Easy Language-Plain Language-Easy Language Plus: Balancing Comprehensibility and Acceptability*, Berlin, Frank & Timme.
Matzenberger J. (2023) *Visuelle Aufmerksamkeit beim Schriftdolmetschen: eine Eyetracking-Studie*, unpublishd MA Thesis, Universität Wien, <https://utheses.univie.ac.at/detail/69464/>.
Neuper C. (2012) *Untertitelung und Informationsverlust: eine Analyse der inter- und intralingualen Übertragung des Films der Knochenmann*, unpublished MA Thesis, Universität Graz, <https://unipub.uni-graz.at/obvugrhs/content/titleinfo/225062>.
Nofftz B. (2014) “Written interpreting in individual countries”, <https://efhoh.org/wp-content/uploads/2017/04/Birgit-Nofftz-Germany-2014-Speech-to-Text.pdf>.
Norberg U. / Stachl-Peier U. (2017) “Discourse prosody and real-time text interpreting: Making live speech visible”, in: A. Łyda / K. Holewik (eds) *Interdis-*

*ciplinary encounters: Dimensions of interpreting studies*, Katowice, Wydawnictwo Uniwersytetu Śląskiego, 161–188.
Norberg U. / Stachl-Peier U. / Tiittula L. (2015) "Speech-to-text interpreting in Finland, Sweden and Austria", *Translation & Interpreting* 7/3, 36–49.
Pechhacker A. (2022) *Live-Untertitelung mittels Respeaking: eine Workflow-Analyse*, unpublished MA Thesis, Universität Wien, < https://utheses.univie.ac.at/detail/63622/#>.
Perego E. (2020) *Accessible Communication: A Cross-Country Journey*, Berlin, Frank & Timme.
Pirelli G. (2006) "Le necessità dei sordi: la sottotitolazione in tempo reale all'università", Special Issue of *inTRAlinea*, <https://www.intralinea.org/specials/article/Le_necessita_dei_sordi_la_sottotitolazione_in_tempo_reale_all-universita>.
Platter J. (2015) *Translation im Spannungsbereich von Mündlichkeit und Schriftlichkeit. Schriftdolmetschen in Österreich – eine textbasierte Analyse*, unpublished PhD Thesis, Universität Wien.
Platter J. (2019) „Schriftdolmetschen", in M. Kadric (ed.) *Besondere Berufsfelder für Dolmetscher*innen*, Wien, Facultas, 117–149.
Platter J. (2021) „Barrierefreie Kommunikation – Aspekte der Professionalisierung im DACH-Raum", in S. Pöllabauer / M. Kadric (eds) *Entwicklungslinien des Dolmetschens im soziokulturellen Kontext: Translationskultur(en) im DACH-Raum*, Tübingen, Narr Francke Attempto, 77–109.
Platter J. (2022) „Schriftdolmetschen in Österreich: Einblicke in den und Erkenntnisse aus dem laufenden Professionalisierungsprozess", in G. Hebenstreit / P. Hofeneder (eds) *Translation im Wandel: Gesellschaftliche, konzeptuelle und didaktische Perspektiven*, Berlin, Frank & Timme, 69-92.
Rink I. (2020) "Kommunikationsbarrieren", in C. Maaß / I. Rink (eds) *Handbuch Barrierefreie Kommunikation*, Berlin, Frank & Timme, 29–66.
Romero-Fresco P. (2020) *Subtitling Through Speech Recognition: Respeaking*, United Kingdom, Taylor & Francis.
Roy C.B. / Napier J. (2015) *The sign Language Interpreting studies reader*, Amsterdam/Philadelphia, John Benjamins Publishing Company.
Sakalidis T. (2021) *Interlinguale Live-Untertitel: Direkt vs. Relaisgestützt*, unpublished MA Thesis, Universität Wien, < https://utheses.univie.ac.at/detail/58275/>.
Schubert K. (2016) "Barriereabbau durch optimierte Kommunikationsmittel: Versuch einer Systematisierung", in N. Mälzer (ed.) *Barrierefreie Kommunikation Perspektiven aus Theorie und Praxis*, Berlin, Frank & Timme, 15–33.
Statistik Austria (2020) "Gesundheitliche Beeinträchtigungen", <https://www.statistik.at/statistiken/bevoelkerung-und-soziales/gesundheit/gesundheitszustand/gesundheitliche-beeintraechtigungen>.
Stinson M. / Elliot L. / Kelly R. (2017) "Deaf and Hard-of-Hearing High School and College Students' Perceptions of Speech-To-Text and Interpreting/Note Taking Services and Motivation", *Journal of Developmental and Physical Disabilities* 29/1, 131–152.
Szarkowska A. / Krejtz K. / Dutka L. / Pilipczuk O. (2018) "Are interpreters better respeakers?", *The Interpreter and Translator Trainer*, 2-20.

Tiittula L. (2006) “Schriftdolmetschen-Mündlichkeit im schriftlichen Gewand”, *A Man of Measure. Festschrift in Honour of Fred Karlsson on his 60th Birthday – a special supplement to SKY Journal of Linguistics* 19, 481–488.
Tiittula L. (2018) “Schriftdolmetschen: Gesprochenes in schriftlicher Form”, in A. Deppermann / S. Reineke (eds) *Sprache im kommunikativen, interaktiven und kulturellen Kontext*, Berlin/Boston, De Gruyter, 193–218.
Tirinnanzi S. (2016) “Qualität beim Schriftdolmetschen”, in U. Gross-Dinter (ed.) *Dolmetschen 3.0 – Einblicke in einen Beruf im Wandel*, Berlin, Frank & Timme, 135–183.
Trapani T. (2012) *La sottotitolazione intralinguistica in tempo reale per sordi. Analisi comparativa tra Italia e Francia*, unpublished MA Thesis, University of Macerata.
Tseng J. (1992) *Interpreting as an emerging profession in Taiwan — A sociological model*, unpublished MA Thesis, Fu Jen Catholic University Taipei, Taiwan.
UNCRPD (2006) “Convention on the Rights of Persons with Disabilities and Optional Protocol”, <https://www.un.org/disabilities/documents/convention/convoptprot-e.pdf>.
Valente A. (2024) *Schriftdolmetschen/Interpretazione scritta: una nuova forma di accessibilità nell’interpretazione*, unpublished MA Thesis, SSLMIT, University of Trieste.
Wagner S. (2005) “Intralingual speech-to-text conversion in real-time: Challenges and Opportunities”, in H. Geryzmisch-Arbogast / S. Nauert (eds) *MuTra 2005 – Challenges of Multidimensional Translation*, Saarbrücken, 2–6 Maz 2005: Conference Proceedings, 210–219.
Witzel J. (2020) “Ausprägungen und Dolmetschstrategien beim Schriftdolmetschen”, in C. Maaß / I. Rink (eds) *Handbuch Barrierefreie Kommunikation*, Berlin, Frank & Timme, 303–325.

**The Interpreters' Newsletter Special Issue 2024 (29 bis), 59-78**
DOI: 10.13137/2421-714X/36765 ISSN 1591-4127 (print) ISSN 2421-714X (online)
https://www.openstarts.units.it/dspace/handle/10077/2119

# Automatic Speech Recognition in conference interpreting: an exploratory study on consecutive interpreting assisted by Sight-Terp

Michele Restuccia
University of Trieste

Abstract

*This article aims to observe potential interpreter-machine approaches in Consecutive Interpreting (CI) assisted by Automatic Speech Recognition (ASR). Sight-Terp was used in the study, a CAI tool for CI equipped with ASR and Machine Translation (MT). Six student-interpreters at the end of their second year of the master's degree in Conference Interpreting at SSLMIT, Trieste were asked to interpret one speech in the traditional consecutive manner and one using Sight-Terp in a randomised order. The results show that a) the transcript produced by Sight-Terp does not always enable the interpreters to concentrate more on speech listening and comprehension, b) the duration of the interpreters' renditions tends to increase, owing to an increased dependency on the written text and to additional cognitive efforts involved in the rendition phase, and c) ASR can be a useful resource for the retrieval of names and figures (although there may be errors in the transcript), but it presents limitations in the transcription of spontaneous spoken language. In conclusion, despite the significant advances in ASR technologies, more research is needed to further explore the advantages of ASR-assisted consecutive interpreting and strategic approaches.*

Keywords

Assisted consecutive interpreting, ASR, CAI tool, interpreting technologies, Sight-Terp.

## Introduction

Over the past century, since the emergence of modern interpreting, technology has consistently played a crucial role in conference interpreting (Kalina/Ziegler 2015). This is particularly prevalent in simultaneous interpreting, though possibly less so in consecutive interpreting. However, it has often taken time for any advance in the field to gain acceptance among interpreters (both practitioners and researchers). For example, there was even initial reluctance towards simultaneous interpreting (Gaiba 1998: 163; Fantinuoli 2016), now considered one of the most significant inventions of the 20th century (Stoll 2001). This pattern persists in the present era of the "technological turn" (Fantinuoli 2018), in which interpreters are confronted with Computer-Assisted Interpreting (CAI) tools, the relentless rise of Remote Simultaneous Interpreting (RSI), and, recently, Machine Interpreting (MI).

Unlike other pivotal moments in the brief history of conference interpreting, the current technological turn presents unprecedented characteristics. Firstly, new practices made possible by cutting-edge technologies have rarely been brought about by interpreters, but rather external circumstances, such as the Covid-19 pandemic which contributed to the extraordinary growth of RSI (Fantinuoli 2021: 509). Secondly, unlike previous technological solutions – particularly those developed for terminology management – the most sophisticated tools for interpreters (such as AI-based CAI tools and MI) have a direct impact on the interpreting task itself and the cognitive processes that underpin it. Such tools require interpreters to redistribute their cognitive resources and develop new strategies to obtain genuine quality gains.

There are several factors that explain the scepticism of professional interpreters, especially among older generations. Indeed, many interpreters still tend to prefer either traditional approaches with no interference from CAI tools, reliance on their own skills acquired through professional experience (Prandi 2020), or technological resources that can be used for preparation rather than during the task itself (Corpas Pastor/Fern 2016). The factors underlying this stance include the cost of CAI tools, their lack of functionality and adaptability to the specific needs of individual interpreters, and their influence on the interpreting process, which is already cognitively demanding (Tripepi Winteringham 2010, Corpas Pastor/Fern 2016). On the other hand, the use of technologies seems to be almost unavoidable outside the interpreting booth, namely during preparation, when interpreters usually make use of non-specific resources, such as the Internet, or office tools such as Word or Excel (Moser-Mercer 1992; Berber-Irabien 2010; Corpas Pastor/Fern 2016). The doubts and criticisms voiced by professional interpreters partly account for the delay in research on new interpreting technologies. However, an interest in these new resources and more in-depth research are essential for future developments, not only because the working reality is constantly evolving, but also because other stakeholders – such as RSI and software providers – may contribute to changing the interpreting market and its practices without considering the profession's standards and code of ethics. If interpreters do not take an interest in technologies that genuinely meet their needs, "it is possible that interpreters may soon have to use tools they have not personally selected" (Prandi 2023: 51).

## 1. Consecutive interpreting and Automatic Speech Recognition

Unlike simultaneous interpreting, which has been the focus of numerous studies on CAI tools, few contributions have been devoted to technology-assisted (or technology-enhanced) consecutive interpreting. One reason assisted consecutive interpreting has been studied less than assisted simultaneous interpreting could be its lower diffusion in the present working reality, as well as the fact that the practice itself is realised in two steps. Another reason is likely to be that in communicative settings the use of a CAI tool is logistically less practical than a CAI tool for simultaneous interpreting, which is used in an isolated booth. Furthermore, in formal and high-stakes settings where consecutive interpreting (CI) is generally needed – such as a meeting of two delegations, a meeting of a working group, inaugural ceremonies, or welcoming a foreign delegation (Riccardi 2003: 110) – interpreters seem to prefer the traditional pen and notepad approach, mainly for logistical (outdoor venues, battery life, inadvertently hit buttons on the tablet screen, screen brightness) and confidentiality reasons (Goldsmith 2018).

The first contributions on technology-based CI were focused on the use of tablets for note taking (Drechsel/Goldsmith 2016; Goldsmith 2018; Altieri 2020), which does not however represent a true automation of CI by means of a CAI tool. In the last few years, particular attention has been devoted to the use of Automatic Speech Recognition (ASR) tools as a support for CI (Wang/Wang 2019; Chen/Kruger 2023; Ünlü 2023). ASR-based consecutive interpreting aims to overcome typical problem triggers in note taking, such as numbers, proper names and acronyms. The availability of a transcript is supposed to ensure additional accuracy and completeness in terms of content, since the risk of missing or misunderstanding numbers or names is theoretically reduced compared to traditional note taking, which involves the activation of a number of cognitive efforts. However, questions may arise on the impact that the use of ASR, or rather an overreliance on ASR, may have in terms of reformulation and transparency of the message in the target language, communicability, ergonomics, and reliability and readability of the transcript. Besides, if transcribing a speech were key to conveying a message correctly, interpreters would only need to use shorthand writing. This is not the case, as the drawback of shorthand is that "aucune sténographie ne fixe une conversation, ni même un discours improvisé"[1] (Seleskovitch 1975: 84). In fact, in addition to the gap between oral and written language, decrypting shorthand writing entails not only a process of deciphering, but also a process of converting the message into the target language, which will inevitably slow down the interpreter.

### 1.1 Automatic Speech Recognition

Automatic Speech Recognition (ASR) is the process of converting human speech into a sequence of written words via a computer programme (Jurafsky and Martin 2009 in

1 "There is no shorthand writing that fixes a conversation, nor an impromptu speech" (Seleskovitch 1975: 84). The quotation was translated into English by the author of this article.

Fantinuoli 2017: 26). Until the late 2000s, the application of ASR to interpreting was largely the subject of theoretical speculation (Tripepi Winteringham 2010). However, since the advent of new computational approaches, the quality of ASR technologies has improved to the point where such tools can achieve an error rate comparable to that of human listeners (Cavallo/Ortiz 2018: 23). In particular, advancements in neural networks and deep learning have significantly enhanced the accuracy of ASR and have therefore driven increasing interest in applying it to interpreting (Fantinuoli 2017: 28). However, ASR is still far from being a perfect transcribing tool for interpreters, as there are both intrinsic accuracy/quality limitations and external linguistic elements that pose potential obstacles to its effectiveness, such as nuances, linguistic variation, non-verbal communication, accents, emotions, metaphors, intonation and irony (Cavallo/Ortiz Schild 2018). Additionally, ambiguity, background noise, speech rate, and body language can also affect the quality of the ASR output (Fantinuoli 2017: 28).

### 1.2 Speaker-dependent and speaker-independent speech recognition

An important criterion to categorise speech recognition systems is their adaptability to the speaker's voice. *Speaker-dependent* speech recognition systems can adapt to the characteristics and pronunciation of an individual speaker, thereby improving over time. By contrast, *speaker-independent* systems are not designed to adjust to a specific voice and are generally more effective in contexts where the message is predictable. However, such systems do not achieve the same level of accuracy as *speaker-dependent* ones. An example of a *speaker-dependent* system is *Dragon Naturally Speaking*[2], primarily used in respeaking, while *Sight-Terp* (§2.2) is a *speaker-independent* speech recognition system.

Under ideal conditions, *speaker-independent* speech recognition systems can reach an accuracy rate of 95-99% regarding term recognition. Yet, factors such as noise, music, and spontaneous oral speech affect the quality of recognition, thus representing a limitation for this technology (Fantinuoli 2017). Especially spontaneous oral speech may pose obstacles for ASR systems. Digressions, false starts and incomplete sentences, inferences and implicit references, changes in the speaker's intonation, which typically appear in such speech, are often difficult to be transcribed in an easily readable fashion. For an ASR system to be competitive and reliable for an interpreting task, three features are especially crucial:

1. Speaker-independence, enabling the ASR to adapt to a range of speakers and accents.
2. A low word error rate (WER rate), thus ensuring reliability without imposing an excessive cognitive load on the interpreter.
3. Low latency, allowing the ASR to fit in with the average delay of interpreters (3-5 seconds; Fantinuoli 2017: 28-29).

2 See https://dragon-naturally-speaking-premium.en.download.it/.

## 2. Research methodology

The aim of the present study is to investigate the interpreter-machine interaction in ASR-assisted consecutive interpreting to explore how and to what extent the ASR in this case study, Sight-Terp (§2.2), can enhance consecutive interpreting. Based on the assumption that the ASR can produce excellent results with clear and well-structured speeches (e.g. Defrancq/Fantinuoli 2020; Ünlü 2023), it was decided to investigate the effectiveness of the ASR tool in Sight-Terp when coping with spontaneous oral speech (§1.2), which is common in the interpreter's working life. Secondly, it was assumed that the use of ASR for CI reduces the cognitive load in the first phase, since the effort of note taking is partly replaced by the transcript produced by the ASR. The last assumption is that an automatic generated transcript is a useful source to enhance the interpreting accuracy regarding numbers and proper names. However, it leads the interpreters to stick closer to the text and/or to longer renditions in the target language. In fact, a transcript has a higher word density compared to traditional interpreting notes. If using an ASR tool in CI, the interpreters can choose to use the transcript either as a source for a sight translation or as a backup text for the retrieval of individual elements (such as figures and named entities, see §2.2).

Three hypothesises can be postulated as follows:

1. Interpreters can concentrate more during the source language (SL) listening and comprehension phase thanks to the transcript being produced in real time.
2. The interpreter's ASR-assisted delivery requires more time than in the traditional consecutive mode, especially when the interpreters tend to stick close to the text.
3. While the ASR can assist the interpreter in retrieving specific elements, such as numbers and proper names (or named entities, NE), it is less reliable when attempting to transcribe some features that are typical of spontaneous oral speech (such as repetitions, digressions, false starts, etc.).

### 2.1 Methodology

For this study, data triangulation was employed with the aim of observing the same phenomenon from various perspectives, thereby identifying analogies or differences in the interpreters' behaviour and their interaction with Sight-Terp. Moreover, this approach allowed us to assess the same aspects in the most objective manner possible (see Bowker 2022: 403).

To minimise potential bias, as many variables as possible were kept under experimental control. The two speeches interpreted by the participants (§2.3) contained several dependent variables, such as numbers, proper names and figurative language. These elements were distributed as evenly as possible in the texts (e.g. Seeber 2011; Prandi 2018).

## 2.2 Sight-Terp[3]

Sight-Terp[4] is the CAI tool that was used for this study, is a prototype of *speaker-independent* ASR-based CAI tools for consecutive interpreting. It was developed by Cihan Ünlü (2023) for his experimental MA dissertation at Ankara University. As a prototype of a CAI tool, Sight-Terp was employed exclusively as a case study. Therefore, the results obtained (see §4) do not attempt to come to any prescriptive conclusions about the tool itself.

The interface of Sight-Terp, which can be used on personal computers, tablets and smartphones, is divided into two text areas: on the left hand-side of the screen, Sight-Terp provides the interpreter with a full transcript of the original speech, running in real-time through continuous speech recognition, while on the right hand-side of the screen a translation in the target language is provided, running with a delay of few seconds. The aim is to replace traditional note taking with a transcript and its translation, which can eventually be used as a support for sight translation.

The production of the ASR output and the Machine Translation (MT) output is made possible by the Speech translation function (ST), also known as Automatic speech-to-text translation, which is the distinguishing trait of Sight-Terp. Such an innovative model unifies the ASR tool and the MT tool and therefore goes beyond the conventional cascade approach[5], where ASR and MT are used separately. After pressing the start recognition button, "[Sight-Terp] initiates an automatic speech recognition session using Microsoft Speech Translation API, which is based on an end-to-end system using deep neural network-based modelling" (Ünlü 2023: 67).

The average latency of the transcription has been observed as around two seconds, but it is strongly influenced by the stability of the Internet connection. The result displayed on the screen is a text segmented into utterances that are usually made up of three or four sentences. A full stop is placed at the end of an utterance, when a unit of meaning (or semantic unit) is recognised as such by the AI-based speech model incorporated in Sight-Terp. In addition to full stops and text segmentation, the *TrueText* function by Microsoft initiates a text-normalisation and automatic punctuation process. Such a process "enhances the readability of the text […] and therefore increases the accuracy rate of the neural machine translation" (Ünlü 2023: 68), on which the MT tool is based. Each utterance is then numbered and aligned with the corresponding utterance in the translation, preceded by the same number. This is how the app facilitates the retrieval of the various speech passages and their translation. Clearly, inappropriate segmentation of the speech may have significant repercussions for the

3 For a more detailed description of this tool, see Cihan Ünlü's unpublished MA Thesis "Automatic Speech recognition in consecutive interpreter workstation: computer-aided interpreting tool 'sight-terp'".

4 See https://www.sightterp.net/.

5 Current Speech translation systems are based on two approaches: the cascade and the innovative end-to-end approach. If the cascade approach is applied, the ASR tool and the MT tool work separately; once the ASR has generated its output, this serves as an input for the MT, which eventually produces the final output (the translated text). If the end-to-end approach is applied, the separation into two stages is substituted by a unified process where the source language speech is directly translated into the target language.

readability of the transcript and therefore on the interpreter's comprehension effort, as other studies have already highlighted (cfr. Wang/Wang 2019, Chen/Kruger 2023). For this reason, the *Automatic Text Segmentation* function provided by Sight-Terp "is deployed with the aim to display the reference text in an easy-to-read fashion and allow the user (interpreter) to follow up the source segment with its target MT output thanks to the enumerated style" (Ünlü 2023: 69). It must be noted that other ASR systems convert a speech into written words without paying attention to punctuation and the layout of the text (for example Dragon Anywhere[6], where punctuation marks need to be dictated by the speaker, or the ASR incorporated into Word or Google Drive). Nevertheless, there remains a drawback in the transcription by Sight-Terp: a default text segmentation setting which splits the text whenever the speaker takes a pause of two or more seconds, which can pose a potential obstacle to the correct interpretation of the text and be a potential source of error for the MT tool.

Finally, Sight-Terp offers two more features: the *Named Entity Recognition* (NER) function highlights proper names, acronyms, numbers and time references, and the *Digital Notepad* enables users to take notes in the lower part of the tablet.

### 2.3 Participants and speeches

The participants were six students[7] at the end of their two-year master's degree in Conference Interpreting at the University of Trieste. They had all accumulated more than 72 hours of practice in German-Italian consecutive interpreting during their regular classes and passed their exam "Consecutive from German into Italian 2". Two of them had German as a B-language, while the other four participants had German as a C-language[8]. As the participants were not familiar with Sight-Terp, they took part in an 11-hour seminar composed of one theoretical session (two hours) and three practice sessions (nine hours, three hours each session). The first theoretical session was focused on the presentation of Sight-Terp and on preparing for an assignment using AI-based tools, while the other three sessions gave the student-interpreters the opportunity to develop their own approach to interacting with Sight-Terp. The interpreters could consolidate their approach during these three practice sessions focused on several speeches.

6 See how Wang/Wang (2019) employed Dragon Anywhere in their study.

7 All six students took part at the present study voluntarily, they did not receive any remuneration. Before they were recruited, they had all been informed that the recordings of their deliveries would only be analysed within this study and that their names would be anonymised to guarantee the confidentiality of their personal data. Moreover, they all signed a Declaration of Consent for their deliveries to be processed for the purposes and procedures of the present study.

8 The B language is a language which the interpreters master at a level close to mother tongue. This is the language into which the interpreters can provide a fluent and accurate interpretation. The C language is the language which the interpreters can fully understand and from which the interpreters can interpret into their mother tongue. The definitions above refer to the European language profiles described on the Internet page of the European Union (see https://europa.eu/interpretation/freelance.html).

The data collected refers to two speeches, which the interpreters translated from German into Italian. The former, entitled "*Niassa Naturreservat*", focuses on the Niassa natural reserve in Mozambique, and the latter, "*GAFAM-Unternehmen*" focuses on the biggest tech companies worldwide. Both speeches were prepared by interpreting students, reviewed by an interpreting student who was a German native speaker and delivered by a native speaker from Germany, who was asked not to stick to the written words and to deliver the speech naturally and vivaciously.

| | Length (words) | Duration (mins) | Speech rate (words/min) |
|---|---|---|---|
| Speech A | 621 | 5'35" | 110 |
| Speech B | 610 | 6'02" | 101 |

Tab. 1. Features of the interpreted speeches.

| | Figures | Named entities |
|---|---|---|
| Speech A | 13 | 9 |
| Speech B | 10 | 27 |

Tab. 2. Target elements (figures and named entities) of the interpreted speeches.

## 2.4 Data collection procedure

During the practice sessions leading up to the data collection, no technique nor strategy was suggested, since one of the aims was to observe how the interpreters would deal with ASR-assisted consecutive interpreting in the classroom.

In the second phase of this study, the interpreters were asked to interpret two speeches from German into Italian, one using traditional note taking and one using Sight-Terp. The data collection process took place on the 30th of November and the 1st of December 2023. The data collection was spread over two days so that the participants would have to provide only one interpreting performance per day, thereby decreasing the risk that certain variables (such as stress or fatigue) could affect the quality of the second performance.

Since it was decided to reproduce a typical interpreting assignment, the participants were provided with a topic briefing for each speech. The briefings consisted of an introduction on a fictional conference setting, an abstract of the speech in Italian and in German, some in-depth insights on the main topic, and a thematic glossary (35 entries for speech A and 42 entries for speech B). The interpreters had ten minutes to read the briefing and prepare for the assignment. Before the interpreting task, the interpreters were asked to use Sight-Terp by adopting the approach they personally developed during the practice sessions. The above conditions remained unchanged on both days.

The group of six student-interpreters was randomly divided into two subgroups: the first subgroup interpreted speech A with Sight-Terp and speech B with the help of the notes taken by hand, and the second group did the opposite. The random distribu-

tion of participants into two subgroups made it possible to counterbalance the variable related to the intrinsic features of the speeches and/or of the interpreters and look at the data in most rigorous way possible (Gile 2016: 220). Moreover, the random distribution of independent variables is one of the aspects of experimental design that makes it possible to obtain "maximum efficiency [and] reduce the effect of confounding variables and bias" (*Ibid.*).

Since Sight-Terp is based on a *speaker-independent* ASR and its performance strongly depends on a stable Internet connection, the speeches were pre-recorded and played in a quiet room with adequate internet access. This enabled Sight-Terp to transcribe them with maximum efficiency. The transcription and the MT output by Sight-Terp were also pre-recorded on a tablet screen. Finally, the source language recordings and screen recordings were synchronised. The screen recording allowed to avoid potential differences in the transcription or possible disruption of the app, which would compromise the entire data collection: for instance, a suddenly slower Internet connection or sudden background noises might have resulted in different ASR outputs if Sight-Terp had been used in real-time with each individual participant. This approach allowed all participants to be exposed to the same stimuli and therefore to use the same transcript (i.e. screen recording) as an aid to their consecutive performances.

### 2.5 Focus group discussion and questionnaire

As the number of participants did not warrant quantitative generalisations, a focus group discussion was organised at the end of the data collection phase (e.g. Frittella 2023: 67). This method proved extremely relevant within this study, since the focus group discussion highlighted qualitative aspects of ASR-assisted consecutive interpreting, such as strategic and technical elements. The discussion was moderated by thought-provoking questions (for instance, what was your approach to the use of Sight-Terp in assisted consecutive interpreting? What did you use the transcript for? How did Sight-Terp influence your performance? What type of interpreting strategies did you use in assisted consecutive interpreting?). The discussion also provided insights into the psychological impact of using a transcript for consecutive interpreting. The main topics concerned the strategies adopted to look at notes and to check ASR transcript when delivering the target language speech, as well as strategies to cope with the limitations of ASR and difficulties arising from its use in consecutive interpretation.

After data collection and the focus group discussion, the interpreters were asked to evaluate their experience through a questionnaire. The questionnaire was distributed via Google Forms and consisted of 31 questions of a technical, interpreting and strategic nature – these were thematically divided into five sections:

1. Preparing for an assignment with AI.
2. Sight-Terp-assisted consecutive interpreting.
3. Interaction with Sight-Terp.
4. The performance of Sight-Terp.
5. Future prospects.

Both open and closed-ended questions were adopted to find out how Sight-Terp was used and what strategies were deployed by the interpreters, and to investigate to what extent they were satisfied with their own performance, the transcript produced by Sight-Terp, and Sight-Terp's functionalities.

### 2.6 Limitations of the study

The first limitation of the present study is the number of participants, which is often one of the major limitations of experimental interpreting studies. Moreover, to collect reliable data on ASR-assisted consecutive interpreting, the participants were given the opportunity to practice with this interpreting approach during three sessions, for a total of nine hours (§2.3). The findings obtained must therefore be interpreted with reference to the present study and may not be applied to the whole category of professional interpreters, especially because the participants were six final-year students (§2.3). In addition, the number of professionals who use such tools in their work is still believed to be very limited.

A second limitation is the participants' limited preparation. More targeted training would potentially yield better results in the assisted consecutive approach.

The third limitation is the very fact that Sight-Terp itself was used. Hence, the findings may not be applied to all ASR systems, since at this stage Sight-Terp is still a prototype of a freely accessible *speaker-independent* ASR system. The results of this study cannot be compared to those obtained via a specialised *speaker-dependent* ASR system (such as Dragon Naturally Speaking). However, within this study the key aspect is not really the quality of the ASR transcript, but rather its suitability for CI.

## 3. Results and Analysis

### 3.1 Modes of use

The interpreters' behaviour during the data collection phase was observed at the same time as they interpreted both speeches and subsequently investigated through the answers collected in the questionnaire. Three different modes of using Sight-Terp stood out.

Interpreters 1 and 5 took their notes on a traditional notepad for CI and looked up names, time references or numbers in the transcript produced by Sight-Terp. Their rendition was then generally based on the ASR transcript for short extracts and on their notes for longer extracts.

Interpreters 3 and 4 took their notes on a paper notepad to analyse and process the speech. They then used the transcript to sight translate the text, which was more comprehensive than their notes. Occasionally, they looked at their notes to decipher extracts which were not clear in the transcript.

Interpreters 2 and 6 based both the first and second phase of CI on Sight-Terp without taking any traditional interpreting notes, although they used a paper notepad to note down key concepts and/or transcription errors, such as segmentation or formatting errors (see §3.3).

The final questionnaire confirmed these three approaches, as it revealed that Sight-Terp was used by each interpreter, albeit in different manners and proportions. Out of six student-interpreters, four used Sight-Terp frequently, while two made little use of it. Moreover, five interpreters reported that they mostly used Sight-Terp in the rendition phase of CI (two of them also used it in the phase of listening and analysis). By contrast, one student mainly used Sight-Terp in the first phase; namely, to compare the notes with the transcript or to retrieve missing information. Regarding the use of the transcript and the MT output, five students clearly preferred to use the transcript either to look up individual elements (figures, named entities, terminology) or to perform an on-sight translation. This trend is also confirmed by the fact that the interpreters prevalently used Sight-Terp to retrieve numbers and information they had missed, followed by proper names and lastly by terminology. Only one student also tried to integrate the MT output in the rendition phase. The interpreters evaluated the interaction with Sight-Terp as, on average, 5.5 on a scale out of ten, where ten stands for "easy interaction". One interpreter evaluated the interaction as "one", representing "too difficult" on the scale.

In terms of human-machine interaction, Sight-Terp appears to have been a distracting factor for four interpreters out of six. This was also confirmed in the focus group discussion; indeed, interpreters 2 and 6 stated that the awareness of having a complete transcript at their disposal distracted them during the listening and analysis phase. Similarly, interpreter 5 described a sort of "psychological burden". In the focus discussion she described a feeling of nervousness before the traditional consecutive performance, as is often the case, but this feeling disappeared during the performance itself. By contrast, the opposite occurred in the ASR-assisted consecutive performance: the knowledge of having a full transcript, which is more comprehensive than the interpreting notes, was comforting before the beginning of the performance. However, the effort of looking at Sight-Terp and the worry of watching out for possible transcription errors resulted in feelings of nervousness during the consecutive performance itself.

Moreover, according to two out of six interpreters, Sight-Terp had an adverse impact on eye contact. According to four out of six, it had an impact that was more negative than positive, with five out of six interpreters claiming that Sight-Terp adversely affected the fluidity of their rendition. However, the interpreters had a positive attitude towards the use of ASR for CI, as they stated that more practice with it could eventually lead to better performances in assisted consecutive interpreting. Out of six interpreters, only one had a negative attitude to the use of ASR in consecutive interpreting. In addition, a deeper understanding and awareness of the strengths and limitations of ASR could potentially have a positive impact on the performance.

### 3.2 Interpreters' renditions

The following table compares the duration of the original speeches with that of the interpreters' renditions.

| Interpreter | Original speech (mins) | Rendition (mins) | Mode | Time difference (mins) |
|---|---|---|---|---|
| Int. 1 | 5’35” | 5’13” | ST | -0’22” |
| Int. 2 | | 5’16” | CONSEC | -0’19” |
| Int. 3 | | 4’20” | CONSEC | -1’15” |
| Int. 4 | | 4’22” | ST | -1’13” |
| Int. 5 | | 4’48” | CONSEC | -0’48” |
| Int. 6 | | 7’07” | ST | +1’32” |

Tab. 3. Duration of speech A and duration of interpreters’ renditions. ST = Sight-Terp assisted consecutive, CONSEC = traditional consecutive

| Interpreter | Original speech (mins) | Rendition (mins) | Mode | Time difference (mins) |
|---|---|---|---|---|
| Int. 1 | 6’02” | 5’42” | CONSEC | -0’20” |
| Int. 2 | | 6’22” | ST | +0’20” |
| Int. 3 | | 5’58” | ST | -0’04” |
| Int. 4 | | 4’45” | CONSEC | -1’08” |
| Int. 5 | | 5’14” | ST | -0’48” |
| Int. 6 | | 4’52” | CONSEC | -1’10” |

Tab. 4. Duration of speech B and duration of interpreters’ renditions. ST = Sight-Terp assisted consecutive, CONSEC = traditional consecutive

In the traditional consecutive approach, the duration of the interpreters’ renditions was always shorter than the original speech. By contrast, two interpreters produced a longer rendition than the original speech when using Sight-Terp for their consecutive performance (interpreter 6 Tab. 3, interpreter 2 in Tab. 4). Interestingly, these were the two interpreters who exclusively relied on Sight-Terp for their rendition, thus performing a sight translation.

| Time difference in CONSEC | Median | Time difference in ST | Median |
|---|---|---|---|
| -0’04’’ | (-0’20’’) + (-0’48’’): 2 = *0’34”* | +1’32’’ | (-0’22’’) + (-0’48’’): 2 = *0’40”* |
| -0’19’’ | | +0’20’’ | |
| *-0’20”* | | *-0’22”* | |
| *-0’48”* | | *-0’48”* | |
| -1’08’’ | | -1’08’’ | |
| -1’10’’ | | -1’15’’ | |

Tab. 5. Calculation of the median of time difference. ST = Sight-Terp assisted consecutive, CONSEC = traditional consecutive

Since the values of time difference noticed in consecutive interpreting performances and in Sight-Terp assisted performances were heterogeneous and the number of participants was limited, a calculation of the average time difference would not be sufficiently representative. Instead, the median of time difference was calculated, which can be considered as the middle value (Tab. 5)

Generally speaking, the use of ASR for CI does not always result in a longer rendition than the original speech. However, it can be observed that the interpreters' renditions in assisted consecutive interpreting tend to be longer compared to their performances in the traditional consecutive mode. The median of time difference is 34 sec. in the traditional mode, while the median of time difference is 40 sec. in the assisted consecutive mode.

### 3.3 The performance of Sight-Terp

| | WER | Figures | Proper names | Incorrect segmentation |
|---|---|---|---|---|
| Speech A | 13,6% | 9/13 | 10/18 | 9 |
| Speech B | 8,2% | 8/10 | 23/28 | 12 |

Tab. 6. The performance of Sight-Terp

In speech A, the ASR correctly transcribed nine figures out of thirteen. The remaining four figures were considered as not immediately readable, rather than errors as such. Indeed, two figures were transcribed in full words (*4.500 Quadratkilometern ✍ viertausendfünfhundert, 168 Milliarden ✍ Einhundertachtundsechzig Milliarden*), while two figures were not immediately readable because of their format (*100.000 Dollar ✍ 100 00 $0, 50.000 Dollar ✍ 50 00 0€*). On one occasion, the ASR correctly transcribed the figure, but not its reference (*50% des Einkommens ✍ 50% des Einkaufs*). However, since the ASR system in Sight-Terp does not include the joint recognition of figures and their references, the latter figure was not considered as an error. In speech B, the rate of correctly transcribed figures amounts to 8/10. Once again, one figure was transcribed in full words (*über 130.000 Beschäftigte ✍ über einhunderdreißigtausend Beschäftigte*) and another was partially transcribed (*über 1 Milliarde Stunden ✍ Milliarde Stunden*).

With regard to named entities (NE), here the ASR system seems to have produced reliable results, particularly when transcribing well-known NE (*Afrika, Mosambik, Google, YouTube, Apple, Airbnb*). However, on two occasions there were errors (*Mosambik ✍ Osnabrück, NGO ✍ MGO*). Conversely, less widely known NE, such as the name *Niassa*, the acronyms NATU and GAFAM were not correctly transcribed. In such cases the transcript generated words which are phonetically similar to the original ones (*Gafam ✍ Gafar, Gafan, NATU ✍ Lato, Niassa ✍ erster, nasser, Iassa*).

When the accuracy results on figures and NE are correlated with the interpreters' satisfaction, it appears that the aforementioned errors did not particularly destabilise them. Indeed, according to the questionnaire, the interpreters did not feel that the transcription caused them to commit errors they would have not committed in the

traditional consecutive mode. Incorrect numbers and names, as well as omitted words, did not result in errors in the interpreters' renditions.

For the quality assessment of the transcripts, the Word Error Rate (WER) was calculated using *Amberscript*[9], a software programme which also offers an automatic WER tool. The WER values reported in Tab. 6 indicate that the transcript produced by Sight-Terp is definitely not flawless, but not bad either. Yet, the WER itself is not sufficient to determine whether a transcript is suitable for the purpose of interpreting. For this reason, it was necessary to examine how and where the ASR segmented the text. The ASR feature in Sight-Terp generally produces complete sentences. However, the ASR segmented certain utterances incorrectly: for instance, by separating the subject from its verb. Wrong segmentation occurred more often when the speaker paused to plan the sentence or self-correct, made digressions or had false starts, thereby making it harder for the ASR to transcribe spontaneous speech in an easily readable fashion. This drawback, which in itself is one of the current limitations of ASR, stems from the distance between mostly unplanned spoken language and well-structured written language. Although the readability of the text was not significantly compromised, incorrect segmentation may require more effort, and thus more time, to make sense of the transcribed speech.

By comparing the above-mentioned results with the results of the questionnaire, the relatively good quality of the ASR transcript was also confirmed by the participants: four interpreters reported being satisfied with the transcript quality, and two of them very satisfied. Moreover, according to the questionnaire, three interpreters tended to rely on the transcript, while the other three tended not to rely on the same output, and therefore always preferred to double-check their speech comprehension by means of their traditional notes.

Furthermore, sudden changes in the speech rate or in the speaker's intonation seem to have affected not only text segmentation, but also transcription, leading to errors or omissions of single words (e.g. *in meinem heutigen relativ kurzen Beitrag möchte ich auf eine Frage [eingehen]*), which may also explain the incorrect transcription of well-known named entities (e.g. *Mosambik, NGO*). This is an undeniable limitation of current ASR systems, which should always be taken into account, especially when using such systems in authentic work settings, where sound disruptions or background noise could influence the performance of such CAI tools. As regards the interpreters' opinion about text segmentation and incorrectly transcribed or omitted words, it appears that four interpreters felt they did not commit any errors related to the use of the transcript. One interpreter excluded this possibility, and one interpreter got the impression that some errors could have been avoided in the traditional consecutive mode.

## 4. Discussion

This section discusses the results presented in section §3, in an attempt to give an answer to the hypotheses formulated in section §2.

9 See https://www.amberscript.com/en/resources/wer-tool/.

### 4.1 Listening and comprehension skills

As indicated by the questionnaire, the transcript produced by Sight-Terp is mostly employed as an aid in the renditions of the interpreters, who looked up figures or performed an on-sight translation. One student-interpreter used the transcript mostly to integrate information in note taking or to double-check that the notes had been correctly understood. Otherwise, the transcript was generally perceived by the interpreters as a potential distraction, because it often forced them to weigh up and monitor the solutions given by the ASR. More specifically, during the focus group discussion interpreter 2 and interpreter 6 stated that knowing that a full transcript is available may inadvertently disrupt their listening. This, in turn, adversely affects both speech comprehension and the thorough analysis of its logical structure.

By contrast, four interpreters out of six used Sight-Terp as a support for the retrieval of numbers and sometimes names, without giving up their traditional note taking, which was confirmed to be necessary for deeper comprehension. The transcript itself proved to be a reasonably reliable source of content, which was, however, not a product of the interpreter's elaboration; as such, the meaning of a text with a high word density is not as immediately transparent as the interpreter's notes, which are the product of an in-depth internalisation of the speech.

Therefore, the first hypothesis was disproved, as the participants were not able to concentrate more on listening and comprehension, but instead needed to coordinate additional efforts both in stage 1 and stage 2 of CI. This was particularly evident in the interpreters who decided to rely exclusively on the transcript produced by Sight-Terp.

### 4.2 Speech rendition duration

Out of six ASR-based consecutive interpreting performances, two renditions turned out to be longer than the original speech, whereas all the traditional consecutives were shorter. Therefore, the use of an ASR-generated transcript cannot be said to lead to a longer rendition every single time, but ASR-enhanced consecutives were always slightly longer compared to the traditional ones.

The two longer renditions were produced by interpreter 6 and interpreter 2 (see Tab. 3 and Tab. 4), who based their interpretation exclusively on the Sight-Terp transcript. It can be concluded that both did not process the speech as thoroughly as in traditional consecutive interpreting via note taking. Moreover, during the focus group discussion both interpreters mentioned feeling distracted because they knew they would have a full transcript at their disposal. Therefore, it can be concluded that the lack of a deeper analysis of the speech must have affected their listening and comprehension. This was subsequently compensated for by the attempt to reconstruct the sense of the transcript and to rephrase it during the second phase of consecutive interpreting. Indeed, it must be considered that traditional note taking not only enables a thorough analysis of the speech, but also a preliminary translation process. By contrast, the interpreter who chooses to rely exclusively on the ASR transcript to perform a sight translation, shifts the entire translation process to the rendition phase, therefore requiring more effort. The accumulation of additional effort in the second phase of

CI, which should only entail the free production of the target language speech, may result in an increase in the length of the interpreters' rendition (see interpreter 2 and interpreter 6 and their approach to Sight-Terp, §3.1). Thus, similar to simultaneous interpreting, a tightrope effect (Gile 1999) can be observed in the second phase of assisted consecutive interpreting.

As regards the other interpreters, the use of a transcript may have resulted in an attempt to be more exhaustive and to stick to the transcribed text relatively closely (i.e. trying to reproduce all its parts), because of the psychological burden they felt, as was mentioned by the interpreters during the focus group discussion.

In conclusion, the second hypothesis was partly confirmed by the results. The renditions produced in the assisted consecutive mode are not always longer than the original speech, and a key role is played by the way the ASR tool is employed. However, in comparison with the renditions produced in traditional consecutive interpreting, ASR-assisted consecutive interpreting entails a slight duration increase of the speech rendition.

### 4.3 Readability of automatic transcription

As expected, the ASR feature in Sight-Terp produced acceptable transcripts. Indeed, the positive WER rates (Tab. 6) are also reflected in the interpreters' satisfaction with the transcription: four interpreters reported being satisfied, and two of them very satisfied. The same is also valid for the recognition of numbers and proper names, despite some differences in the accuracy rates between speech A and speech B (Tab. 6).

Nevertheless, in terms of its application to interpreting, the transcripts must be examined in their entirety, taking into account the coherence and cohesion of the texts, as well as their readability. The text segmentation function in Sight-Terp aims to spread the text contents into enumerated chunks, thus distributing the information load. Yet, this very function posed an obstacle to readability, especially when syntactic units were segmented in the wrong place, thereby making the retrieval of the logical structure more difficult. Incorrect segmentations can be attributed to long pauses taken by speakers (pauses of reflection or related to sentence planning efforts), digressions or false starts. Furthermore, emphasis must be placed on the linearity of the transcription process. Unlike a human interpreter, the ASR transcription proceeds in a linear sense, and therefore reproduces the original speech as closely as possible, including its irregularities. Interpreters who choose to rely on an insufficiently transparent transcript run the risk of transferring the task of inferring the deep meaning of the speech to the audience, whose understanding will be hindered rather than facilitated.

It is worthwhile mentioning that transcription errors affecting individual (especially well-known) words or their individual omission usually belonging to collocations did not pose an obstacle to the participants. This can be explained in light of the interpreters' language skills and the briefing, which gave the interpreters the possibility to prepare themselves and anticipate the key ideas of the speech, as would be the case in a real setting. It is believed that interpreters with a good level preparation are able to spot general errors committed by the machine immediately, especially if such errors are likely to be anticipated. In the example "in meinem heutigen relativ kurzen Beitrag

möchte ich auf eine Frage [eingehen]"[10], the only verb that is supposed to belong to this collocation is the verb "eingehen"[11].

In conclusion, the third hypothesis concerning the way ASR deals with spontaneous oral speech was confirmed, since ASR is more accurate than a human being in the retrieval and transcription of technical elements, such as numbers and proper names.

## 5. Conclusion

The present study identified three approaches to performing an ASR-based consecutive interpretation. In all these approaches, an ASR tool is applied to transcribe the source language speech in the first stage of consecutive interpreting, while its output (the transcript) is used by the interpreters in the second stage of CI, either as a source text for a sight translation or as a backup text for the traditional interpreting notes. Both the potential and drawbacks of this approach were discussed, including whether the ASR transcript might be a useful aid to enhance consecutive interpreting (i.e. for the retrieval of numbers and named entities), and whether an overreliance on it can inadvertently hinder the interpreter's performance (less fluency, longer rendition, additional analysis and rephrasing effort). However, the use of a transcript produced by an ASR tool can ultimately enhance consecutive interpreting under certain conditions: for instance, if such tool is applied in a consecutive setting with a stable Internet connection and no background noises, or if it is used as a backup for note taking, which in fact enables a thorough speech analysis.

The interpreting profession is undeniably at a turning point, as new technologies, such as AI and Automatic Speech Recognition (ASR), are significantly influencing the way interpreting is performed. While simultaneous interpreting has been the focus of a number of studies on CAI tools, new ways of performing consecutive interpreting potentially involving Automatic Speech Recognition and/or Machine Interpreting are being experimented with. It is thus imperative that interpreters develop a deeper understanding of the new tools available to develop strategic approaches for their use, especially as they still present some intrinsic limitations. New studies in this field can give the opportunity to continue exploring how CAI tools can be strategically integrated into the interpreting workflow, without adversely affecting the process, but rather offering interpreters actual advantages. Only with their active engagement can interpreters uphold quality and working standards in their profession.

10 "In my relatively short speech today, I would like [to address] a question". The translation into English of the original German example was done by the author of this article.

11 For non German-speakers: an important strategy in the interpretation from German into a romance language is the anticipation of the verb at the end of the sentence. An extensive and solid knowledge of German collocations is crucial to anticipate the meaning of the sentence before the speaker concludes it. In the example reported above, the interpreters who work from German do not need to wait too long to understand that the verb belonging to "auf eine Frage …" is almost certainly "eingehen".

References


Altieri M. (2020) "Tablet interpreting: étude expérimentale de l'interprétation consécutive sur tablette ", *The Interpreters' Newsletter* 25/2020, 19-35.

Baigorri-Jalón J. (2014) *From Paris to Nuremberg, The birth of conference interpreting*, Amsterdam/Philadelphia, John Benjamins.

Berber-Irabien D. (2010) *Information and communication technologies in conference interpreting*, unpublished PhD Thesis, Terragona, Universitat Rovira I Virgili.

Bowker L. (2022) "Computer-assisted translation and interpreting tools", in F. Zanettin / C. Rundle (eds.) *The Routledge Handbook of Translation and Methodology*, London, Routledge, 392-409.

Cavallo P. / Ortiz Schild L.E. (2018) "Computer-assisted interpreting tools (CAI) and options for automation with Automatic Speech Recognition", *TradTerm*, 32/2018, 9-31, <https://www.researchgate.net/publication/330207613_Computer-Assisted_Interpreting_Tools_CAI_and_options_for_automation_with_Automatic_Speech_Recognition> (28 October 2024).

Chen S. / Kruger J. (2023) "The effectiveness of computer-assisted interpreting. A preliminary study based on English-Chinese consecutive interpreting", *Translation and Interpreting Studies*, 18/3, 399-420.

Chernov S. (2016) "At the dawn of simultaneous interpreting in the USSR. Filling the gaps in history", in J. Baigorri-Jalón, K. Takeda (eds.) *New Insights in the History of Interpreting,* Amsterdam/Philadelphia, John Benjamins, 135-166.

Corpas Pastor G. / Fern L.M. (2016) *A survey of interpreters' needs and practices related to language technology*, <http://www.lexytrad.es/assets/Corpas-Fern-2016.pdf> (28 October 2024).

Defrancq B. / Fantinuoli C. (2020) "Automatic Speech Recognition in the booth: Assessment of system performance, interpreters' performances and interactions in the context of numbers", *Target*, 33/1, 73-102.

Drechsel A. / Goldsmith J. (2016) *Tablet Interpreting. The evolution and uses of mobile devices in interpreting*, <https://www.academia.edu/36017504/Tablet_Interpreting_The_evolution_and_uses_of_mobile_devices_in_interpreting> (28 October 2024).

Fantinuoli C. (2016), "InterpretBank. Redefining computer-assisted interpreting tools", in J. Esteves-Ferreira / J.M. Macan / R. Mitkov / O.M. Stefanov (eds.) *Proceedings of the 38th Conference Translating and the Computer*, Geneva, Tradulex, 42-52, <https://aclanthology.org/2016.tc-1.5.pdf> (28 October 2024).

Fantinuoli C. (2017) "Speech Recognition in the Interpreter Workstation", in J. Esteves-Ferreira / J.M. Macan / R. Mitkov / O.M. Stefanov (eds.) *Proceedings of the 39th Conference Translating and the Computer*, Geneva, Tradulex, 25-34, <https://www.asling.org/tc39/wp-content/uploads/TC39-proceedings-final-1Nov-4.20pm.pdf> (28 October 2024).

Fantinuoli C. (2018) "Interpreting and technology: The upcoming technological turn", in C. Fantinuoli (ed.) *Interpreting and technology*, Berlin, Language Science Press, 1-12.

Fantinuoli C. (2021) “Conference interpreting and new technologies”, in M. Albl-Mikasa / E. Tiselius (eds.) *Routledge Handbook of Conference Interpreting*, Routledge, 508-522.
Frittella F.M. (2023) *Usability research for interpreter-centred technology. The case study of SmarTerp*, Berlin, Language Science Press.
Gaiba F. (1998) *The origins of simultaneous interpretation. The Nuremberg Trial*, Ottawa, University of Ottawa Press.
Gile D. (1999) “Testing the Effort Models’ tightrope hypothesis in simultaneous interpreting – A contribution”, *HERMES*, 12/23, 153-172.
Gile D. (2016) “Experimental research”, in C.V. Angelelli, B.J. Baer (eds) *Researching Translation and Interpreting*, New York, Routledge, 220-228.
Goldsmith J. (2018) “Tablet interpreting. Consecutive interpreting 2.0”, in N.K. Pokorn / C.D. Mellinger (eds.) *Community Interpreting, Translation, and Technology*, Special Issue of *Translation and Interpreting Studies* 3, 342-365.
Herbert J. (1980/1952) *Manuel de l’interprète. Comment on devient interprète de conférence*, Genève, Librairie de l’université Georg.
Kalina S., Ziegler K. (2015) “Technology”, in F. Pöchhacker, N. Grbić, P. Mead, R. Setton (eds.) *Routledge Encyclopedia of Interpreting Studies*, New York, Routledge, 410-412.
Moser-Mercer B. (1992) “Terminology Documentation in Conference Interpretation”, *Terminologie et Traduction*, 2/3, 285-304.
Prandi B. (2018) “An exploratory study an CAI tools in Simultaneous Interpreting: theoretical framework and stimulus validation”, in C. Fantinuoli (ed.) *Interpreting and technology*, Berlin, Language Science Press, 29-59.
Prandi B. (2020) “The use of CAI tools in interpreter training: where are we now and where do we go from here?”, Special Issue of *inTRAlinea: Technology in Interpreter Education and Practice*, <https://www.intralinea.org/specials/article/the_use_of_cai_tools_in_interpreter_training> (28 October 2024).
Prandi B. (2023) *Computer-assisted simultaneous interpreting. A cognitive-experimental study on terminology*, Berlin, Language Science Press.
Riccardi A. (2003), *Dalla traduzione all’interpretazione. Studi d’interpretazione simultanea* (*From translation to interpretation, studies on simultaneous interpreting*), Milano, Led.
Seeber K.G. (2011) “Cognitive load in simultaneous interpreting: Existing theories – New models”, *Interpreting*, 13/2, 176-204.
Seleskovitch D. (1975) *Langage, langues et mémoire : étude de la prise de notes en interprétation consécutive*, Paris, Lettres Modernes.
Stoll C. (2001) “Neue Technologien im Konferenzdolmetschen: DigiLab, ein Forschungsprojekt des Instituts für Übersetzen und Dolmetschen der Universität Heidelberg” (“New Technologies in Conference interpreting, DigiLab, a Research project of the Institute for Translation and Interpreting at Heidelberg University”), in A.F. Kelletat (ed.) *Dolmetschen*, Frankfurt am Main, Peter Lang.
Tripepi Winteringham S. (2010) “The usefulness of ICTs in interpreting practice”, *The Interpreters’ Newsletter*, 15/2010, 87-99.

Ünlü C. (2023) *Automatic Speech recognition in consecutive interpreter workstation: computer-aided interpreting tool 'sight-terp'*, unpublished MA Thesis, University of Ankara.

Wang C. / Wang X. (2019) "Can computer-assisted interpreting tools assist interpreting?", *Transletters. International Journal of Translation and Interpreting*, 3, 109-139.

**The Interpreters' Newsletter Special Issue 2024 (29 bis), 79-99**
DOI: 10.13137/2421-714X/36766 ISSN 1591-4127 (print) ISSN 2421-714X (online)
https://www.openstarts.units.it/dspace/handle/10077/2119

# Working memory and simultaneous interpreting – Linking different levels of cognitive load to different levels of source text difficulty

Martina Behr,
Markus Martini
and Pierre Sachse*
University of Innsbruck

Abstract

*Working memory holds information in an immediately accessible state for ongoing cognition. It is thought to play an important role in successful simultaneous interpreting. This study investigated the relationships between, on the one hand, variations in cognitive load (high vs. low) in a working memory task and, on the other hand, the accuracy of rendition of a source text of predefined difficulty (easy vs. difficult) in a simultaneous interpreting task. The results show that more information was retained in the working memory task when the cognitive load was low than when it was high. Correlations with accuracy scores in the simultaneous interpreting task revealed that working memory performance in both high and low cognitive load conditions was related to difficult source text passages, but not to easy ones. The results of this study provide a more nuanced picture of the relationship between working memory and simultaneous interpreting than heretofore.*

Keywords

Working memory, simultaneous interpreting, cognitive load, source text difficulty.

* Authorship contribution statement: Martina Behr and Markus Martini: study conceptualisation, data curation, methodology, project administration, resources, software, data analysis, visualization, writing (original draft, review and editing); Pierre Sachse: resources, review and editing.
Affiliations: Martina Behr: Department of Translation & Interpreting Studies, University of Innsbruck; Markus Martini & Pierre Sachse: Department of Psychology, University of Innsbruck.

According to the Tightrope Hypothesis, "interpreters work close to saturation" most of the time in simultaneous interpreting (SI) (Gile 2009: 182; cf. also Mizuno 2005). During SI, interpreters work under a high cognitive load as they have to simultaneously understand the incoming speech, retain the relevant information and interpret it into another language. Working memory (WM), defined as a cognitive system that holds information available for ongoing cognitive processes (Oberauer/Hein 2012; Martini *et al*. 2015), is thought to play a critical role in this process (Timarová 2008). WM capacity refers to the individual difference construct that reflects the limited capacity of a person's WM (Wilhelm *et al*., 2013), e.g., the number of items that can be temporarily retained in WM while other information is processed simultaneously. WM capacity has been linked to a wide range of cognitive abilities, including language comprehension (Just/Carpenter 1992), fluid intelligence (Conway *et al*. 2002), and the integration of pre-existing domain knowledge (Hambrick/Engle 2002). Thus, WM is also a core cognitive construct for explaining SI performance (Timarová 2012).

Experimental research into the cognitive processes involved in conference interpreting began in the 1970s (cf. Daró 1997). Today, it can be broadly divided into three areas: a) research focusing on the executive functions (cf. Nour *et a*l. 2020: 164; Timarová 2008: 21), b) attempts to prove the Interpreter Advantage Hypothesis (García 2014; cf. Chmiel 2018), and c) research on the relationship between working memory and interpreting in general. This relationship, which is the subject of this study, was analysed in the meta-study carried out by Mellinger/Hanson (2019). In line with Timarová (2008), they divided the studies analysed into two groups: a) studies comparing professional interpreters with one or more control groups (e.g. students), and b) studies relating interpreting performance to interpreters' WM capacity. To the best of our knowledge, so far there have been ten studies which can be assigned to the latter group: Christoffels *et al.* 2003; Liu *et al.* 2004; Tzou *et al*. 2012; Timarová *et al*. 2014, Timarovà *et al*. 2015; Injoque-Ricle *et al*. 2015; Macnamara/Conway 2015; Dong *et al*. 2018; Anssari-Naim 2021 and Bae/Jeong 2021). Three of these studies focus on the correlation between WM and SI performance, i.e. accuracy, in the spoken language, by testing either interpreting students or experts. Tzou *et al*. (2011) compared interpreting students with a control group of untrained bilingual persons. They found that a reading span task was positively correlated with SI performance. Injoque-Ricle *et al*. (2015) tested professional interpreters and found that a listening span task with articulatory suppression predicted SI performance, whereas listening span without articulatory suppression did not. Bae/Jeong (2021) tested interpreting students and found a positive correlation between an operation span task and SI performance. The three studies used different approaches to measure SI performance. Tzou *et al*. (2011) asked two experienced interpreters to rate on a five-point scale the overall quality, i.e. accuracy, voice and interpreter's confidence, and on a scale from 0 to 3, the interpretation of ten sentences equally distributed in the speech. In the study conducted by Injoque-Ricle *et al*. (2015), two bilingual raters assessed fluency, delay and accuracy on a scale from 1 to 10. Bae/Jeong (2021) had two interpreting teachers rating the SI performance by giving a maximum of 100 points for accuracy (40%), fluency (40%) and delivery (20%). None of the three studies provides information on the level of difficulty of the source text (ST).

## 1. The study

The present study contrasted the effects of WM load (high vs. low) with ST difficulty in SI (easy vs. difficult). The aim was to examine experimental variations in the specific WM process of retention and experimental variations of ST difficulty in SI, and to investigate how these relate to each other. Such manipulation of both WM and SI tasks would provide a deeper understanding of the relationships between WM and SI and the conditions under which WM is related to SI.

The WM task in the present study focused on the retention interval between encoding and retrieval. Retaining information in an immediately accessible state so that it can be retrieved at the moment it is needed is a core function of WM and probably essential for successful SI. Therefore, the WM task was manipulated and administered under low and high cognitive load conditions. Cognitive load in this context refers to the attentional demands of a concurrent processing task while relevant information has to be held in WM. Accordingly, cognitive load can be operationalised as the ratio of the time that the processing task occupies attention to the duration of memory retention. The longer attention is occupied by a secondary task, the less time is left to refresh the to-be-retained information, resulting in a decrease in WM performance (Barrouillet *et al.* 2011). According to the time-based resource-sharing model (Barrouillet *et al.* 2004), memory traces decay over time. However, decaying memory traces can be maintained in WM by attentional refreshing, i.e. by bringing the to-be-remembered items back into the focus of attention (Camos *et al*. 2018). Refreshing is thought to occur in the brief moments between two processing tasks (for alternative explanations, see Langerock *et al*. 2024). Reducing the time available between two processing tasks, e.g. by presenting more items to be remembered in the same amount of time, increases cognitive load and decreases WM performance. As a result, fewer items can be maintained in WM. Conversely, increasing the time between two processing tasks, e.g. by reducing the number of items to be processed in the same amount of time, decreases cognitive load and increases WM performance. This means more items to be remembered can be maintained in WM because more time is available to refresh the items. This cognitive load effect has been replicated many times (e.g., Barrouillet *et al*. 2004; Langerock *et al*. 2014; Vergauwe *et al*. 2010) and has been proposed as a Priority-A benchmark that any model of WM should be able to explain (Oberauer *et al*. 2018).

In the WM task, participants had to memorise six digits, followed by a sentence-verification task. The cognitive load manipulation related to the number of sentences they had to judge before recalling the digits in the correct order. In the low WM load condition, four sentences had to be judged; in the high load condition, as many sentences as possible had to be judged. This methodological approach made it possible to directly measure the effect of the discrete WM function of temporarily maintaining information under different interference conditions, while holding information load (number of encoded items) and retention interval (20 seconds) constant.

The SI task focused on a variation in the difficulty of the ST. Experimental studies on the role of WM for SI do not always provide exhaustive information (and reflection) on the characteristics of the ST. Tzou *et al.* (2011) specify that the speech they used was read by an American native speaker at 130 words per minute. Injoque-Ricle *et al*. (2015) state that their speech was pre-recorded and that it was about WM and

dyslexia. Bae/Jeong (2021) used final exam speeches from English into Korean, read by a native speaker at 110-120 words per minute. The aim of the present study was not to use authentic live speeches for the sake of ecological validity but to manipulate the ST for the specific purposes of the study. Accordingly, the ST contained easy sections that should not be challenging, even for beginners, alternating with 'difficult' sections that were expected to lead to errors because of the high cognitive load they should represent for the interpreters, unless strategic decisions allow them to deal with the difficulties successfully.

Assuming that interpreting competence – i.e. among others, the successful use of interpreting strategies – is only visible above a certain level of ST difficulty, the differentiation of ST difficulty should help to test whether experimental results differ according to the difficulty of the ST. The scoring system to measure the interpreting performance was not based on general marks for the overall interpretation, but by assessing propositional accuracy, thus allowing high local granularity.

For the WM task, participants were expected to retain more items in the low cognitive load condition than in the high cognitive load condition (Barrouillet *et al*. 2004, 2007, 2011; Oberauer *et a*l. 2018; Vergauwe *et al*. 2010). For the SI task, participants were expected to show higher propositional accuracy for easy text paragraphs than for difficult text paragraphs. For the relationship between WM performance and SI performance, it was predicted that the higher both WM cognitive load and ST difficulty of SI, the greater the positive correlation.

## 2. Method

### 2.1 Participants

Twenty-eight students on a master's degree programme in conference interpreting from seven different universities in Germany (Mainz/Germersheim, Heidelberg and Cologne) and Austria (Innsbruck, Vienna and Graz) took part in the online experiment[1]. Due to problems with the WM task, such as task interruptions and repeated starts during the experiment, five participants had to be excluded from the analysis. Three participants did not complete the interpretation task, so no file was recorded for them. The analysed data set consisted of twenty participants (20 female, mean age = 23.5 years, standard deviation (SD) = 2.07 years, age range = 21-30 years). The experiment was approved by the ethics committee of the University of Innsbruck. All participants gave their explicit consent before the experiment.

All 20 participants had German as their A language and English as their B language (45%) or C language (55%). They were in their first year of study in the master's programme (85% in the first semester, 15% in the 2nd semester). Nine of them had taken a one-semester introduction to SI from English into German for 45 minutes per week while studying for their bachelor's degree.

1 After being contacted personally by us, interpreter trainers at the seven institutes presented the study in class. Students who wished to take part were offered the prospect of being informed of their test results and could register by e-mail. The response rate was low, with 28 participants from a total of seven institutes.

## 2.2 Procedure

The study was conducted entirely online.[2] After registration, all participants received a link to the study website, which contained information in the following tabs: *Welcome and general information*, *Consent form*, *Instructions* (general procedure, standard browsers (Chrome and Firefox) and download links, Deactivation of AdBlocks), *Task 1* (the WM task under two conditions), *Task 2* (the SI task from English into German), *Prize draw* (all participants were entered into a prize draw for 8 x 70 euros) and *Contact details* of the study managers. All information on the website was presented in German, the participants' L1.

The tabs for *Task 1* and *Task 2* contained information about the nature and procedure of the task, as well as a link to start the task. At the end of Task 2, participants completed a general questionnaire on socio-demographic and related information. Participants were asked to complete all parts of the study within two weeks. The website was created in Google Sites, and the tasks were programmed and hosted using PsychoPy and Pavlovia (Pierce et al. 2019, 2022). All data were automatically recorded under a subject code.

## 2.3 Working memory task and cognitive load manipulation

To measure cognitive load effects in WM, participants performed a WM task under two cognitive load manipulation conditions: part 1 in the low WM load condition and part 2 in the high WM load condition. Each part consisted of a block of ten trials. Part 1 had to be completed first. The time between part 1 and part 2 could be chosen freely. Participants were encouraged to complete each part in a quiet, undisturbed environment and in a rested state.

At the beginning of the task, participants were informed that they would need headphones. They were given the opportunity to test the volume. Each part started with two test runs in the respective condition. During the task, the background colour of the screen was black and the font colour was white. The entire experiment was conducted in German.

An example of an experimental trial is shown in Figure 1. In each condition, the task started with a written instruction: "In the following, you will hear 6 digits between 1 and 9. Remember the digits in the heard order". Digits were generated with speech-to-voice software (https://speechelo.com/). The voice was female. Each audio file lasted ~4.5 seconds (~450ms/digit and an inter-stimulus interval of ~300ms). For the whole duration of the presentation of the digits, participants saw a red "+" in the middle of the screen. After an auditory presentation of six digits, participants were

2 Conducting studies online can help save time and money as well as increase the number of participants by removing time and geographical constraints. Implementing the different tasks so that they can be completed online is time-consuming but has proved to be a reliable way of obtaining the data needed for this type of study, so the same experiment was carried out with professional interpreters in October/November 2024. However, the sound quality of the online recordings made by students may not be sufficient for analysing prosodic characteristics of the interpretations.

presented with simple sentences for 20 seconds (retention interval). The instructions were: "Directly after the series of numbers, you will be presented with simple sentences. Your task is to check whether the sentence is correct. If it is (e.g. "Apples are good for your health"), press the right arrow key for 'the sentence is correct'. If the sentence is incorrect (e.g. All people eat apples) press the left arrow key for 'the sentence is incorrect'. These instructions were followed by a picture of the position of the fingers and the request to position the fingers as shown in the picture for the duration of the task. A pool of 344 German sentences was generated. These sentences were randomly assigned to the participants.

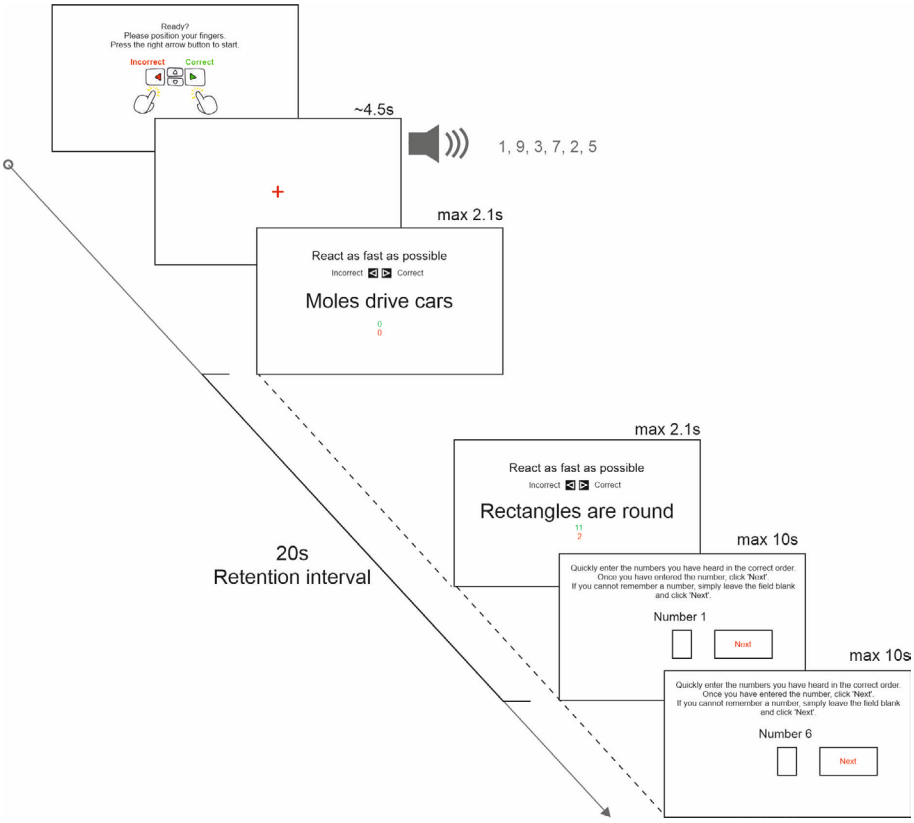


Figure 1. Example of a single experimental trial of the working memory task.
Participants encoded six auditorily presented digits followed by a 20-second retention interval. During the retention interval, participants verified sentences as correct or incorrect. In the low WM load condition, four sentences were presented during the retention interval; in the high WM load condition, participants were instructed to verify as many sentences as possible. The retention interval was followed by a recall phase in which participants were asked to enter, in the order of presentation, the previously heard digits in the fields provided. Background and font colours were changed for presentation purposes; for details, see text.

In the low load condition, four sentences were presented and had to be rated during the 20-second retention interval. In the high load condition, the retention interval was the same, but participants were asked to rate as many sentences as possible during these 20 seconds (in this condition, the test persons rated an average of 14 sentences: M = 14.10, SD = 1.11). Each sentence was presented for a maximum of 2.1s. If no rating of the sentence (correct/incorrect) was given within this time, the sentence was counted as incorrect. The response time began 800ms after a sentence appeared on the screen (maximum response time was 1.3s). The latency between responding and the next sentence was 3.9s in the low load condition and 0ms in the high load condition. As shown in Figure 1, above the sentence presentation, participants constantly saw the cues "Please react as fast as possible" and that the left arrow key referred to 'incorrect' and the right arrow key to 'correct'. Continuous feedback on how many sentences were answered correctly (number coloured green) and how many were answered incorrectly (number coloured red) was displayed below the sentence to increase the participants' motivation to give correct answers. Immediately after rating the sentences, participants were asked to recall the digits in the order in which they had heard them. They were given the following instructions: "Quickly enter the digits you heard in the correct order. Once you have entered a digit, click 'Next'. If you cannot remember a digit, just leave the field blank and click 'Next'". Participants recalled the digits at their own pace, within a maximum of 10 seconds per digit.

To measure whether participants had rehearsed the digits during the processing task, they were asked at the end of both parts of the WM task whether and when they had rehearsed the digits using a three-point scale (1 = "not at all", 2 = "parallel to the sentences", 3 = "in between sentences").

### 2.4 Interpreting task

The interpreting task was to simultaneously interpret a four-minute general speech from English into German.

#### 2.4.1 Source text

The written speech (duration of speech: 4:10 min., length: 439 words, 676 syllables, 33 sentences, speech rate: 107 words per minute, 164 syllables per minute) was read and recorded by a native speaker of British English. It was a speech on the topic of the Red Cross in the fight against HIV/AIDS.

The ST was manipulated by defining two levels of difficulty: easy and difficult. Easy text paragraphs consisted of simple syntax, frequent words and expected sequences (cf. Kalina 2011: 169), obvious idiomatic equivalents in the target language (cf. Mankauskienė 2016: 146f.) and a maximum of one adjective accompanying a noun (cf. Gile 2009: 193; Wang 2015: 69f.; e.g.: "Many countries don't have enough doctors and nurses. Other countries don't have enough money to pay for hospitals. Our volunteers are trying to fill these gaps. But this is not enough."). Difficult paragraphs contained higher demands on the cognitive level due to hypotaxis, less frequent words, terms with less obvious equivalents in the target language and syntactic challenges, i.e. segments specifically placed at the beginning or in the middle of a

sentence which would need to be placed later in the sentence in the German rendition (e.g. "As your experience is an asset allowing us to learn more about the issues that are closest to you, the conference's open format will create an energizing and inspiring context for the day.").

There was a total of 13 paragraphs (on average 2.6 sentences with an average of 31 words). The seven easy paragraphs contained an average of 3.6 sentences with an average of 31.7 words and the difficult paragraphs contained an average of 1.5 sentences with an average of 30.1 words. The speech started with two short sentences that were classified as very easy to make it easier to start interpreting and were not counted. The text then alternated between easy and difficult paragraphs, starting with an easy paragraph. The two levels of difficulty of the ST can be seen in Table 1.

| | whole text | difficult paragraphs | easy paragraphs |
|---|---|---|---|
| words | 439 | 206 | 231 |
| Flesch Reading Ease[3] | 58<br>fairly difficult | 42.1<br>difficult | 70.1<br>fairly easy |
| LIX[4] | 37.33<br>low complexity | 53.48<br>medium complexity | 29.52<br>very low complexity |

Table 1. Levels of difficulty of the source text

### 2.4.2 Procedure

On the study's website, the participants were informed that this task involved interpreting a text simultaneously from English into German, that they would need headphones and a microphone, and that the computer would have access to the microphone. The details of the speech included information on the length of the speech and the topic. The task was preceded by a rehearsal with a 3-minute test speech (on the topic 'The coexistence of generations in our society'). After a recommended break of five minutes, the main speech was to be interpreted.

The participants were asked to find a place where they could concentrate on the task undisturbed. After checking the volume, the test task and the main task could be started by clicking on a link. The participants had the option of contacting one of the experimenters in case of technical difficulties, but none of the 20 participants made use of this option.

### 2.4.3 Scoring system

The accuracy of the interpretations was assessed using propositional analysis. The evaluation grid was created from a combination of automatic and manual counting of the propositions (drawing upon CPDIR, based on Turner/Green 1977, and following the approach proposed by Bovair/Kieras (1985) (cf. Brown *et al*. 2008)). Each

3 Readability score from 0 to 100 (100 = extremely easy to read)
4 Björnsson (1968)

proposition was assigned a value of 1, resulting in a total score of 234 points (119 in total for the easy paragraphs and 115 in total for the difficult paragraphs). The units to be assessed were therefore the propositions, i.e. no units of meaning were formed, nor could such units have been weighted in terms of their relevance (cf. Gieshoff/ Albl-Mikasa 2022). Although in case of very fast speeches, the omission of less important segments can occur and is a rather good strategic choice by the interpreter, the term accuracy not only refers to correctness and precision of the interpretation but also to completeness, i.e., everything the speaker says has to be transferred into the target language.

The evaluation using a stable scoring system was limited to propositional accuracy to obtain a reliable numeric value to compare to the WM performance. Other criteria generally used to assess an interpreting performance – e.g. prosodic features, situation and audience orientation (cf. Behr 2015), or omissions for strategic reasons – were not taken into account. In the event of a significant error such as a *contre-sens*, i.e. a misinterpretation, where the interpreter says the opposite of what the speaker said in the original, raters could mark the deduction of one point in the scoring scheme. The evaluation grid was used by three raters, all of whom have a degree in conference interpreting, German as A language and English as B or C language; two raters have had extensive professional and teaching experience. The inter-rater reliability was high (.953).

Transcripts were created with Whisper AI for all automatically recorded interpretations and corrected manually, where necessary. The raters used these transcripts to assess accuracy.

## 3. Results

Data were analysed using JASP (JASP Team, 2024).

### 3.1 Working memory performance

Performance on the sentence validation task was on average over 70% in both the low load condition (M = 73.88%, SD = 13.99%) and high load condition (M = 79.29%, SD = 9.23%) and did not differ across conditions based on a paired samples t-test, t(20) = -1.85.08, p = .080.

Mean memory performance is shown in Figure 2. Participants retained significantly more digits in the low load condition (M = 67.75%, SD = 21.03%) than in the high load condition (M = 58.92%, SD = 21.02%), t(19) = 3.13, p = .003 (one-tailed), d = .699. Memory performance in the high and low load WM conditions were positively correlated, r(20) = .819, p < .001.

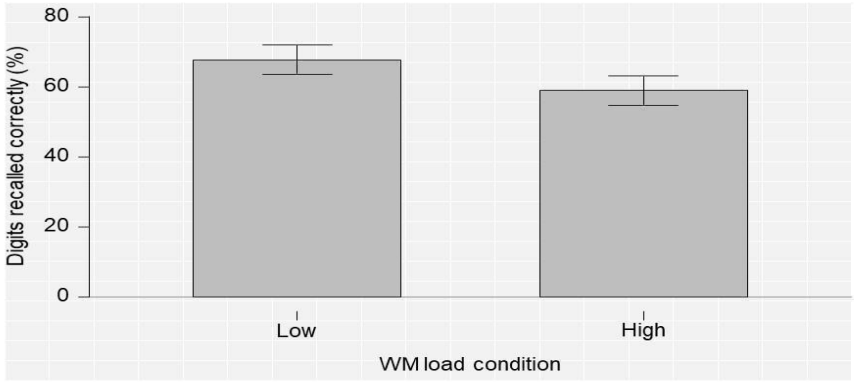


Figure 2. Memory performance in the working memory task conditions
Note: Mean number of digits correctly recalled in the low (left) and high (right) load working memory conditions. Error bars represent 95% confidence intervals.

### 3.2 Rehearsal strategy in the WM task conditions

In the low WM load condition, 13 participants (65%) reported that they rehearsed the digits simultaneously with reading and judging the sentences in the sentence-verification task, and 6 participants (30%) reported that they rehearsed the digits between sentences. In the high WM load condition, 14 participants (70%) reported that they rehearsed the digits simultaneously with the sentences, 2 participants (10%) rehearsed them between the sentences, and 4 participants (20%) reported that they did not rehearse the words during sentence-verification. 14 participants (70%) reported that they did not change their rehearsal strategy, and 5 participants (25%) reported that they changed their rehearsal strategy from the low WM load to the high WM load condition. A binomial test revealed that the proportion of participants who reported changing their rehearsal strategy from the low WM load to the high WM load condition was not significantly different from the expected test value of .55, $p = .112$.

### 3.3 Interpreting task

The distribution of accuracy scores in percentages is shown in Figure 3 and descriptive statistics are shown in Table 2. The overall accuracy scores were around 60%. A Shapiro-Wilk test indicated that the overall accuracy scores were normally distributed, $W = .94$, $p = .193$.

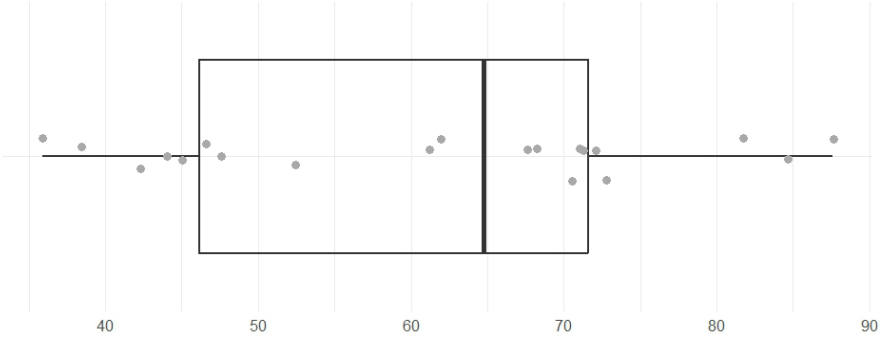


Figure 3. Percentage of accuracy in the SI task
Note: Boxplot for the accuracy scores in percentages. The accuracy scores in percentages (x-axis) were calculated by dividing the individual accuracy score by the maximum accuracy score and multiplying the quotient by 100. The dots represent participants' accuracy scores.

| Accuracy | Mean | SD | Minimum | Maximum |
|---|---|---|---|---|
| Overall | 61.13 | 15.97 | 35.9 | 87.61 |
| Easy text paragraphs | 73.03 | 16.38 | 35.29 | 96.64 |
| Difficult text paragraphs | 48.83 | 18.07 | 18.26 | 78.26 |

Table 2. Descriptive statistics for the overall accuracy, as well as the easy and difficult paragraphs of the SI text, SD = Standard deviation

Accuracy scores for the easy paragraphs show that 73% were correctly reproduced, compared to 48% for the difficult paragraphs (Table 2). Shapiro-Wilk tests showed that the scores for the easy paragraphs, $W = .94$, $p = .260$, and the difficult paragraphs, $W = .97$, $p = .659$, were normally distributed. Accuracy scores were significantly higher for the easy paragraphs than for the difficult paragraphs, $t(19) = 8.38$, $p < .001$, $d = 1.873$.

### 3.4 Relation between working memory performance and interpreting accuracy

As shown in Table 3, Pearson correlations revealed a significant correlation between memory performance in both WM load conditions and overall SI accuracy (easy + difficult ST). Furthermore, memory performance in the low WM load condition was significantly correlated with SI accuracy on the difficult ST passages but not with easy ST passages. Memory performance in the high WM load condition was significantly correlated with SI accuracy on the difficult ST passages but not with easy ST passages.[5]

5 Additional statistical analyses: High and low load WM tasks were strongly correlated, $r(20) = .819$, $p < .001$. Although the sample size was small, two linear regression models were applied. The first model included accuracy scores for easy text passages in the SI task as the

| WM load condition | ST difficulty | $r_{Pearson}$ (p) |
|---|---|---|
| low | easy + difficult | r(20) = .51 (p = .011)*[1] |
| high | easy + difficult | r(20) = .38 (p = .046)*[1] |
| low | easy | r(20) = .31 (p = .184) |
| low | difficult | r(20) = .63 (p = .003)** |
| high | easy | r(20) = .10 (p = .692) |
| high | difficult | r(20) = .61 (p = .005)** |

* p < .05; ** p < .01; [1] = one-tailed; WM = working memory; ST = source text

Table 3. Correlations between WM load conditions and ST difficulty

## 4. Discussion

The present study aimed to further investigate the relationship between WM and SI by varying cognitive load in WM and ST difficulty in the SI task. The results show that (i) in line with our hypothesis, memory performance in the low WM load condition was higher than in the high WM load condition, and (ii) memory performance in both the low and high WM load conditions correlated with SI accuracy for difficult ST passages but not for easy ST passages.

The results showing that increasing cognitive load in WM decreases memory performance are consistent with existing findings (Barrouillet *et al*. 2004; Langerock *et al*. 2014; Vergauwe *et al*. 2010, but also Ricker/Vergauwe 2020). The general finding that task performance in WM is related to SI performance supports theories suggesting that temporary active maintenance of information (in WM) plays an important role in interpreting (Darò/Fabbro 1994; Gerver 1975; Moser 1978) and also supports studies that have found a relationship between WM performance and SI performance (Tzou et al. 2011; Injoque-Ricle *et al*. 2015; Bae/Jeong 2021; see also Mellinger/Hanson 2019). The present study complements previous studies by experimentally varying the parameters of both the WM and the SI tasks.

dependent variable and the two WM tasks (low load vs. high load) as covariates. The results for the first model showed that the ANOVA model was non-significant, F(2,19)=1.78, p = .199, adjusted $R^2$ = .076. The second model included accuracy scores on difficult text passages in the SI task as the dependent variable and the two WM tasks (low vs. high load) as covariates. The results for the second model showed a significant ANOVA model, F(2,19) = 6.12, p = .010, adjusted $R^2$ = .350, but both the low load WM task, β = .397, t = 1.30, p = .236, and the high load WM task, β = .280, t = .87, p = .398, were non-significant. Together with the high standard errors (SE) for the regression coefficients (SE low load WM task = .462; SE high load WM task = .462), these results tend to indicate a multicollinearity problem. One way to reduce the problem of multicollinearity is to exclude one of the correlated variables. In the case of the present study with only two predictor variables, excluding one predictor variable would lead to the results of the correlational analyses (see Results section). Accordingly, the low load WM task shared most of the variance with the accuracy score for the difficult source text passages in the SI task.

Memory performance in both the low and high WM load conditions were significantly related to difficult ST passages. A central assumption of the study was that increasing the time between two sentence-verification items would increase the possibility of refreshing the items to be maintained (low WM load condition), and decreasing that time would decrease the possibility of refreshing the items to be maintained (high WM load condition; e.g. Barrouillet *et al*. 2004). Accordingly, it was assumed that participants who are better at maintaining information, in our case digits, while simultaneously processing other information, in our case sentence-verification, should also show higher SI accuracy. The results of the study support this assumption. We expected a stronger relationship with SI accuracy under the high WM load condition than under the low load condition, as it might be an advantage for interpreters to maintain relevant information in a highly accessible state for a longer temporal interval while simultaneously processing other relevant information. Therefore, the low WM load condition was specifically designed to allow participants to refresh the digits in the 20s-retention interval where distraction through the sentence verification task was low, as there were only four sentences to verify. In contrast, the high WM load condition was designed to maximally minimise participants' possibilities to refresh the digits through a continuous, high frequency engagement in the sentence verification task in the 20s-retention interval. However, the result that in both the high and low WM load conditions WM performance correlated with difficult ST passages may indicate that the two WM load conditions measured different abilities relevant to SI. It can be speculated that in the low WM load condition, the refreshing of information over a prolonged period was measured, whereas, in the high WM load condition, where the focus of attention is bound to process new relevant information, it was the maintaining of information outside the focus of attention in a highly accessible state over a prolonged period. However, it should be noted that performance on the sentence-verification task was below 80% in both WM task conditions. This may indicate that the maximum time allowed for sentence-verification was too short to make a correct decision, or that participants were preoccupied with refreshing the digits to be remembered. While this may have less of an effect in the low WM load condition, where participants had enough time to refresh the digits, it may have a more substantial impact on the measured memory performance in the high WM load condition, where the intention was to maximise attention to the sentence-verification task and, thus, minimise the opportunities to refresh the digits. This is a possible reason why the measures of the two tasks were highly correlated and shared a similar amount of variance with SI accuracy for the difficult ST passages.

Contrary to our expectations, the results of the present study show that the two WM load conditions were not correlated with accuracy in the case of easy ST passages. Easy compared to difficult ST passages were characterised, for example, by simple syntax, frequent words and expected sequences (see Methods section). WM is a construct that keeps relational representations in a directly accessible state and interacts strongly with long-term memory. Reducing the number of relational bindings formed in interaction with long-term knowledge helps WM to divert information away from the focus of attention, making WM more flexible, i.e. more WM capacity is available for processing new incoming information (Rhodes/Cowan 2018). Accordingly, it is conceivable that no relationship between the low WM load condition and easy ST passages was found because the WM load manipulations measured WM-related process-

es, whereas interpreting easy ST passages measured more long-term memory-related processes. The interpretation of easy ST passages may require less WM involvement due to existing language chunks in long-term memory, which can be directly accessed during SI, helping to free WM capacity (Thalmann *et al.* 2019). This view is supported by studies showing that WM capacity is increased when information can be related to long-term knowledge. For example, Fincher-Kiefer et al. (1988) found that participants with high baseball knowledge outperformed participants with low baseball knowledge on a reading span task when sentences were baseball-related, whereas no differences were found when sentences were free of domain-specific knowledge. The efficient use of long-term knowledge is a characteristic of expertise resulting from intensive training. Studies have shown that deliberate practice (training) can improve performance in various domains, such as chess, music or sports (Ericsson *et al.* 2006). Accordingly, practice improves the accuracy, efficiency and ease of the cognitive processes involved and, thus, the execution of actions. Training is said to increase the number and connectivity of information units in long-term memory, which can lead to more effective information processing despite the limitations of WM (Hambrick/Engle 2002; Ericsson/Kintsch 1995). This finding is in line with the assumption that interpreters use their WM more efficiently by reverting to knowledge and strategies specific to SI (Liu *et al*. 2004: 36; for experiential knowledge, see also Wen/Dong 2019: 779; Ghiselli 2022: 73). Several studies have shown that experienced interpreters recur more often to certain strategies than student interpreters (Köpke/Signorelli 2011; Kohn/Kalina 1996; Riccardi 1996; Jörg 1997), and the relevance of the strategic use of anticipation was expressed by Chernov as early as the 1970s: "Our hypothesis is that the basic mechanism that makes SI possible is probability anticipation of the development of the message" (Chernov 2004: 91; see also Mizuno 2005: 749). Up to now, little is known about a possible relationship between cognitive load and specific interpreting strategies (cf. Dong/Li 2020), although these strategies allow cognitive resources to be saved due to automation (cf. Ericsson *et al*. 2006; for automation by and in interpreting training, see Balakhonov 2023).

Higher accuracy scores for the easy paragraphs than for the difficult ones demonstrate the relevance of ST difficulty for interpreting accuracy as one of the central quality criteria and for experimental design. This result confirms that paragraphs with hypotaxis, less frequent words, terms with less obvious equivalents in the target language and syntactic challenges are more difficult than paragraphs without such characteristics. Thus, when analysing what happens when difficult sections are interpreted, easy sections can be useful as reference values. At the same time, this result suggests that differences in performance are not visible below a certain threshold of difficulty. Therefore, it is reasonable to conduct studies where interpreting scores do not refer to a whole text but only to selected passages in the interpretations, e.g. critical sentences (Liu *et al*. 2004), pre-selected and manipulated variables (Timarová *et al*. 2015) or selected sentences (Tzou *et al*. 2011) – provided they are sufficiently difficult. Furthermore, defining the level of difficulty of the whole ST makes it possible to control this variable, which is particularly useful when considering other aspects of interpretation, such as stamina, types of errors, use of strategies.

The present study has several limitations. Firstly, the results are limited to interpreting students at the master's level. The average performance level of 75% is relatively low and should be higher for professional interpreters who have developed the

necessary interpreting skills. Although in the present study, memory performance in the low WM load and accuracy in the difficult ST passages shared almost 40% of their variance, several other factors relevant to SI performance were not measured. One such factor is the level of expertise, i.e. the amount of SI practice. The meta-studies by Wen/Dong (2019) and Ghiselli (2022) provide evidence that the level of expertise is one possible predictor of SI performance. A relevant cognitive ability within expert performance seems to be a better selection of relevant information (Liu *et al*. 2004: 19; 36) – a process that can also be related to WM (Unsworth/Engle 2007; Unsworth et al. 2012; but see also Meinz/Hambrick 2010, who show that WM capacity is a central predictor of sight-reading piano playing independent of practice level). Second, interpreting quality was assessed solely based on propositional accuracy. Accuracy is considered one of the most essential quality criteria, but a thorough evaluation of an interpreting performance must also include aspects such as prosody, communication skills and register (cf. Bühler 1986; Riccardi 2002; Collados Aís *et al.* 2011).

In conclusion, the present study shows a positive relationship between WM and SI, but only under certain conditions. Our results indicate that maintenance of information in the face of distraction is significantly related to SI accuracy for difficult text passages. This finding is a further step toward a better understanding of the relationship between WM and SI, as it highlights that specific WM processes are related to specific processes in SI. Our findings, therefore, suggest that it is important to break up WM and SI processes to get a better understanding of the conditions when WM is related to SI – which is masked when sub-processes on both sites are neglected.

## References


Bae M. / Jeong C. J. (2021) "The role of working memory capacity in interpreting performance: An exploratory study with student interpreters", *Translation, Cognition & Behavior* 4/1, 26-46.

Balakhonov V. (2023) "Corpus-assisted segmentation exercises for simultaneous interpreting", in M. Djovčoš / O. Klabal / P. Šveda (eds) *Bridge: Trends and Traditions in Translation and Interpreting Studies*, Vol. 4/1, 92-108

Barrouillet P. / Bernardin S. / Camos V. (2004) "Time constraints and resource sharing in adults' working memory spans", *Journal of Experimental Psychology: General* 133/1, 83-100. https://doi.org/10.1037/0096-3445.133.1.83

Barrouillet P. / Bernardin S. / Portrat S. / Vergauwe E. / Camos V. (2007) "Time and cognitive load in working memory", *Journal of Experimental Psychology: Learning, Memory, and Cognition* 33/3, 570-585. https://doi.org/10.1037/0278-7393.33.3.570

Barrouillet P. / Camos V. (2015) *Working memory: Loss and reconstruction.* Psychology Press.

Barrouillet P. / Portrat S. / Camos V. (2011) "On the law relating processing to storage in working memory", *Psychological Review* 118/2, 175-192. https://doi.org/10.1037/a0022324

Bartsch L.M. / Oberauer K. (2023) "The contribution of episodic long-term memory to working memory for bindings", *Cognition* 231, 105330. https://doi.org/10.1016/j.cognition.2022.105330

Behr M. (2015) “How to Back the Students – Quality, Assessment and Feedback”, in D. Andres / M. Behr (eds) *To Know How to Suggest... – Approaches to Teaching Conference Interpreting*, Berlin, Frank & Timme, 201-217.

Björnsson, C.-H. (1968) *Lesbarkeit durch LIX*. Pedagogiskt centrum, Stockholms skolförvaltn.

Bovair S. / Kieras D.E. (1985) “A guide to propositional analysis for research on technical prose”, in B.K. Britton / J.B. Black (eds) *Understanding Expository Text*, Hillsdale, NJ, Erlbaum, 315-362.

Brown C. / Snodgrass T. / Kemper S.J. / Herman R. / Covington M.A. (2008) “Automatic Measurement of Propositional Idea Density from Part-of-Speech Tagging”, *Behavior Research Methods* 40/2, 540-545. https://doi.org/10.3758/BRM.40.2.540

Brown J. (1958) “Some tests of the decay theory of immediate memory”, *Quarterly Journal of Experimental Psychology* 10/1, 12-21. https://doi.org/10.1080/17470215808416249

Bühler H. (1986) “Linguistic (semantic) and extra-linguistic (pragmatic) criteria for the evaluation of conference interpreters and interpretation”, *Multilingua* 5/4, 411-439.

Camos V. / Johnson M. / Loaiza V. / Portrat S. / Souza A. / Vergauwe E. (2018) “What is attentional refreshing in working memory?”, *Annals of the New York Academy of Sciences* 1424/1, 19-32. https://doi.org/10.1111/nyas.13616

Chen J. / Yang H. / Han C. (2022) “Holistic versus analytical scoring of spoken-language interpreting: A multi-perspectival comparative analysis”, *The Interpreter and Translator Trainer* 16/4, 558-576. https://doi.org/10.1080/1750399X.2022.2084667

Chernov G. V. (2004) “*Inference and Anticipation in Simultaneous Interpreting. A probability prediction model”*, edited by R. Setton / A. Hild, Amsterdam/Philadelphia, John Benjamins. https://doi.org/10.1075/btl.57

Chmiel A. (2018) “In search of the working memory advantage in conference interpreting – Training, experience and task effects”, *International Journal of Bilingualism* 22/3, 371–384. https://doi.org/10.1177/1367006916681082

Christoffels I.K. / de Groot A.M.B. / Kroll J.F. (2006) “Memory and language skills in simultaneous interpreters: The role of expertise and language proficiency”, *Journal of Memory and Language* 54/3, 324–345. https://doi.org/10.1016/j.jml.2005.12.004

Collados Aís Á. / Iglesias Fernández E. / Pradas Macias E.M. / Stévaux, E. (2011) *Qualitätsparameter beim Simultandolmetschen. Interdisziplinäre Perspektiven*, Tübingen, Narr.

Conway A.R.A. / Cowan N. / Bunting M.F. / Therriault D.J. / Minkoff S.R.B. (2002) “A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence”, *Intelligence* 30, 163-184. https://doi.org/10.1016/S0160-2896(01)00096-4

Cowan N. (1999) “An Embedded-Processes Model of Working Memory”, in A. Miyake / P. Shah (eds) *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control*, Cambridge, Cambridge University Press, 62-101. https://doi.org/10.1017/CBO9781139174909.006

Cowan N. (2001) “The magical number 4 in short-term memory: A reconsideration of mental storage capacity”, *The Behavioral and Brain Sciences* 24/1, 87-185. https://doi.org/10.1017/s0140525x01003922

Darò V. (1997) “Experimental Studies on Memory in Conference Interpretation”, *Meta: Journal Des Traducteurs / Meta: Translators’ Journal* 42/4, 622-628. https://doi.org/10.7202/002484ar

Darò V. / Fabbro F. (1994) “Verbal memory during simultaneous interpretation: Effects of phonological interference”, *Applied Linguistics* 15, 365-381. https://doi.org/10.1093/applin/15.4.365

Dong Y. / Li P. (2020) “Attentional control in interpreting: A model of language control and processing control”, *Bilingualism: Language and Cognition* 23/4, 716-728. https://doi.org/10.1017/S1366728919000786

Engle R.W. / Tuholski S.W. / Laughlin J.E. / Conway A.R.A. (1999) “Working memory, short-term memory, and general fluid intelligence: A latent-variable approach”, *Journal of Experimental Psychology: General* 128, 309–331. https://doi.org/10.1037/0096-3445.128.3.309

Ericsson K.A. / Charness N. / Feltovich P.J. / Hoffman R.R. (2006) *The Cambridge Handbook of Expertise and Expert Performance*, Cambridge, Cambridge University Press.

Ericsson K.A. / Kintsch W. (1995) “Long-term working memory”, *Psychological Review* 102, 211-245. https://doi.org/10.1037/0033-295X.102.2.211

Fincher-Kiefer R. / Post T.A. / Greene T.R. / Voss J.F. (1988) “On the role of prior knowledge and task demands in the processing of text”, *Journal of Memory and Language* 27, 416-428. https://doi.org/10.1016/0749-596X(88)90065-4

García A.M. (2014) “The interpreter advantage hypothesis”, *Translation and Interpreting Studies* 9/2, 219-238. https://doi.org/10.1075/tis.9.2.04gar

Gerver D. (1975) “A psychological approach to simultaneous interpreting”, *Meta Translators’ Journal* 20, 119-128. https://doi.org/10.7202/002885ar

Ghiselli S. (2022) “Working memory tasks in interpreting studies: A meta-analysis”, *Translation, Cognition and Behavior* 5/1, 50-83. https://doi.org/10.1075/tcb.00063.ghi

Gieshoff A.C. / Albl-Mikasa M. (2022) „ Interpreting accuracy revisited: a refined approach to interpreting performance analysis”, Perspectives, 32(2), 210–228. https://doi.org/10.1080/0907676X.2022.2088296

Gile D. (2009) *Basic Concepts and Models for Interpreter and Translator Training*, rev. ed., Amsterdam/Philadelphia, John Benjamins.

Hambrick D.Z. / Macnamara B.N. / Campitelli G. / Ullén F. / Mosing M.A. (2016) “Beyond born versus made: A new look at expertise”, in B. H. Ross (ed.) *The Psychology of Learning and Motivation*, Elsevier Academic Press, 1-55.

Hambrick D.Z. / Engle R.W. (2002) “Effects of domain knowledge, working memory capacity, and age on cognitive performance: An investigation of the knowledge-is-power hypothesis”, *Cognitive Psychology* 44, 339-387. https://doi.org/10.1006/cogp.2001.0769

Han C. (2018) “Using rating scales to assess interpretation. Practices, problems and prospects*”*, *Interpreting* 20/1, 59-95. https://doi.org/10.1075/intp.00003.han

Hu M. / Fan W. (2021) “The interpreter advantage in executive functions—A systematic review and meta-analysis”, *Forum for Linguistic Studies* 3/1, 131-161. https://doi.org/10.18063/fls.v3i1.1251

Hudjetz A. / Oberauer K. (2007) “The effects of processing time and processing rate on forgetting in working memory: Testing four models of the complex span paradigm”, *Memory and Cognition* 35/7, 1675-1684. https://doi.org/10.3758/BF03193501

Injoque-Ricle I. / Barreyro J.P. / Formoso J. / Jaichenco V.I. (2015) “Expertise, Working Memory and Articulatory Suppression Effect: Their Relation with Simultaneous Interpreting Performance”, *Advances in Cognitive Psychology* 11/2, 56-63. https://doi.org/10.5709/acp-0171-1

JASP Team. (2024) JASP (Version 0.18.3) [Computer software].

Jörg U. (1997) “Bridging the gap: Verb Anticipation in German-English Simultaneous Interpreting”, in M. Snell-Hornby / Z. Jettmarová / K. Kaindl (eds) *Translation as Intercultural Communication – Selected Papers from the East Congress – Prague 1995*, Amsterdam/Philadelphia, John Benjamins, 217-228.

Just M.A. / Carpenter P.A. (1992) “A capacity theory of comprehension: Individual differences in working memory”, *Psychological Review* 99, 122-149. https://doi.org/10.1037/0033-295X.99.1.122

Kalina, S. (2011) “Maß für Maß. Eine vergleichende Profilanalyse von Diskursen beim Dolmetschen”, *trans-kom* 4/2, 161-175.

Kohn K. / Kalina S. (1996) “The strategic dimension of interpreting”, *Meta* 41/1, 118-138. https://doi.org/10.7202/003333ar

Köpke B. / Signorelli T.M. (2011) “Methodological aspects of working memory assessment in simultaneous interpreters”, *International Journal of Bilingualism* 16/2, 183-197. https://doi.org/10.1177/1367006911402981

Langerock N. / Vergauwe E. / Barrouillet P. (2014) “The maintenance of cross-domain associations in the episodic buffer”, *Journal of Experimental Psychology: Learning, Memory, and Cognition* 40/4, 1096-1109. https://doi.org/10.1037/a0035783

Langerock N. / Oberauer K. / Throm E.V. / Vergauwe E. (June 20, 2024) “The cognitive load effect in working memory: Refreshing the empirical landscape, removing outdated explanations”, https://doi.org/10.31234/osf.io/3zbe2

Liu M. / Schallert D.L. / Carroll P.J. (2004) “Working memory and expertise in simultaneous interpreting”, *Interpreting* 6/1, 19-42. https://doi.org/10.1075/intp.6.1.04liu

Mankauskienė D. (2016) “Problem Trigger Classification and its Applications for Empirical Research”, *Procedia - Social and Behavioral Sciences* 231, 143-148.

Martini M. / Furtner M.R. /Maran T. / Sachse P. (2015) “Information maintenance in working memory: an integrated presentation of cognitive and neural concepts”, *Front. Syst. Neurosci*. 9:104. doi: 10.3389/fnsys.2015.00104

Meinz E.J. / Hambrick D.Z. (2010) “Deliberate practice is necessary but not sufficient to explain individual differences in piano sight-reading skill: The role of working memory capacity”, *Psychological Science* 21/7, 914-919. https://doi.org/10.1177/0956797610373933

Mellinger C.D. / Hanson T.A. (2019) "Meta-analyses of simultaneous interpreting and working memory", *Interpreting* 21/2, 165-195. https://doi.org/10.1075/intp.00026.mel

Meinz E. J. / Hambrick, D. Z. (2010) "Deliberate practice is necessary but not sufficient to explain individual differences in piano sight-reading skill: the role of WMC", *Psychological Science* 21(7), 914-919. http://dx.doi.org/10.1177/0956797610373933.

Mizuno A. (2005) "Process model for simultaneous interpreting and working memory", *Meta* 50/2, 739-752. https://doi.org/10.7202/011015ar

Moser B. (1978) "Simultaneous interpretation: A hypothetical model and its practical application", in D. Gerver / H. W. Sinaiko (eds) *Language Communication and Interpretation*, New York, Plenum Press, 353-368. https://doi.org/10.1007/978-1-4615-9077-4_31

Moser-Mercer B. (2000) "Simultaneous interpreting: Cognitive potential and limitations", *Interpreting* 5/2, 83–94. https://doi.org/10.1075/intp.5.2.03mos

Nour S. / Struys E. / Woumans E. / Hollebeke I. / Stengers H. (2020) "An interpreter advantage in executive functions? A systematic review", *Interpreting* 22/2, 163-186. https://doi.org/10.1075/intp.00045.nou

Oberauer K. / Hein L. (2012). "Attention to Information in Working Memory", *Current Directions in Psychological Science* 21/3, 164-169. https://doi.org/10.1177/0963721412444727

Oberauer K. / Lewandowsky S. / Awh E. / Brown G.D.A. / Conway A. / Cowan N. / Ward G. (2018) "Benchmarks for models of short-term and working memory", *Psychological Bulletin* 144/9, 885-958. https://doi.org/10.1037/bul0000153

Oberauer K. / Lewandowsky S. / Farrell S. / Jarrold C. / Greaves M. (2012) "Modeling working memory: An interference model of complex span", *Psychonomic Bulletin and Review* 19/5, 779-819. https://doi.org/10.3758/s13423-012-0272-4

Oberauer K. / Süß H.-M. / Wilhelm O. / Wittmann W. (2008) "Which working memory functions predict intelligence?", *Intelligence* 36, 641-652. https://doi.org/10.1016/j.intell.2008.01.007

Peterson L. / Peterson M.J. (1959) "Short-term retention of individual verbal items", *Journal of Experimental Psychology* 58/3, 193-198. https://doi.org/10.1037/h0049234

Peirce J. / Gray J.R. / Simpson S. / MacAskill M. / Höchenberger R. / Sogo H. / Kastman E. / Lindeløv J. K. (2019) "PsychoPy2: Experiments in behavior made easy", *Behav Res* 51, 195-203. https://doi.org/10.3758/s13428-018-01193-y

Peirce J. W. / Hirst R. J. / MacAskill M. R. (2022) *Building Experiments in PsychoPy*. 2nd edition. London: Sage. https://uk.sagepub.com/en-gb/eur/building-experiments-in-psychopy/book273700

Rhodes S. / Cowan N. (2018) "Attention in working memory: attention is needed but it yearns to be free", *Ann N Y Acad Sci* 1424(1), 52-63. https://doi.org/10.1111/nyas.13652

Riccardi A. (1996) "Language: Specific strategies in simultaneous interpreting", in C. Dollerup / V. Appel (eds) *Teaching Translation and Interpreting 3*, Amsterdam/Philadelphia, John Benjamins, 213-222.

Riccardi A. (2002) “Evaluation in interpreting. Macrocriteria and microcriteria”, in E. Hung (ed.) *Teaching Translation and Interpreting: Building Bridges*, Amsterdam/Philadelphia, John Benjamins, 115-125. https://doi.org/10.1075/btl.42.14ric

Ricker T.J. / Vergauwe E. (2020) “Consistent failure to produce a cognitive load effect in visual working memory using a standard dual-task procedure”, *Journal of Cognition* 3/1:15, 1-10. https://doi.org/10.5334/joc.108

Ricker T.J. / Vergauwe E. (2022) “Boundary conditions for observing cognitive load effects in visual working memory”, *Memory and Cognition* 50/6, 1169-1185. https://doi.org/10.3758/s13421-022-01320-3

Shao Z. / Chai M. (2021) “The effect of cognitive load on simultaneous interpreting performance: An empirical study at the local level”, *Perspectives* 29/5, 778-794. https://doi.org/10.1080/0907676X.2020.1770816

Signorelli T.M. / Obler L. (2012) “Working memory in simultaneous interpreters”, in J. Altarriba / L. Isurin (eds) *Memory, Language, and Bilingualism: Theoretical and Applied Approaches*, Cambridge, Cambridge University Press, 95-125. https://doi.org/10.1017/CBO9781139035279.005

Thalmann M. / Souza A. S. / Oberauer K. (2019) “How does chunking help working memory?”, *Journal of Experimental Psychology: Learning, Memory and Cognition* 45, 37-55. https://doi.org/10.1037/xlm0000578

Timarová Š. (2008) “Working memory and simultaneous interpreting”, in P. Boulogne (ed.) *Translation and Its Others. Selected Papers of the CETRA Research Seminar in Translation Studies 2007*. http://www.kuleuven.be/cetra/papers/papers.html (16 August 2024)

Timarová Š. (2012) *Working memory in conference simultaneous interpreting*, unpublished PhD dissertation, Charles University, Prague/KU Leuven.

Timarová Š. / Čeňková I. / Meylaerts R. / Hertog E. / Szmalec A. / Duyck W. (2014) “Simultaneous interpreting and working memory executive control”, *Interpreting* 16/2, 139-168. https://doi.org/10.1075/intp.16.2.01tim

Timarová Š. / Čeňková I. / Meylaerts R. / Hertog E. / Szmalec A. / Duyck W. (2015) “Simultaneous interpreting and working memory capacity”, in A. Ferreira / J.W. Schwieter (eds.) *Psycholinguistic and Cognitive Inquiries into Translation and Interpreting*, Amsterdam/Philadelphia, John Benjamins, 101-126. https://doi.org/10.1075/btl.115.05tim

Turner A. / Greene E. (1977) *The construction and use of a propositional text base* (Technical Report 63), Boulder, CO, University of Colorado, Institute for the Study of Intellectual Behavior, https://www.colorado.edu/ics/sites/default/files/attached-files/77-63.pdf (30 June 2024)

Tzou Y.-Z. / Eslami Z. R. / Chen H.-C. / Vaid J. (2011) “Effect of language proficiency and degree of formal training in simultaneous interpreting on working memory and interpreting performance: Evidence from Mandarin–English speakers”, *International Journal of Bilingualism*, *16*(2), 213-227. https://doi.org/10.1177/1367006911403197

Unsworth N. / Engle R.W. (2007) “The nature of individual differences in working memory capacity: active maintenance in primary memory and controlled search from secondary memory”, *Psychol Rev.* 114(1), 104-32. https://doi.org/10.1037/0033-295X.114.1.104

Unsworth N. / Spillers G.J. / Brewer G.A. (2012) "Working memory capacity and retrieval limitations from long-term memory: an examination of differences in accessibility", *Q J Exp Psychol (Hove)* 65(12), 2397-410. https://doi.org/10.1080/17470218.2012.690438

Unsworth N. / Fukuda K. / Awh E. / Vogel E.K. (2014) "Working memory and fluid intelligence: Capacity, attention control, and secondary memory retrieval" *Cognitive Psychology 71*, 1-26. https://doi.org/10.1016/j.cogpsych.2014.01.003

Vergauwe E. / Barrouillet P. / Camos V. (2010) "Do mental processes share a domain-general resource?", *Psychological Science* 21/3, 384–390. https://doi.org/10.1177/0956797610361340

Wang, Hairuo (2015) "Error Analysis in Consecutive Interpreting of Students With Chinese and English Language Pairs", *Canadian Social Science* 11 (11), 65–79.

Wen H. / Dong Y. (2019) "How does interpreting experience enhance working memory and short-term memory: A meta-analysis", *Journal of Cognitive Psychology* 31/8, 769-784. https://doi.org/10.1080/20445911.2019.1674857

Wilhelm O. / Hildebrandt A. / Oberauer K. (2013) "What is working memory capacity, and how can we measure it?" *Frontiers in Psychology* 4. https://doi.org/10.3389/fpsyg.2013.00433

Yuan R. (2021) "Material development for beginner student interpreters: How does text structure contribute to the difficulty of consecutive interpreting?", *The Interpreter and Translator Trainer*. https://doi.org/10.1080/1750399X.2021.1950979

Yudes C. / Macizo P. / Bajo T. (2011) "The influence of expertise in simultaneous interpreting on non-verbal executive processes", *Frontiers in Psychology* 2/309, 1-9. https://doi.org/10.3389/fpsyg.2011.00309

# Contributors

Clare Donovan graduated from ESIT in 1982 as a Conference Interpreter. She worked as a freelance interpreter for International Organizations, Ministries and Conferences, drawing on her professional experience to underpin her research and her doctoral thesis on Fidelity in Interpretation, submitted in 1990. She taught at ESIT, Sorbonne Nouvelle, from 1984 to 2021 in various capacities, including for many years Director for the Conference interpreting Master. She currently teaches at ISIT and at Lomé University in Togo. She headed the Interpretation Service at the OECD for ten years. A founding member of EMCI and of PAMCIT and a registered trainer at AfIA (African Interpretation Academy), she has organized training for trainer workshops and training courses worldwide, with a particular emphasis on curriculum development and the specificities of interpreting into a B language. She is on the Editorial Board of several journals.

Cecilia Manzoni obtained her master's degree in Specialised Translation and Conference Interpreting from the University of Trieste in 2024. Her thesis was entitled 'Distance interpreting and its impact across the board: standards, trends and the case of the European institutions'. During the master's course, Cecilia Manzoni participated in the EMCI (European Master's in Conference Interpreting) curriculum, obtaining the EMCI certificate in her final exams. In 2023 she was awarded her diploma from the Collegio Universitario Luciano Fonda, a university honours college based in Trieste. She now works as a conference, liaison and community interpreter.

Judith Platter is senior lecturer at the Centre for Translation Studies at the University of Vienna, where she teaches speech-to-text interpreting (STTI), plain and easy

language, audio description and specialised translation at BA and MA level. She also teaches STTI at the University of Graz and at the Postgraduate Centre of the University of Vienna. She holds an MA in Interpreting Studies (German, Italian, English and French) and a PhD in Transcultural Communication from the University of Vienna, specialising in accessibility and speech-to-text interpreting. Her dissertation focused on the professionalisation of STTI in Austria.
She is a member of the Vienna Interpreting Research Group. Her main research interests are interpreting/translation didactics, accessibility and translation, intermodality/intermediality in translation and professionalisation. A trained and certified speech-to-text interpreter herself, Judith Platter is a founding member and current chairwoman of the Austrian Association of Speech-to-Text Interpreters (OESDV) and represents professional interests at national and international level, including within FIT. In 2022, she chaired the organising committee of the 6th European Conference of Speech-to-Text Interpreters, ECOSVienna 2022.

Alessia Valente completed her Master's degree in Specialised Translation and Conference Interpreting at the University of Trieste in 2024. Her Master's thesis, entitled 'Written interpreting: a new form of accessibility in interpreting', explored the emerging field of written interpreting and its crucial role in improving accessibility. During her Master's degree, Alessia Valente followed the curriculum of the European Master's in Conference Interpreting (EMCI) and obtained the EMCI certificate. She currently works as a Network Operation Centre Specialist at M-Cube, focusing on providing technical support to Dutch and English-speaking clients.

Michele Restuccia graduated in 2024 with a Master's degree in Specialised Translation and Conference Interpreting at the University of Trieste. His thesis focused on the use of Automatic Speech Recognition to support consecutive interpreting from German into Italian. In November 2024 he started his PhD in Linguistic and Literary Studies at the University of Udine and Trieste. His research project explores how professional interpreters adopt computer-assisted interpreting (CAI) tools and seeks to devise strategies for integrating these tools into consecutive interpreting between German and Italian.

Martina Behr is a full professor of interpreting studies at the University of Innsbruck, Austria. Her main research interests focus on conference interpreting quality and assessment, training and cognition in conference interpreting, and system dynamics in conference interpreting.

Markus Martini is an associate professor at the University of Innsbruck, Austria. His research focuses on learning and memory, in particular the biasing effects of the post-encoding phase on memory.

Pierre Sachse holds the Chair of General and Experimental Psychology at the University of Innsbruck, Austria. His main research interests include eye-movement research (macro-cognition).

CALL FOR PAPERS
THE INTERPRETERS' NEWSLETTER
ISSUE 30 (2025)

Editors: Vincenzo Lambertini (University of Turin) and Natacha Niemants (University of Bologna)

Issue 30 of *The Interpreters' Newsletter* welcomes submissions that contribute to describing **the interpreting activity in all the situational contexts in which it is required**, by using either traditional or innovative analytical methodologies.

**Topics of interest for issue 30/2025**

Topics of interest include but are not limited to the following areas:

- Conference and dialogue interpreting education, through innovative and collaborative approaches
- Artificial Intelligence, CAI tools and machine interpreting
- Distance/Remote interpreting practices and developments
- The role(s) and agency of interpreters in different modes and settings
- Interpreting ethics
- Corpus-based interpreting studies
- Discourse analytical approaches to interpreting
- (Multimodal) Conversation Analysis of interpreter-mediated events
- Observational and cognitive research on interpreting
- Systematic and scoping reviews

**Practical information and deadlines**

Papers must be submitted in English or French (thoroughly proofread if authors are not native speakers of these languages) and describe original research which is neither published nor currently under review by other journals or conferences. Submitted manuscripts will be subject to a process of double-blind peer review.

Authors must follow the **guidelines** and the **style sheet** before submission. They are available at: https://www.openstarts.units.it/communities/36bbf0cb-888b-4f2a-aa4f-7d4f10ee3601.

Manuscripts should be between 6,000 and 7,500 words long, including references, abstract (150-200 words) and keywords (5-10) and should be sent as Word attachments to: interpretersnewsletter@units.it. **Email subject**: NL30 PAPER; **File Name**: author's name_NL30.

| | |
|---|---|
| Manuscript submission: | 30th March 2025 |
| Results of peer-reviewing process: | 30th June 2025 |
| Publication: | December 2025 |