International Journal of Language, Translation and Intercultural Communication

Vol 11 (2026)

The Sustainability of Interpreting as a Profession in the Era of Artificial Intelligence (early access; full release by the end of February 2026)

The rising star of Speech-to-text Interpreting

Judith Platter, Daniela Eichmeyer-Hell

doi: 10.12681/ijltic.42531

To cite this article:

Platter, J., & Eichmeyer-Hell, D. (2026). The rising star of Speech-to-text Interpreting : how AI transcription tools impact its sustainability within the sensitive accessibility ecosystem. *International Journal of Language, Translation and Intercultural Communication*, *11*, 27–42. https://doi.org/10.12681/ijltic.42531

The rising star of speech-to-text interpreting: how AI transcription tools impact its sustainability within the sensitive accessibility ecosystem.

Judith Platter
Centre for Translation Studies, University of Vienna
judith.platter@univie.ac.at

Daniela Eichmeyer-Hell
Centre for Translation Studies, University of Vienna
de@delfinterpret.com[1]

Abstract

This paper explores forms of intraprofessional and interprofessional collaboration in the field of speech-to-text interpreting (STTI) in Germany and Austria, taking into account the changing technological landscape with AI-transcription tools as the latest technological evolution. To this end, we first give an overview of accessibility, accessibility regulations and how STTI is established therein as an accessibility service. Then, we depict entities within the STTI network, in terms of human and technology involved. In part three, we present two STTI settings: a remote broadcasting setting for a broad variety of users and an on-site community interpreting setting for an individual user. By analyzing the data, we shed light on the service provision and decision-making parameters which were taken into account by the STTI network entities, providing STTI professionals with arguments to ensure successful and inclusive communication involving human interaction and highlighting respective challenges and opportunities.

Keywords*: speech-to-text interpreting (STTI), intraprofessional and interprofessional collaboration; AI-based transcription tools, economic and professional sustainability; user-centred decisions*

1 Introduction

A post on LinkedIn in September 2025 shows, as just one example, the general threat that AI represents not only to professional sustainability but also to accessibility:

> The quality of a lot of subtitles being broadcast right now is embarrassing.
> And who's to blame?
> Not just the companies churning them out at bargain-bin rates, but the broadcasters and streamers who accept them. Who sign off on captions that look like they've been dictated by a drunk parrot, then thrown into a tumble dryer.
> That's not accessibility. That's contempt in Arial font, white on black.
> Because let's be honest, if you actually cared, you wouldn't accept subtitles that cut off mid-sentence or having line breaks which render a sentence unreadable. You wouldn't shrug at captions that say (MUSIC) for an entire concert. You wouldn't let dialogue vanish faster than the biscuits in an edit suite.
> But here we are. Broadcasters demand it cheap. Vendors oblige. And the end result is subtitles that feel less like communication and more like a crossword puzzle written by someone who's never even seen a crossword, never mind compiled one.
> And the absurd bit is this.
> You'd never let it happen anywhere else. Imagine airing a drama where half the dialogue was dipped, or a nature documentary where Attenborough's voice kept freezing mid-syllable while a note on screen says (PROBABLY A BIRD). You wouldn't dare.
> So why is it fine for subtitles? Why is it fine when the audience is d/Deaf or hard of hearing?

[1] Although this article is a result of joint discussion, Platter is the author of the sections dedicated to the STTI CI setting, Eichmeyer-Hell of the sections dedicated to the STTI BC setting.

> Accessibility is not a nuisance. It's not a discount option. It's part of the show. Treat it properly, or admit you're broadcasting contempt.
> d/Deaf viewers deserve better. Always have. Always will.
> The question is, when will the people in charge stop pretending the garbage being broadcast right now counts as inclusion?" (Henderson 2025)

This LinkedIn post highlights the lack of shared responsibility in quality and the lack of opportunity for feedback. In this paper, we will elaborate on the ecosystem surrounding STTI as an accessibility service as well as the networks of those whose aim is to foster accessibility and grant it to deaf[2] and hard-of-hearing people in communicative settings. We first describe the basis of accessibility and STTI, then we elaborate on aspects of collaboration and cooperation in interpreting, highlighting which processes and entities are the main network agents within STTI; subsequently, we will present two different settings of STTI provision: one being STTI for an anonymous, heterogeneous user group in the context of live captioning for private broadcasting providers, which we consider a special case of STTI facility, the other being STTI as a public service interpreting setting for a single individual user known to the STT professionals. In both settings, the STTI service is organised and coordinated by a consultant interpreter (Downie, 2020, p. 92) and performed by trained and certified STTI professionals. Outside this inner interpreting network, the consultant interpreter interacts with different entities for the efficient provision of the service, such as users, institutional stakeholders, and the wider public (Skaaden 2021); our paper will specifically address those entities from the perspectives of STTI.

By analyzing the major network correlation with particular focus on the user's perspective, we try to showcase how AI tools impact networks' equilibrium regarding STTI as an accessibility service. We will shed light on the service provision and decision-making entities in the extended interpreting network. We hereby elaborate on the parameters which are taken into account when opting for human STTI or AI-based speech-to-text (STT) services, as well as on the challenges and opportunities which arise for STTI professionals in this specific context of accessibility service aimed at successful and inclusive communication. Furthermore, we will try to establish arguments to make the profession of human STTI sustainable.

2 Accessibility and the STTI ecosystem

2.1 Accessibility and STTI

Accessibility is a right granted by law: on an international level, the United Nations Convention on the Rights of Persons (UN CRPD) with disabilities came into force on May 3, 2008. In the Definitions of this regulation, the terms "communication" and "language" are elaborated, referring explicitly to spoken and signed languages. STTI, as a live conversion of spoken language into a readable, comprehensible text as we define it within this paper, is only implicitly specified in expressions like "accessible multimedia as well as written [...] and alternative modes, means and format of communication, including accessible information and communication technology" (UN CRPD, Definitions). Article 4 of the same Convention underlines that necessity to take the specific needs of people with disabilities into account, stating that "information and communication technologies [...] suitable for persons with disabilities" should be researched and developed

[2] We here refer to deaf people, taking into consideration that in most cases, Deaf people consider Sign Language to be their native language and, therefore, preferred communication mode.

(UN CRPD, Article 4, g) and that "the training of professionals [...] working with persons with disabilities [...] to better provide the assistance and services guaranteed" should be promoted (ibid, i). Additionally, the Convention underlines that priority should be given to technologies at an affordable cost (ibid, g).

At the European level, the European Accessibility Act from April 17, 2019, in its introductory paragraph directly refers to the UN CRPD and highlights the necessity to ensure the accessibility of products and services for "persons with disabilities". As regards STTI, it lists "access to audiovisual media services" and the respective requirements of "subtitles for the deaf and hard of hearing, audio description, spoken subtitles and sign language interpretation" (EEA, Article 3, Definitions, 6) as well as "real time text" as "form of text conversation in point-to-point situations or in multipoint conferencing where the text being entered is sent in such a way that the communication is perceived by the user as being continuous on a character-by-character basis" (ibid, 14). In the Annex section, the EEA calls for "adequate quality" of accessibility components (EEA, Annex I, Section IV).

In Germany and Austria, diverse existing disability regulations on national levels have been harmonized, in the fairway of these regulations; in some cases, they received new denominations, highlighting a shift in focus from disability to accessibility: Since June 28, 2025, the former Equal Rights Act for people with disabilities (Behindertengleichstellungsgesetz BGG in Germany and BGStG in Austria) is called Accessibility (Enforcement) Acts (Barrierefreiheitsstärkungsgesetz BFSG in Germany and Barrierefreiheitsgesetz BaFG in Austria); both regulations underline the necessity for information to be accessible in general, to be perceivable by means of more than one sensory channel, and to be presented in an understandable and, for the user, perceivable and appropriate written style. In this case, again, STTI is not mentioned explicitly. The Austrian regulation repeats the regulations of the EEA, highlighting that subtitles and SL interpretations have to be "complete, in a suitable quality required for a correct appearance" (BaFG, section 4 b, bb – our translation).

As an outcome of this regulatory basis and the more general terminology used in it, the principles and modes of implementing accessibility measures are rather vague, which leads to flexibility, on the one hand, and a large bandwidth of possible interpretations on the other. This is even more obvious when it comes to STTI services, as they are not explicitly excluded, but also not expressly required, in contrast to SLI, which is already mentioned within these regulations. Furthermore, the included parameters of "completeness", "correctness", "adequateness", and "suitable quality" leave room for interpretation as well, as there is no quality assessment method mentioned, by which it would be possible to decide whether these requirements are met. Although there might be clearer parameters within translation and interpreting (T&I) studies (for quality assessment models on STTI, see e.g. Alonso-Bacigalupe & Romero-Fresco 2024; Eichmeyer-Hell 2021), outside our discipline, they tend to be paired with and put on the same level as "affordable costs", as those are clearly measurable in numbers when it comes to the cost to be paid for STTI services. This last principle might be particularly problematic in the age of AI, where cost-effectiveness as a quantitative entity shown in figures is one of the main selling points, and "suitability" as a qualitative concept would require more complex evaluation schemes or becomes a more subjective parameter.

In this context, the rising popularity of "accessibility" is a decisive factor; due to the aforementioned regulations, public awareness and requests for prospective services are promoted.

This comes along with an increasing interest in services offered by technology-driven companies with clear profit-orientation.

Unlike at its beginnings (around 2002 in Germany and 2010 in Austria), there is, nowadays, a considerable number of trained and certified professional STT interpreters who offer their services as freelancers; they can be found on their individual websites, on experts lists of the professional associations in Germany (BSD 2025) and Austria (ÖSDV 2025), or through advisory offices installed by regional governments or paragovernmental institutions. T&I institutions offer training in accessibility, amongst others, in STTI (Platter 2025, Eichmeyer-Hell forthcoming); this results in a considerably higher number of T&I practitioners aware of quality aspects for accessibility services. They are trained in STTI competences, dictating with user-dependant speech recognition software, interpreting strategies, and the use of suitable technology, amongst them AI-based speech recognition software.

2.2 The STTI network

2.2.1 Human entities within the STTI network

As a given reality and as in many other professions, for the delivery and performance of their services on the regulatory basis mentioned above, STT interpreters are part of professional networks. In T&I studies, these networks are defined as groups of professionals working closely together on the same project, in a shared working context with a high degree of interaction and coordination (e.g. Risku et al. 2016; Kadrić & Iacono 2023), which are needed to efficiently put into action four key mechanisms: responsibilities, routines, roles, and relationships. In literature, it is often stressed that the result and the success of projects are highly influenced by the fact that entities within the network share a common goal – providing a service or a quality product – and are well aware that the project could not be managed by individuals alone. From that perspective, they are highly dependent on collaboration and cooperation, terms often used as synonyms (for definitions of collaboration and cooperation in SLI, see Gile & Napier 2020), but distinguishing between inter-professional collaboration and intra-professional collaboration. The latter refers to collaboration activities between persons of the same professional group, in our case, STT interpreters, while inter-professional collaboration in this paper will be considered in the sense of collectively coordinated actions borne by various professional groups (Kadrić and Iacono 2023).

Although research into conference interpreting has already addressed interaction and collaboration entities, parameters and activities for interpreters collaborating in booth or for interpreters in teams (e.g. Nogueira 2022; Hoza 2022, Platter et al. 2025), for STTI, the definition of collaboration entities, parameters and activities involved in the organization and implementation of STTI services has not been given much attention, so far. Undoubtedly, though, in line with other interpreting professions, there are collaborative process phases which enable, facilitate and enhance the provision of the service in the broader social and economic context.

In interpreting, Kutz 2010 (289–290) describes the interaction of individual interpreters and coordinating interpreters within a broader collaboration context characterized by specific processes, aims and results. He highlights specifically the importance of organizational preparation, designating its aims as “ensuring appropriate working conditions, providing preparation material and ensuring the functioning of the technical devices” (our translation, 289). In this context, he also elaborates on the fact that steps of organisational preparation are often undertaken by language service providers (LSPs), consulting interpreters or professional

associations, as contractors often have limited knowledge about how interpreting services are put into practice. As a result, interpreters need to be active outside their core activity, being proactive in organisational and coordinating tasks within projects and project networks, before, during and after the interpreting assignment itself (Kadrić and Iacono 2023: 169), never losing sight of the product they are working on. This certainly also applies to STTI, as a newer discipline.

When it comes to STTI in the audiovisual context, research can be found on live-subtitling, focusing on collaboration practices for the delivery of the product (Pražák et al. 2020), naming the cost-parameter as one leading factor (206):

> Our goal in a five-year research project funded by the Czech Government and co-funded by Czech Television was to explore possibilities for fully automatic live subtitling (i.e., without respeaking) and to develop a technology for live subtitling through respeaking under the following constraints:
> – cheap operation with one respeaker and no correctors, reaching common live subtitling accuracy
> – remote operation over the internet to support part-time respeakers working from home
> – minimum live subtitle latency, especially for sports programs – semi-supervised respeaker training process (Pražák et al. 2020: 206)

Other research elaborates on networking parameters in terms of technique, demand for STTI and respective training needs, highlighting "the increasing demand for STTI worldwide" (Alonso-Bacigalupe & Romero-Fresco 2024: 541). These aspects can be integrated with studies focusing on the visibility and physical presence of (spoken-language and SLI) interpreters in TV or for the media in general or on-screen and off-screen interpreting (Pöchhacker 2018). For STTI in broadcasting, the authors draw on his following statement, implying that the network entities all share a necessary kind of team spirit aligned towards a common goal – in this case, the audiovisual product including live text (aka live-captions) – and, therefore, the willingness to assume shared responsibility for it (Kadrić & Iacono 2023: 170; 171):

> Alternatively, if the screen viewed by the audience is considered as a product comprising both visual and audio signals, media interpreters can never be off-screen; rather, they would invariably be part of the audiovisual product, and the on-screen/off-screen distinction would become pointless (Pöchhacker 2018: 258.

Differing from the cited publications, for the aim of this paper and in line with interpreting as interaction (Wadensjö 1998), we assume that the users – varying from individually known users in triadic conversational events to a heterogenous, mostly anonymous group of users in AVT context – are direct parts of the project network (Kadrić and Iacono 2023: 174, Norberg & Stachl-Peier 2018).

2.2.2 Technology within the STTI network

As an interpreting specialization, STTI has always been impacted and even defined by available and feasible technological tools, starting from the earliest approaches of handwriting, (conventional and specialized), keyboard-based, and, in the later stages, speaker-dependent speech-recognition systems. Professionals are aware of the fact that technology is just one aspect of professional, specialized, and, therefore, efficient live rendition of spoken into written texts, and besides the mastering of those tools, STT interpreters need user-centred strategies.

With the dawn of AI-based, apparently free to use, tools which are available 24/7, entities within the STTI network nevertheless seem at least to be tempted to try mere AI-based technologies or even substitute human STT interpreters with them, in different kinds of interpreting settings. We therefore consider technology to be an additional entity within the network.

Let us draw the line: There are professionals, and there are tools which – used by professionals – deliver the best user- and product-oriented quality. Unfortunately, and paradoxically resulting from a fostered accessibility need and awareness paired with a more flexible approach in terms of acceptable quality, entities taking decisions within STTI networks assume that accessibility services can be done by human interpreters but are better done by some technical tool, disregarding the fact that accessibility requires interpretation practice, ethical approaches and human skills, such as empathy and cultural sensibility, and awareness and understanding of context. Given the number of professionals, this "need" for solely technological solutions, for example, AI-generated transcripts, seems even more unfounded and negates the questioning of quality.

3 Research setting and methodology

We present case studies based on observation protocols. For the broadcast examples, we want to refer to three cases, all three from 2025 (and partially also referring to the years before). The first case study analyzes the Wimbledon tennis tournament in June 2025, the second one STTI for Champions League matches, and the third one STTI for Austrian national football league matches. In the broadcast context, STTI is generally referred to as "live captioning", however, taking into account the translational task encountered, we stick to the term STTI , be it intra- or interlingual.

For the community interpreting (CI) setting, the observation data were collected between June and September 2025, and the observation protocols were drawn on communications prior and after three events in education via e-mail, messenger, telephone calls, and observations of the interpreted event itself. The data were collected and analyzed in a qualitative way based on the network entities and practices mentioned under section 3.2.

3.1 Broadcasting setting

We refer to three cases of STTI for broadcasters (BC), all sports entertainment: the Wimbledon tournament, the Champions League, and the Austrian national football league. In 2024, human STTI was provided for the Wimbledon tournament and was planned for 2025, too; however, a couple of weeks before the tournament's start, human STTI was cancelled; this decision was withdrawn, as 10 days before the tournament, the consulting interpreter's company providing STTI was called to again set up a team of professional STT interpreters. Overall, 25 STTI professionals were needed and had to be coordinated but only got a contract for one week. A few days after the tournament and the STTI started, the company was informed that the second week would be captioned by AI. Watching the AI-generated captioning, it became quite evident that the AI tool was trained using the human output during the first week, but still, the quality apparently lagged behind the human output, and for the users, some strange details popped up, such as conversations that had nothing to do with the broadcast content, and comments corresponding to a different court. That shows that AI, in this case, was not able to correctly interpret the context and therefore transcribed utterances from other courts and commentators' booths.

The second example: human STTI was provided for Champions League matches broadcast by Amazon Prime between 2021 and 2025. In September 2025, a sudden change to AI-captioning took place. The same effects were observed as in the Wimbledon setting.

The third example is different: there was a call for human STTI of the Austrian Bundesliga matches broadcasted by Sky. Interestingly, the provider(s) had to be located in the UK. There was

an Austrian company offering the possibility to invoice Sky Austria, as apparently, an important taxation issue existed. The offer was declined, with the insistence that the service must be provided in the UK. When watching the matches, an extremely small amount of text was delivered, and the live text was definitely not produced by a native or even B2-professional. To quantify the amount of content delivered in these live texts, we transcribed a 2-minute section of both the source text and the target text, and assessed the target text using the WIRA-model (Eichmeyer-Hell forthcoming). The result of the assessment shows that only 16 % of the content was rendered, and accuracy was 96.36 %, far below the minimum requirement of 98 %.

All these examples have one thing in common: there was no opportunity for user to make complaints, and the users were not taken into account at all.

3.1.1 Entities and actions within the BC STTI network

The entities within the network are the contracting broadcaster, and the technical provider of the live text into the stream, who subcontracts the STTI-provider, in this case a partnership of two interpreters, who, in turn, set up the team composed of professional freelance speech-to-text interpreters. The user, being an unknown heterogeneous entity, is not taken into account as part of that network, as mentioned above. There is limited interaction within the BC network, as Figure 1 shows.

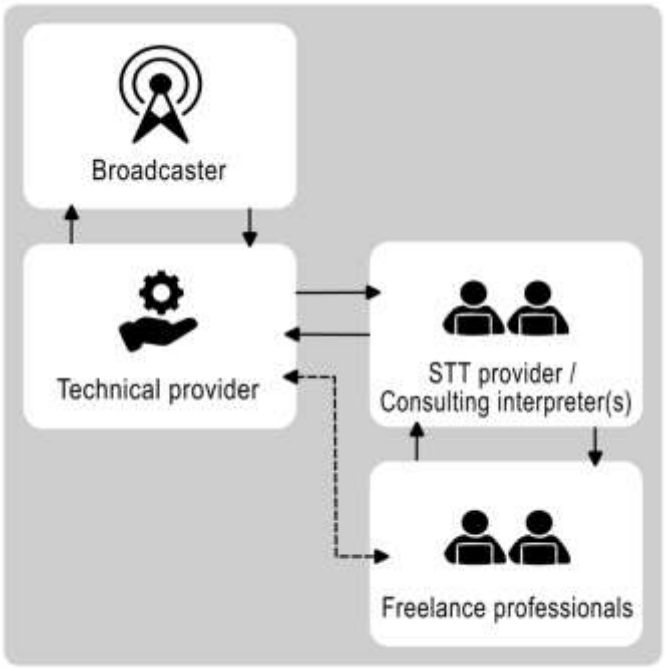

Figure 1: Interaction within the BC STTI network

In the first two cases mentioned above, the broadcaster working with technical providers of live texts did not have any contact with the STTI provider nor with the professionals delivering the live texts. The core interaction – in the form of cooperation – takes places between the technical provider and the STTI provider; however, the freelancers will have direct contact with the technical staff during the assignment (dashed arrow), but neither before nor after. Rates, terms and conditions, as well as operational details, are agreed upon between the technical provider and the

STTI provider (cooperation). Eventually, there is also collaboration involved, as some information on content and names may be given to the STTI provider. The STTI provider elaborates a scheme for collaborative preparation. The freelancers, together with the STTI provider, manage the files containing the preparation material. The feedback loop takes place between the technical provider and the STTI provider, who forwards the received feedback to the freelancers, who, in turn, give their feedback to the STTI provider who forwards it, if applicable, to the technical provider.

The invoicing and payment workflows are as follows: the freelancers invoice the STTI provider, who, in turn, invoices the technical provider, and the latter invoices the broadcaster. The payment workflow is exactly the other way round, starting with the broadcaster, and the freelancers being the last ones to get paid.

3.1.2 Economic parameters within the BC STTI network

In all three settings, the pricing of the service was established based on an offer and acceptance thereof. Of course, the broadcasting sector has always been price-sensitive, so the negotiated rates were reasonable, but acceptable.

3.2 CI setting

The community interpreting (CI) setting is constituted of three STT-interpreted events at educational institutions on the outskirts of Vienna. In all events, the STT user was a hard-of-hearing adult attending parent evenings, the first one at an after-school care facility, the second one held at the primary school, the third one at the kindergarten. All dialogic encounters involved the user, and other participants, including teachers and other parents, as well as STT interpreters. The teachers at kindergarten had previous experience with STTI for the user, the teachers at primary school and at after-school care encountered their first experience with STTIs.

3.2.1 Entities and actions within the CI STTI network

The entities within the network are the user, a hard-of-hearing person in her 40s, a consulting STT interpreter, two other STT interpreters, the local educational community, the regional educational authority, and the regional office for social services. The other participants, teachers and parents, interacting during the STTI event, will be excluded in respect of this analysis, as they were not actively included in the organization.

Their actions in respect of the successful STTI process are as follows:

- As an experienced user of STTI in previous years, the user was aware of her active role in organizing the STTI event. She was very proactive in ensuring the availability of STT interpreters and the appropriate working conditions on site:
    - She informed the consulting interpreter in June 2025 of the dates, time, and the estimated duration of the events in September 2025 (1–2 hours), taking into account that due to the location, the availability of a team of two interpreters for the events requires planning.
    - Also in June, she informed the teachers at all three institutions of the presence of the STT interpreter team and their specific needs (equipment to be made available on site). Taking into account that the primary school and the after-school care

facility were inexperienced with STTI, she advised them to get into contact with the consulting interpreter if more detailed information on equipment and payment of STTI by public institutions were needed.

- She instructed her partner to talk to the local educational community in order to explain that she intended to use STT interpreters for the event in the primary school and that the local educational community had to pay for it. She asked them to get in touch with the consulting STT interpreter if further information would be required.
- The last week of August, she reminded the teachers of the equipment needed for the STT interpreters and the fact that they had to access the location at least 30 minutes before the beginning of the event.

– The consulting STT interpreter has known the user for 5 years and has been organizing interpreter teams ever since. She knew that in view of institutions and persons experiencing STTI for the first time, a more active involvement in coordination would be required:

- In June 2025, she contacted a pool of possibly available STT interpreters, informing them about the planned dates and the fact that the events would require the presence of at least one STT interpreter on site. She informed the interpreters that the interpreter on site would require a car. Subsequently, she established a team of three STT interpreters to interpret: interpreter A (on site) and B (online) covered the event at the after-school care facility on September 1 and in primary school on September 2; interpreter B (online) and C (on-site) covered the event at kindergarten on September 11th. She shared a file containing the relevant facts on the location premises and the overall information on the assignments: number of estimated participants, information on the user requirements and preferences, technical tools required to be provided by the STT interpreter on site, and information on fees to be paid to the interpreters, mentioning also that the invoicing to the regional office for social services was to be done by the consulting STT interpreter. Furthermore, she provided the technical equipment to the STT interpreters (an online STTI platform, a video conferencing tool, a technical information sheet and a microphone set for use on site); with interpreter A, she had a test session to ensure the functioning and smooth handling of the equipment on August 29th. She collected the interpreters' invoices, established a general invoice for each of the three events, attached the signed services form to them and sent them to the regional office for social services. After receipt of the payment, she forwarded the payment to the STT interpreters.
- The consulting interpreter contacted the regional education authority to clarify the responsibility of the local education community to pay for STTI at the primary school in July and August; she had three e-mail conversations and two phone calls, amounting to 4 hours in total. Furthermore, she explained the role of STTI in the expected events to the local educational authority representative in a phone call, held on August 13th, with a duration of 45 minutes. She mentioned to him that based on the user's preferences, mere AI transcription tools were not the right solution for interactive, multi-party communication systems, nor were they compliant with privacy expectations the other participants in the events might have. She elaborated on the fact that the user has clear needs and preferences, in terms of projection tools

(a hand-held tablet which she can position on her knees, in case there was no table on which to put it) and highlighted that professional STT interpreters adapt the rendition style to the users' preferences in the matter of legibility and readability; furthermore, they have the possibility to correct and intervene, in the event that this is needed.

- The STT interpreter team consisted of three trained and certified STT professionals, with experience of 11 (interpreter A), 4 (interpreter B) and 7 (interpreter C) years, all of them having a background as trained conference interpreters with around 20 years' experience. Interpreter C acted as consulting interpreter:
    - All STT interpreters prepared collectively for the assignment by sharing background information on the topics to be treated and the terminology of the events: they communicated by e-mail and via an established messenger group for instant communication. They individually prepared the shortcut and speech-recognition inputs on their devices. Interpreters A and C worked on site: They calculated the needed travel time individually, ensuring that they would be on site at least 30 minutes prior to the start of the event; they installed the internet connection for the video conferencing and STTI live-text tool, checked with interpreter B in the messenger service, that everything was working and connected the microphones to their laptop. They interacted with the user on the events, ensuring that they brought a tablet for her and that the live-text would meet her requirements in the specific setting (they ensured a suitable font size and font colour, taking into account the lighting conditions and the room size). They interacted with the other participants – teachers and parents alike – on site, explaining their presence, the use of the microphone, and the availability of the live-text, and answered questions in this regard. For invoicing, they ensured that the service form was signed by the user and the teacher after the assignment. Additionally, they asked for feedback from the participants and the user alike.
- The authorities were the local educational community, the regional educational authority and the regional office for social services. The regional office for social services and the regional educational authority had previous experience with human STTI, while for the local educational community, it was the first time.
    - The regional office for social services paid for the STTI services for the user for the previous years, as the events at the kindergarten were covered by the public funding for private needs. They are not usually involved in the organization of the event and are contacted only afterwards when the coordinating STT interpreter sends them the invoice as well as the signed service form.
    - The regional educational authority bears the costs for STTI in schools provided for pupils and students. They were contacted by the user and the consulting STT interpreter at the beginning of August to clarify the fact that the planned event in September in primary school had to be borne by the local educational community. Furthermore, they were asked by the consulting interpreter whether the user has the possibility of objecting to the proposal of the local educational community to use an AI-based transcription tool, instead of the preferred human STTI services. The regional education authority contacted the local educational community to clarify

which approach might be taken in ensuring STTI for the user at the primary school event.

- The local educational community bears the costs of STTI in schools, provided for parents of pupils and students attending the local schools. They contacted the user's partner at the beginning of August, asking for information on the estimated costs for STTI services at schools' events. In the middle of August, they contacted the consulting STT interpreter telling her that the school would provide a smartboard with an AI-based transcription service as an "adequate and financially feasible" solution for the user.

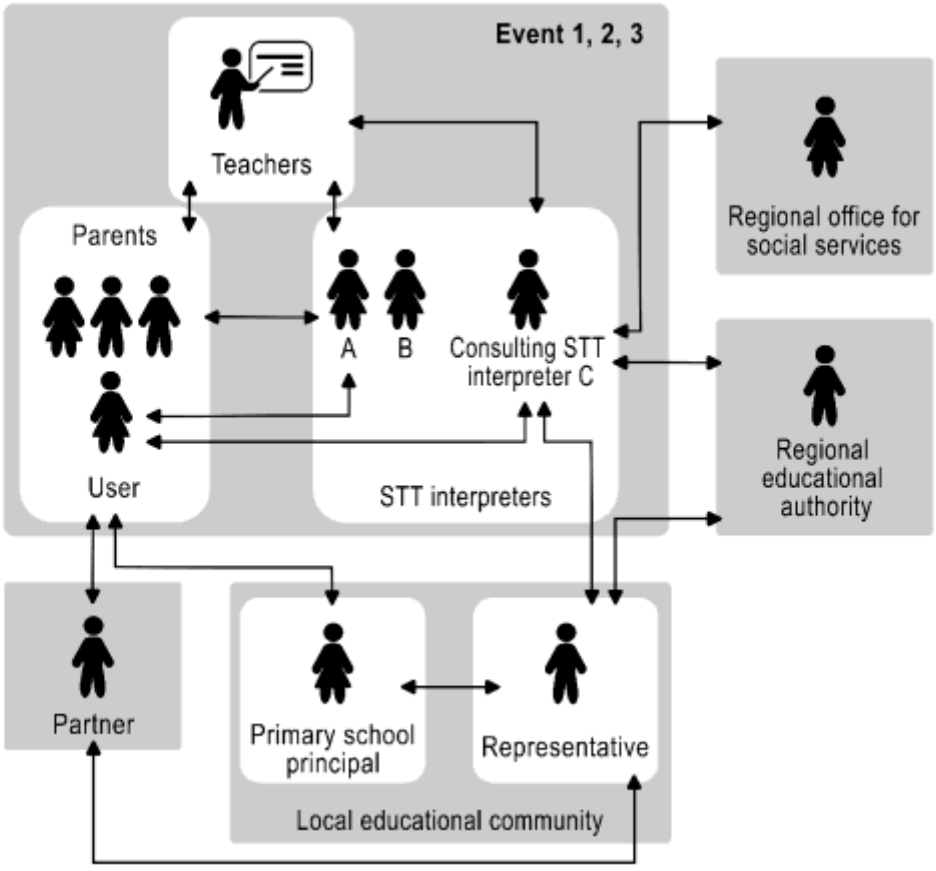

Figure 2: Interaction within the CI STTI network

3.2.2 Economic parameters within the CI STTI network

For STTI services in Austria, users with hearing impairment can ask for a yearly budget at the regional office for social services. This budget is assigned if the user has a hearing loss of more than 50 percent and can be used for events in the "private life sector", covering "important appointments in life and for life training" at schools, in kindergarten, or at the doctor's (Social services 2025). The user intended to use this private budget for two of the three events – the one at the kindergarten and the one at the after-school care facility, taking into account that for compulsory school activities, the local educational community has to bear the costs.

The regional office for social services pays fixed rates for STTI services: these were 33 euros for 0.5 hours of interpreting for the on-site working interpreters, as well as 31 euros per hour of travelling time compensation. Travel costs were reimbursed according to the kilometre allowance. The interpreter working online was paid 1.1 euros per minute of interpreting assignment, as well as 31 euros for technical setup. The payment scheme does not include any additional rates paid for organizational tasks, which in this case were rendered by the coordinating STTI interpreter (Social Ministry Service 2023: 2).

The local educational community was informed about the costs of the previous events by the user's partner and the consulting STT interpreter, adding up to about 500 euros for an event of 2 hours. Summing this information up for all planned primary school events including parents, such as parents' day, they decided that those costs were not affordable. They proposed to use an AI-based transcription tool provided by the smartboard software.

Although broadcasters are known for paying rather low rates, compared to the pre-established social services rates, these were still 67 % higher in the cases referred to in this paper.

3.3 Viability parameters for BC and CI STTI

For the BC STTI setting, viability fully depends on the contractors. If the one and only parameter that counts is price, human-provided STTI will not be viable at all. It seems that there are no quality assurance measures in the loop, nor has the user any possibility to demand an appropriate quality level that allows people with the need for live text, for whatever reason, to access the information given by a broadcast event – sports events, in our examples.

For the CI STTI setting, it happened that the user had a personal budget at her disposal to spend on STTI services. This personal budget was sufficient to cover the expenses for all three events. Although the costs for the event at the primary school had to be borne by the local educational community, the user was ready and willing to spend her personal budget to obtain human STTI services. Still, she had some concerns about the fact that the two authorities might not be cooperative in this regard. She messaged her concerns to the consulting STT interpreter:

> Can we be sure that the regional social office will be paying for the event at the primary school? I ask, because I do not want to get negatively surprised that at the end they do not bear the costs for it because the local educational community should have paid for it or they might argue that I had to accept the refusal of the local educational community and get along with the smartboard as this might be cheaper for the regional social service as well (personal communication, August 26 2025, our translation)

On the other hand, one of the network entities, in this case, the primary school represented by the local educational community, was dealing with STTI services for the first time. Based on the legal obligation to pay for a suitable STTI service, they calculated costs in this regard for the first time and decided that human STTI services were not affordable, arguing that an AI-based transcription smartboard would deliver sufficient quality. If this were not the case, the user should give feedback in this regard:

> She has to tell us if it works, but from our point of view, it should work (personal communication, August 13, 2025, our translation).

This approach proved to be very critical from the user perspective and put additional pressure on her:

> I am really willing to test the smartboard function in some other occasion, but as it is the first event of the school year, I do not want to risk not getting all the information which is shared amongst the parents. On the other hand, I do not want to stand out in a negative way, or bother the principal or the representative of the local educational community. My child has just started school, my younger child will follow in three years. It is important not to displease the important people. It might be simpler if the people making decisions are well disposed towards me as a mother. Unfortunately, in this case, we as a family have to consider the total package! (personal communication, August 26 2025, our translation)

The consulting STT interpreter argued that AI-based tools do not offer sufficient quality in a dialogic encounter involving around 30 people, with no human interpreter on site to intervene in the communication or the target text. Additionally, she asserted that the interpreting team had already been informed about the dates of the events and was available. She highlighted again that due to the importance of the event, the smartboard should be tested on some other occasion.

4 Conclusion

How sensitive is STTI in these specific settings to mere technological solutions? Cost effectiveness seems to become the main decisive factor. So, the first step can be cutting down the costs of human STTI.

In the BC STTI setting, an important power imbalance and an incomplete/interrupted information flow can be observed. For STT interpreters, it is almost impossible to get in contact with those who make the decisions, as on the one hand, the contact in most of the cases is established by a third party, such as a technical provider, and on the other hand, if there is direct contact with the broadcaster, it is, to all intents and purposes, impossible to get in contact with higher-level employees or officials. Any information and arguments, for or against, will not go further than to a lower employee or clerk level. Only in very exceptional cases will even valid reasons lead to an intended or expected kind of reaction from the broadcaster.

In the CI STTI setting, the cost-bearing entity proposed a free-of-cost AI-transcription tool included in the smartboard at the premises of the primary school. The authority did not consider any arguments provided by the consulting STT interpreter and the user herself that this tool might not correspond to the expectations of the user. They took for granted that their limited experience with the quality of the tool would be suitable for an event, where the school personnel, in this case the teachers, as well as the other participants in the communicative event, were to experience STTI for the first time.

It is only due to the commitment of the consulting STT interpreter and the willingness of the user that human STTI services were provided. For the payment for the service, those people assumed the risk of not getting effective payment by the regional social service and bore the financial risk of paying for the services on their own. Additionally, the consulting STT interpreter was willing to accept that some of the costs (for the STTI platform) were not covered, nor was the time which she employed for the preparation and the consulting activity.

So, another question arising is: which members of the interpreting network consider which service levels in opting for human or AI-based services? And, which, if any, quality parameters are taken into account?

In the BC STTI case, it seems that quality is barely taken into account, and it is more about fulfilling formal accessibility requirements, which are supposedly met by providing any kind of live text. In the case of the CI STTI setting, too, the most important factor might be that services

are considered “fit for purpose” in terms of the regulations of the social cost bearers. In both cases, however, the suitability of substituting interpreters with AI-based tools seems never to have been checked. Especially, this is true where AI-generated live captions for BC do not comply with the official guidelines for subtitles, such as requiring a specific duration of the subtitles on screen, to mention just one of the requirements. The AI-generated captions go past very quickly and are, therefore, illegible in many cases. For both settings, it applies that the AI-generated captions are very often neither legible nor comprehensible, as they are mere transcriptions, keeping the shortcomings of natural speech, such as redundancies, false starts, and incorrect grammar, amongst other things. Furthermore, missing or incorrect speaker identification and punctuation can be observed, negatively affecting both legibility and comprehensibility. Last but not least, AI is not, up to now, able to adapt the live text to the local culture where the live text is to be consumed.

In STTI services with their wide range of provision, more and more STTI professionals find themselves in network settings where they have to argue in favour of human services. They can no longer assume that within all entities in the network, the suitability and quality of products and processes is constantly checked by all and that every entity involved is willing to share responsibility for suitability and accessibility. From a financial point of view, STT interpreters are facing increasing “competition”, as other network entities might consider AI-based tools as a substitute for the human service provider with a bearable, or almost no, decrease in quality. Taking into account the practical significance of such project networks, further studies on user expectations and their role in the process of ensuring suitable quality are more than necessary. Additionally, the different elements and levels of service provision in the preparation, service provision and post-event phase that involve interpreters should be addressed in training efforts for future professionals; in that case, they have the opportunity to establish a more comprehensive list of arguments for ensuring the sustainability of their service and their profession in general.

Acknowledgment

The authors would like to thank Regina Rogl for her help with the graphic representations of the interpreting networks.

References

Alonso-Bacigalupe, L. & Romero-Fresco, P. (2024). Interlingual live subtitling: the crossroads between translation, interpreting and accessibility. *Universal access in the information society,* 23(2), 533–543

BFSG – Barrierefreiheitsstärkungsgesetz (Accessibility Enforcement Act). Retrieved 17/11/2025 from https://bfsg-gesetz.de/

BaFG – Barrierefreiheitsgesetz (Accessibility Act). Retrieved 17/11/2025 from https://www.ris.bka.gv.at/GeltendeFassung.wxe?Abfrage=Bundesnormen&Gesetzesnummer=20012316

BSD – Bundesverband der Schriftdolmetscher Deutschlands e.V. (German Speech-to-Text Interpreters’ Association) (2025). Retrieved 17/11/2025 from https://bsd-ev.org/taetigkeit/

Downie, J. (2020). *Interpreters vs machines : can interpreters survive in an AI-dominated world?* 1st ed. London, New York: Routledge, Taylor & Francis Group.

EAA – European Accessibility Act (2019). Retrieved 17/11/2025 from https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32019L0882

Eichmeyer-Hell, D. (2021). Speech recognition (Respeaking) vs. the Conventional Method (Keyboard): A quality-oriented comparison of speech-to-text interpreting techniques and addressee preferences. In: S. Jekat, S. Puhl, L. Carrer & A. Lintner (eds) *Proceedings of the 3rd Swiss Conference on Barrier-free Communication (BfC 2020)*. Winterthur (online), June 29–July 4, 2020. Winterthur: ZHAW Zurich University of Applied Sciences. DOI: 10.21256/zhaw-3001.

Eichmeyer-Hell, D. (forthcoming). Schriftdolmetschen – Realisierungsformen im qualitätsorientierten Vergleich. PhD Thesis. University of Vienna.

Gile, D. & Napier, J.M. (2020). Spoken Language Interpreters and Signed Language Interpreters: Towards Cross-fertilization. *International Journal of Interpreter Education*. 12: 6

Henderson, S. (2025). The quality of a lot of subtitles being broadcast right now is embarrassing. Retrieved 26.11.2025 from https://www.linkedin.com/posts/henderss05_the-quality-of-a-lot-of-subtitles-being-broadcast-activity-7371084953456975873-Wk7G?utm_source=share&utm_medium=member_desktop&rcm=ACoAABVDyM4BF3-GifOrNBfNl0Fs1xMPo-SyKr4

Hoza, J. (2022). Team interpreting. In: C. Stone, R. Adam, R. Müller de Quadros & C. Rathmann (eds) *The Routledge Handbook of Sign Language Translation and Interpreting* (pp. 162–177). 1st ed. New York: Routledge.

Kadrić, Mira, and Katia Iacono. 2023. Interpreting in a project network. Dependencies and interpreters' multidimensional alignment. In C. Zwischenberger, K. Reithofer & S. Rennert (eds) *Introducing new hypertexts on interpreting (studies)* (pp. 168–193). Amsterdam: John Benjamins.

Kutz, Wladimir (2010) Dolmetschkompetenz: Was muss der Dolmetscher wissen und können? *Translatio 1*. Berlin: European Universitätsverlag

Nogueira, T. C. (2022). The interpreting team. The integration of strategies used during the conference interpreting process in the booth. In: C. Stone, R. Adam, R. Müller de Quadros & C. Rathmann (eds) *The Routledge Handbook of Sign Language Translation and Interpreting* (pp. 261–276). 1st ed. New York: Routledge.

Norberg, U. & Stachl-Peier, U. (2018). Quality in speech-to-text interpreting. A study of condensation strategies. In S. Zupan & A. Nuč (eds) *Interpreting Studies at the Crossroads of Disciplines* (pp. 129–155). Berlin: Frank & Timme.

ÖSDV – Österreichischer Verband für Schriftdolmetschen (Austrian Speech-to-Text Interpreters' Association) (2025). Retrieved 17/11/2025 from https://www.oesdv.at/

Platter, J. (2025). Schriftdolmetschen als Teil barrierefreier Schwerpunkte in österreichischen Curricula – Fallbeispiel Zentrum für Translationswissenschaft. In M. Agnetta, A. Schmidhofer & A. Petrova. (eds.). *Bild – Ton – Sprachtransfer: Neue Perspektiven auf Audiovisuelle Translation und Media Accessibility.* (pp. 237–264). Berlin: Frank & Timme.

Platter, J., Iacono, K., & Zwischenberger, M.B. (2025). Do We Act the Same, Similarly or Differently? How Spoken-Language and Speech-To-Text Interpreters Collaborate Within an Extra-Curricular Mock-Conference Training Activity. In P. Šveda, M. Djovčoš & E. Perez (eds) *Confronting Digital Dilemmas in Translator and Interpreter Training* (pp. 237–257). London: Routledge, Taylor & Francis.

Pöchhacker, F. (2018). Media interpreting: From user expectations to audience comprehension. In F. Pöchhacker, Y. Gambier & E. Di Giovanni (eds) *Reception Studies and Audiovisual Translation* (pp. 253–276). London: John Benjamins.

Pražák, A., Loose, Z., Psutka, J., Radová, V. (2020). Live TV subtitling through respeaking with remote cutting-edge technology. *Multimedia tools and applications,* 79(1–2) (pp.1203–1220). New York: Springer US.

Risku, H., Rogl, R., & Pein-Weber, C. (2016). Mutual dependencies: centrality in translation networks. *JoSTrans: The Journal of Specialised Translation*, (25) 1–22.

Skaaden, H. 2021. 'Interpreter's mistake'. Why should other professions care about the professionalization of interpreters? In L. Gavioli & C. Wadensjö, *The Routledge Handbook of Public Service Interpreting* (pp. 261–276). London/New York: Routledge.

Social Ministry Service – Sozialministeriumservice (2023). *Unterstützungsangebote für schwerhörige und gehörlose Menschen.* Retrieved 17/11/2025 from www.sozialministeriumservice.gv.at%2FDownloads%2Fregelungen_-schriftdolmetsch-leistungen.docx.

Social Services Land Niederösterreich (2025). *Gebärdensprachdolmetschen – Förderung*. Retrieved 17/11/2025 from https://www.noe.gv.at/noe/Menschen_mit_Behinderung/Gebaerdensprachdolmetschen.html.

UNCRDP – United Nations Conventions on the Rights of Persons with Disabilities (2006). Retrieved 17/11/2025 from https://www.ohchr.org/en/instruments-mechanisms/instruments/convention-rights-persons-disabilities

Wadensjö, C. 1998. *Interpreting as Interaction.* London/New York: Routledge.