



This is a contribution from *Cultus: the Intercultural Journal of Mediation and Communication* 2019: 12

© **Iconesoft Edizioni Gruppo Radivo Holding**

This electronic file may not be altered in any way.

The author(s) of this article is /are permitted to use this PDF file to generate printed copies to be used by way of offprints, for their personal use only.

The Future of Mediators for Live Events: LTA Project - Academic and Vocational Training

Estella Oncins (TransMedia Catalonia - UAB, Spain)

Carlo Eugeni (SSML - Pisa, Italy)

Rocío Bernabé (SDI München - University of Applied
Languages, Germany)

Abstract

Languages and cultures are intimately related; especially in the age of Information Society, where Technology continuously gives rise to new levels of interaction. In this context, the traditional training of language professionals in Translation Studies is no longer in line with current social, and industry requirements. A common discourse among professional translators is the fact that translator training programs are “inefficient, misleading, too theoretical, and irremediably out of touch with market developments” (Pym, 2011: 6). Current market practices in the translation field are evolving and differentiating in terms of method (crowd-sourcing, relay, and live); working possibilities (in person and remote); distribution opportunities (from massive to individual) and roles (translator, interpreter, and linguistic and cultural mediator). Moreover, the disruption of automatic mediation processes clearly demands a fresh look at the training of future professionals, which is already highlighted by the EMT Expert Group (2009: 7). There is, then, a gap to be bridged in training, which requires training skills to be defined for professionals of language and cultural mediation (real-time intralingual reporters or subtitlers), whilst taking into account technical possibilities and industry requirements.

This article deals specifically with this topic. In particular, it will start from a definition of mediation as a key concept in Media Accessibility; it will then present the results of a survey regarding the skills required of real-time intralingual subtitlers and their role as mediators in several fields of everyday life; lastly, existing technologies in the field of re-speaking which are used to produce real-time intralingual subtitles on TV and standard live settings will be described, as well as the extent of their contribution to the domain of mediation.

1. Introduction

Since the emergence of technology in language-related activities, the traditional training of language professionals within Translation

Studies has been evolving in terms of method, working possibilities, distribution opportunities and roles. Both the study and practice of Translation and Interpretation can be adapted in response to market demands and technological possibilities. One example where this can be applied is in the field of the study and practice of audiovisual translation (AVT), which became established many years ago. An example of AVT among many others would be the ever-important area of Media Accessibility and subtitling for the Deaf and Hard-of-Hearing people (SDH). SDH has been evolving rapidly in recent decades due to new technological developments, such as the Velotype keyboard and ASR systems (Automatic Speech Recognition)¹. These two techniques have made real-time intralingual subtitling possible, bringing this area of AVT closer to established forms of translation and interpreting. Real-time intralingual subtitling involves skills that are closely linked to simultaneous interpreting. Given that interpreters can be easily trained in this, real-time intralingual subtitling can be considered an interesting new niche in Translation Studies (Eugeni, 2008a; Arumí Ribas and Romero-Fresco, 2008; Remael *et al.*, 2016; Eugeni and Bernabé forthcoming).

1.1. Mediation in Media Accessibility

The term ‘mediation’ in Translation and Interpreting Studies has been broadly explored in terms of practice. In this context, ‘the professional distinction is between “interpreting” as spoken mediation and “translation” as written’ mediation (Pym, 2018: 257). An important aspect of AVT in recent decades has been the pervasive nature of technology, which emerges not only as a tool, but also as an agent in the mediation process. According to Pym, ‘mediation, normally applies to situations where there is clearly a present third party: a translator, an interpreter, or perhaps an online machine translation service’ (2018: 257). In all three cases, the mediator becomes a meaning-making agent, whose role it is to create meaning for those who cannot understand what is happening. (Dendrinis, 2014: 143).

¹ Respeaking and Velotyping are the focus techniques of the EU co-funded LTA project, of which this paper is part of the required dissemination work (Ref.: 2018-1-DE01-KA203-004218). Although only these two techniques will be dealt with here, most of what is said about Velotyping also applies to Stenotyping, Palantyping and other techniques making use of a chord keyboard. For more information visit <http://ltaproject.eu> (last accessed 28/08/2019).

In the field of Media Accessibility, ‘mediation’ plays a key role; namely to grant access to information for people when they cannot access audiovisual content because of their disability, such as people who are Deaf and Hard-of-Hearing. This kind of mediation is summarised in the Common European Framework of Reference for Languages (CEFR, 2001: 14)² as follows:

In both the receptive and productive modes, the written and/or oral activities of mediation make communication possible between persons who are unable, for whatever reason, to communicate with each other directly. Translation or interpretation, a paraphrase, summary or record, provides for a third party a (re)formulation of a source text to which this third party does not have direct access.

In this sense, Media Accessibility in the specific field of subtitling for the Deaf and Hard-of-Hearing (SDH) could be considered as ‘intralingual mediation’, the act of interpreting in community-based settings (Pöschhacker, 2008). Intralingual mediation covers any activity that involves reformulation for an audience with specific needs and that share the same language. Part of this mediation process is the result of a linguistic reformulation from a linguistic structure to a different one (simpler, shorter and respecting the characteristics of written language), according to the end users’ need for a new recipient context (Eugeni, 2008b). In addition, other aspects such as the mediation skills of the real-time intralingual subtitler and the increasing use of technology will determine the mediation process.

1.2 Mediation and Deaf and Hard-of-Hearing audiences

Since the adoption of the UN Convention on the Rights of Persons with Disabilities (CRPD), all EU countries have committed to set minimum standards to implement such rights. In this sense, the recently revised Audiovisual Media Service Directive (AVMSD) and the new European Accessibility Act, both by the EU, establish the need to generate new accessibility services to all media: subtitles, audio subtitles, audio description and sign language in live and recorded video, for both linear and non-linear delivery. Accordingly, each country must develop national laws to transpose EU directives and

² Source: CEFR 2001. “Common European Framework of Reference for Languages: Learning, teaching, assessment).

Available online: <https://rm.coe.int/16802fc1bf> (last accessed 31/08/2019)

Acts. Due to the new international regulations and legislation in the Media Accessibility field, EU countries are being obliged to provide accessibility services to cater for the needs of user groups such as the Deaf and Hard-of-Hearing.

The main objective of subtitling for the Deaf and Hard-of-Hearing is to communicate audiovisual information directly to a specific audience. Given that the needs of such an audience, the Deaf and Hard-of-Hearing, who are part of a disabled community, are frequently ignored, it is not difficult to assert that real-time intralingual subtitling is a form of socio-cultural mediation, along with all other forms of Media Accessibility, such as audio description and sign language interpreting. Within this context, Katan (2009: 89-90) synthesizes the process of cultural mediation in a 'logical levels table' in which each level is determined by the context of culture and circumstances that need to be considered. The levels include environment, behaviour, strategies, values/beliefs, identity, role, and mission in society. Each of these levels bears particular significance in the work of real-time intralingual subtitlers. Environment applies to the need for the professionals to know the type of text to be transferred, the setting and the audience specificities; behaviour applies to 'what is to be translated?', which consists of the inclusion/transmission of the non-verbal elements in the target text, to ensure its adequate reception; specifically, in field of Deaf and Hard-of-Hearing accessibility. Values/beliefs apply to the intentions implicit in the Source Text and how to transfer them to the target text. Identity applies to the different agents involved in the communicative process, and therefore, to the needs and requirements of the speaker, the commissioner and the intended audience. The role applies to the meta-competences of the mediator's ability to provide an accessible service. The mission in society deals with the ethics and professional issues from which the task of the real-time intralingual subtitler comprises.

As can be seen, professionals in this field are expected to master specific skills and receive appropriate training that goes beyond the linguistic and cultural knowledge usually provided in current translation studies. In fact, additional mediation skills related to the setting and the needs of the end-users have to be taken into consideration.

1.3 Mediation in real-time intralingual subtitling

The skills used to carry out intralingual mediation in the real-time intralingual subtitling field have hardly been addressed, except for a few studies such as the volume on live subtitling through Respeaking by Romero-Fresco (2011) and that on Respeaking as a technique edited by Eugeni and Mack (2006). Most of these academic works focus on skills used in a TV broadcasting setting and/or describe real-time subtitling practices in specific countries, such as Spain (Orero, 2006) and Belgium (Remael and van der Veer, 2006). As a consequence, not much research can be found that validates the skills and competences applied to settings outside of TV.

In the real-time intralingual subtitling field, professionals are required to relay information from the Source Text to the Target Text for a given communicative purpose and work setting. These mediation skills can be described as competence-based (Cedefop, 2011)³, since real-time intralingual subtitlers have to use various skills and competences to carry out the mediation process. Specifically, when real-time intralingual subtitlers are asked to provide a service, they are in fact acting as mediators. Therefore, they have to understand the information included in the source text and then transfer it to the target text appropriately and accurately in a way that satisfies the requirements of the work setting. The practice of mediation thus entails the use of the source text in particular social contexts in ways that are based on certain social needs.

Real-time intralingual subtitlers as mediators, not only use their sociolinguistic knowledge and language awareness, but must also possess a background in accessibility, with knowledge on how to implement different strategies which will address specific users' needs, and in turn convey a specific meaning through a particular approach to the Target Text, such as verbatim versus sensatim (Eugeni 2008b). To be in a position to do this successfully, they need to activate their mediation awareness according to Katan's 'logical levels' (Katan, 2009) with regard to the specific (Source and/or Target) Text and their familiarity with the topic.

³ Source: Cedefop 2011. "Using learning outcomes - European Qualifications Framework Series: Note 4". Available online: https://www.cedefop.europa.eu/files/Using_learning_outcomes.pdf (last accessed 31/08/2019).

2. Mediation and Technology - impact on mediation services

Ahead of its predecessor, Human-Computer Interaction (HCI) was meant to be “the design, evaluation and implementation of interactive computing systems for human use” (SIGCHI 1992: 5). Nowadays, however, technology in many fields of mediation has superseded human use, by totally or significantly replacing it. In Machine Translation Studies, this is mirrored by Human-Aided Machine Translation, which is increasingly competing with the natural evolution of HCI, meaning Computer-Aided Machine Translation.

Similarly, in Natural Language Understanding, ASR is gradually automating Media Indexing, Reporting and the Subtitling of pre-recorded audiovisual material published on the web. As far as real-time subtitling through Respeaking is concerned, technology is pervasive to such an extent that each step in the process would be impossible without it. In some cases, HCI is so predominantly driven by technology that we can talk of Human-Aided, Real-Time Subtitling and even of Automatic Subtitling. In order to gain a more in-depth understanding of how this technology impacts on real-time intralingual subtitling, the next section will address the three main steps in the process of Respeaking (Source Text, Mid Text, Target Text) and later, concentrate on the nature of technological aspects and their impact on the final product.

2.1. The process of real-time subtitling

With regard to real-time subtitling through Respeaking, Marchionne (2013) shows that the process consists of nine steps that can be grouped into three main areas:

1. The Source Text
 - a. The speaker speaks into a microphone
 - b. The subtitler listens to the speaker, ideally through headphones in a booth or office at the same time as the rest of the audience
2. The Mid Text
 - a. The subtitler speaks into a microphone
 - b. The ASR software program transcribes what the respeaker says

-
- c. The editing software turns what the ASR software has produced into editable text
 3. The Target Text
 - a. The editor identifies, assesses, and corrects significant mistakes
 - b. The edited text is turned into subtitles
 - c. The subtitles are broadcasted
 - d. The audience reads the subtitles while watching the visual component of the setting, e.g. TV programme, conference or parliamentary session.

Even though this process describes the professional activity of the respeaker, it can be easily adapted to that of the velotypist. The only aspect which needs to be modified is the production of the Mid Text, which depends on the technique used. Specifically, in the case of Velotyping, the subtitler types a combination of keys simultaneously (2.a), which the Velotyping programme turns into written text (2.b), which the editing software turns into editable text.

2.2. Technological aspects

As seen, technology in real-time intralingual subtitling is so pervasive that HCI varies from Computer-Aided to Human-Aided to Fully Automated, real-time, intralingual subtitling; depending on factors such as the available resources (technological, human, logistic, and economic); the work setting (face-to-face, online, and by relay) and the context (cultural events, parliamentary sessions, media broadcasts, workplace, educational settings). Because of this, the needs and expectations of the end users may vary accordingly, from situation to situation, as the respondents explained in their answers to the questionnaire mentioned below.

As for the Source Text, technology has a moderate impact. However, there is a huge difference between listening directly to the voice of the speaker in a more or less noisy room; listening to it through headphones in the same room as the speaker or in a cabled soundproof booth; and listening to it via the web or a phone. What varies here is the quality of the reception: listening to the voice of the speaker through headphones in a soundproof booth is the best working condition because the subtitler can clearly listen to the speaker by adjusting the input volume; they can see speakers in their physical context and they are not stressed by background noise or prying eyes.

A similar situation is that of remote subtitling through online connection. However, in a face-to-face situation with no headphones available, the first and third conditions do not apply, and this may diminish the quality of the listening process, which is likely to affect the quality of the end result. In the case of a relay service, the second condition does not apply, thus adding stress to the overall task.

As for the Mid Text, its production is the one most influenced by technology which is, in this instance, related to the technique used to produce the subtitles (Velotyping or Respeaking). In the case of Velotyping, technology is limited to the keyboard. Hence, if the velotypist knows which combinations of keys to press, the final result will not need to be edited, whilst all of the following steps to produce the Mid Text (2.b, 2.c) as well as those to produce the Target Text (3.a, 3.b, 3c) will follow automatically without any further human intervention (except for contexts where subtitles are to be manually cued). The same cannot be said for Respeaking, where more variables can have an impact on the production of the Mid Text, such as the Respeaking mode and the microphone used. With respect to the Respeaking mode, subtitles can be produced either automatically from the voice of the speaker or indirectly via the respeaker. In the first case, step 2.a is skipped and the quality of the final input will depend on the quality of the Source Text. This means that the software can do a good job only if the source text is carefully constructed and without background noise, overlappings, hesitations, self-corrections, extra sounds, or unknown (to the machine) words (Lambourne *et al.*, 2004).

If the Source Text does not have such features, the quality of the Mid Text can be very poor. In this context, homophones and near homophones can generate extra problems, which may lead to the transcription of other words when compared to the ones anticipated. In the second case, the professional respeaker knows how to handle all the aspects above. However, 100% accuracy cannot be guaranteed because the ASR machine may incorrectly recognise a word (e.g. they are > their) or prioritise the recognition of a different word (e.g. but > Butt), thus leading to an incorrect Mid Text (Marsh, 2004). As for the microphone used to dictate the Target Text, differences will depend on the kind of microphone used. If the respeaker uses a professional recording device (condenser, dynamic, ribbon, USB, stereo), the quality of the Mid Text will only depend on the dictation skills of the respeaker and on their capacity to avoid possible mistakes. In the case that a lower-quality microphone is used (mic-in, headphones with incorporated microphone, stenomask), as is common in conference

subtitling; quality may also depend on the background noise. If the respeaker uses a stenomask, the input volume may be adjusted according to reverberation and the individual's voice particularities.

Regarding the production of the Target Text, the technology consists of Editing Software, Subtitling Software, and in some cases, Automatic Translation Software. Editing (subdivided into pre-editing, before starting working; peri-editing, while producing the Mid Text; and post-editing, after the machine releases the Mid Text) is still human-based, but its impact on the Target Text may vary considerably depending on who is in charge of the editing process (i.e., identifying mistakes, assessing their relevance and correcting them). Typically in TV subtitling, at least two people work on the editing separately (Orero, 2006; Remael and van der Veer, 2006).

However this condition largely depends on the country in which it is being carried out. In conference subtitling it is the subtitler who is normally entrusted with editing. In the case of TV subtitling, the first person (the real-time subtitler) will focus on both pre-editing (by adding all possible words that can be enunciated in the Source Text to the software dictionary) and peri-editing (i.e., avoiding dictating or typing words that are difficult to produce through the writing technique used). The second person (the real-time editor) will focus on post-editing; which will take more or less time depending on variables such as the accuracy of the Mid Text, the word-per-minute rate of both the Source Text and the Mid Text, the kind of mistakes produced and the type of editing software used. In particular, if the editor can select possible solutions the software automatically suggests, instead of cutting and re-typing (common in both Velotyping and Respeaking), or re-dictating (not applicable to Velotyping and less common in Respeaking), the quality of the Target Text will be higher. In the second case, the quality of the Target Text will essentially depend on the subtitler's multitasking capacity, i.e. their ability to simultaneously produce the subtitles, browse the Mid Text in search of meaningful mistakes, and correct them. In this case, the effort made may be a further factor impacting on the cognitive load (cf. Gile, 1985) and consequently on the quality of the Target Text. As for the Subtitling Software, the quality of the final product depends on the kind of solution adopted. Subtitles can be sent to the end screen in several ways. In conference subtitling, the editing software may be connected to the screen either through cable (and beamer or video mixer), through USB ports communicating via radio among them, or through a web-based or TV-based solution. Once again, differences

depend on the specific timing of the on-screen appearance of the subtitles. In the case of a cabled screen or of a screen connected via radio, subtitles appear immediately, at the same time as they appear on the subtitling software. The same does not apply in the case of a screen relying on a web-based solution or a TV-based solution. Here the problem is latency, which may vary from 1.5-3 seconds (Eugeni, 2008a), in the case of TV subtitling, up to 30 seconds, as is the case with web subtitling, which also depends on the device used to send and mix the subtitles with the video.

Since there is a lack of well-established and certified training (Eugeni and Bernabé, forthcoming) for the acquisition of the mediation competences needed to carry out real-time intralingual subtitling, section 3 will present the LTA project, which proposes mediation skills for a curriculum designed to train a new breed of real-time intralingual subtitlers. By doing so, it aims to both bridge the gap in the training of these professionals and align both academic and vocational training in the real-time intralingual subtitling field.

2. Mediation skills for two real-time intralingual profiles: respeakers and velotypists

When end users regard the overall performance of a real-time intralingual subtitler as poor or insufficient, it can be assumed that there is a mismatch between their abilities and the standards required by the work setting. According to the 2015 report of the European Federation of Hard-of-Hearing (EFHOH)⁴:

It is essential that quality is maintained alongside increasing quantity; otherwise the purpose of subtitling is lost. One way to ensure quality is maintained is to invest in the training of live subtitling translators and speech-to-text interpreters. Such investment will only increase the accessibility of subtitling. (EFHOH, 2015: 19).

⁴ Source: EFHOH. 2015. "State of subtitling access in EU - 2015 Report". Available online: <https://efhoh.org/wp-content/uploads/2017/01/EFHOH-State-of-Subtitling-2015-English.pdf> (last accessed 31/08/2019).

The LTA project aims to bridge this gap by providing a harmonised modular curriculum following the skills-card structure proposed by the European Certification and Qualification Agency (ECQA), in line with the concepts of competence-based training, inclusion and equality (Cedefop, 2011). It will also provide training materials and implementation routes for higher education and vocational institutions.

LTA gathers educational and non-educational partners (trainers, employers, service providers, an end user association, and a certification association). The consortium aims to provide stable standards that express measurable levels of training quality; and to reach, through certification, a sustainable recognition of the training, profession and profiles of respeakers and velotypists. The LTA curriculum focuses on two techniques to produce real-time intralingual subtitles: Respeaking and Velotyping. The reason is that these techniques cover a higher number of languages and working scenarios, and show a higher employability potential as they respond to the market demands of a balanced outcome regarding costs, accuracy, and language availability. Respeaking uses ASR software to transfer voice to text. Though it is currently the most used technique, end-users still criticise the uneven quality of the delivered text (Romero-Fresco, 2015). Velotype is an EU-developed keyboard that produces whole syllables or words with every keystroke, and trained velotypists can type at the speed of speech for extended periods in over 30 languages. They deliver superior results in noisy environments and in terms of high-quality orthographic typing, but its training is long and costly⁵ compared to Respeaking. According to Lambourne (2006), the training period for velotypists is 12 months, while respeakers training is 2-3 months. In terms of speed, velotypist can reach up to 140-180 wpm and respeakers up to 140-160 wpm. In both cases, accuracy can be 95-98% but it will mainly depend on the source language and the speaker speech rate.

The LTA project consists of five work packages, so-called Intellectual Outputs (IOs), the results of which will be presented here. The goals are two: first, to define the skills and competences required of the real-time intralingual subtitlers; and, second, to explore real-time intralingual subtitling in terms of both vocational training (provided by universities or companies) and professional roles (respeakers or

⁵ The Velotype keyboard costs 1500,- EUR while respeaking software costs about 699,- EUR (information retrieved on 12/12/2019).

velotypists). The results of IO1 are the basis for IO2 (curriculum design), IO3 (creation of learning materials), IO4 (assessment) and certification in IO5.

In IO1 the data collected was both quantitative and qualitative, and the chosen tool for the two online surveys was a questionnaire. The first survey sought to identify the skills needed in the labour market and those demanded by end-users. The second allowed LTA partners to explore current teaching practices. Both surveys began with a skills-card organised by competence areas using learning outcomes as specified by the European Certification and Qualification Agency (ECQA) guidelines. As identified in a comparative study by the European Centre for the Development of Vocational Training (Cedefop, 2016)⁶, learning outcome approaches are used in both vocational and higher-education qualification frameworks for almost all EU countries, and provide a means of transparency, transferability and recognition of training nationally, and throughout Europe.

This section presents the results of the skills survey. The questionnaire was restricted to eight languages and comprised three sections. The first section provided information about the purpose of the project (the goals of the survey) and included the terms of participation and consent. The second gathered demographic data from participants; and the third, data about the competences required of the professional. The two types of questions used were closed questions with a 4-Likert scale (0 – of no importance; 1 – of minor importance; 2 – important; 3 – very important) and multi-choice questions with a free-textbox at the end, which allowed LTA partners to gather meaningful qualitative data. The following sections disclose only the most significant results and focus on the conclusions.

3.1 Demographic data

A total of 121 respondents participated in the survey. The demographic data collected showed that their profile was conducive with the audience targeted by LTA: 57 professionals, 13 trainers, 13 prospective students, 17 providers, 29 end users, 20 other. Respondents under the group ‘Other’ were professionals or providers

⁶ Source: Cedefop 2016. “Application of learning outcomes approaches across Europe: a comparative study” Available online: <https://www.cedefop.europa.eu/en/publications-and-resources/publications/3074> (last accessed 31/08/2019)

of services that had not been mentioned explicitly, for instance, the provision of real-time subtitles at museums, galleries or literary festivals, the provision of technical assistance, and that of remote services. Participants came from Europe, Asia-Pacific, and the U.S. Most of the professionals are between 40 to 59 years old, are freelancers (72.5%), who work mainly in their mother tongue (81%). As for the educational background; 26% of the participants had a Master's degree followed by 24% with a Bachelor's degree. All of the participants had achieved the minimum of a high-school diploma. In terms of work experience, the sample profile shows that 28% of the participants have been working for 20 or more years in the field, whereas approximately 71% of trainers have been training for less than 10 years. 81% of the participants work only in their mother tongue, whereas 15% work in both their mother tongue and another language. In terms of preferred training technique, Velotype, scored 58%, which was much higher than the 2nd preference, Respeaking with only 16%. It should be mentioned that Velotype is mostly used in Dutch speaking countries and that the number of Dutch participants accounted for a higher proportion than those from other countries. This fact may have affected the training techniques preferences. In addition, the remaining 26% correspond to 'Other' using techniques such as Palantype, stenography or the use of QWERTY keyboards.

The demographic data gathered leads to the conclusion that interest in the topic reaches beyond European borders. The participants who work as professionals in the field are middle-aged (over 40 years old), have an educational degree (mainly higher education and undergraduate), work mainly as freelancers and in their mother tongue. The fact that the service has been provided for an average of 20 years while most trainers have been teaching for 0-9 years supports LTA starting point: training is lagging behind.

3.2 Competence areas

The following section sets out the LTA competence framework; expressed in terms of competences to be acquired or mastered as a real-time intralingual subtitler. In line with the European Qualifications Framework (2008), the LTA competence framework outlines which competences are to be achieved, acquired and mastered by the end of the training and the requirements of each given activity, all of which are expressed in a skills-card structure.

The LTA project competences are divided into six areas: Knowledge about accessibility, Linguistic competence, Entrepreneurship and service competence, IT competence, Respeaking competence and Velotyping competence. Each competence area displays a set of skills or knowledge that a real-time intralingual subtitler should acquire and master, which should in turn be expressed in learning outcomes.

All competence areas presented in the survey were presented in the same structure. The same sentence introduced the skills for each competence (“To deliver a good quality service, a real-time intralingual subtitler should be able to [...]”) followed by a specific competence. Respondents rated the skills in a 4-Likert scale from 0 to 3 points (0 – of no importance; 1 – of minor importance; 2 – important; 3 – very important). At the end of the series of questions comprising each competence area, respondents were able to add their comments in a free-text box.

Competence areas ranked as follows:

Linguistic competence: 2.6;

IT competence: 2.4;

Entrepreneurship and service competence: 2.3;

Respeaking: 2.3;

Velotyping: 2.3;

Knowledge about accessibility: 2.1.

It should also be pointed out that Velotype and Respeaking areas may have scored lower because of those respondents who simply do not know or use Respeaking and Velotyping. Also, the higher rating for Linguistic and IT competence areas was probably due to their transversality rather than representing a comment on technique. Knowledge about accessibility, with the lowest points, can also be considered a transversal competence because most of the defined skills apply throughout the other key competence areas, and may determine the real-time intralingual service, i.e. knowing the needs and preferences of the end-users of the service, being able to adjust the output to the working context or advise customers about how to best set-up an accessible working environment. All of these skills may determine the use of the technology and the Respeaking and/or Velotyping service.

The analysis of the qualitative data showed that some respondents relate the overall importance of certain competences to the type of

employment or working context. In this light, they categorised the competence areas such as Knowledge about accessibility, Linguistic competence and competence related to technique (Velotyping or Respeaking)⁷ as necessary for all professionals, whereas Entrepreneurship and service and IT competences are seen as principally important for freelancers. Considering that 72.5% of respondents were freelancers; it can be asserted that the latter competences are considered to be secondary, after having acquired the initial competences, and once the professional has entered the market. It should be highlighted that respondents did not pinpoint the need for additional competence areas.

3.2.1 Knowledge about accessibility competence

This competence area comprise nine skills expressed in Learning Outcomes⁷, relating to three categories:

- a) Basic concepts about accessibility;
- b) Target groups, needs and interaction;
- c) How accessibility is embedded in the environment.

The skills concern the ability to apply knowledge; to use know-how to complete tasks; and to solve problems. The first category details the ability of the real-time intralingual subtitler to explain concepts such as accessibility, disability, multimodality, universal design, type of hearing loss, and particularities of the target audiences. The second category details the ability to help facilitate communication for people with hearing loss, and to adapt the output to the work context. It also includes the cultural particularities of the Deaf, Hard-of-Hearing and Deaf-Blind community along with other end-users of the service. The last category details the ability to identify personal ways of raising awareness, of explaining their role and tasks as professionals, and the ability to advise customers.

The results obtained showed that the skill “adjust the output to the work setting” was rated the highest, with a score of 2.7. This skill refers to a high-level proficiency requiring the professional to be able to evaluate and manage a given working setting for the delivery of the real-time intralingual service, and propose appropriate solutions, i.e. provide basic tips for presenters such as recommending to speak

⁷ They are ‘statements of what an individual should know, understand and/or be able to do at the end of a learning process’ (Source: EQF 2008)

clearly, avoid speaking too fast and use simple language for the description of the visuals for Blind users. The skill with the lowest score (1.4) was Communicate in sign-language in basic situations. This might be due to the fact that sign language is not the mother tongue of all Deaf and Hard-of-Hearing people and also requires specific training.

With regard to data concerning quality, three main topics were identified: the need to clearly describe the role and tasks of the professional and how they differ from other profiles; the uneven provision of accessibility services across countries (Asia-Pacific versus Europe and the U.S.); and the need to establish quality standards for the profession.

3.2.2 Linguistic competence area

This competence area presented eight skills expressed in Learning Outcomes, relating to three categories:

- a) Functionality: accuracy, readability, legibility;
- b) How to cope with speech-related challenges (Exit strategies and ad hoc solutions);
- c) Strategies to acquire and develop specific thematic knowledge.

In this competence area, prospective trainees should be able to understand the concepts and indicators of accuracy, readability, and legibility, and apply them to their work. The ability to perform verbatim (word by word) and sensatim (meaning by meaning) subtitling is expected, as well as the ability to identify and implement exit strategies (e.g. reformulation) and to decide its pertinence according to the setting and the needs of the audience. This competence area also included skills concerning the ability to acquire and develop specific thematic knowledge.

Two skills shared the highest score in the scale responses with 2.8; namely the ability to 'Deliver the accuracy needed in each setting' and to 'Apply readability indicators to the transcribed text'. Both relate to functionality and demand high-level implementation skills, a highly proficient language level (e.g. correct grammar, spelling of basic and difficult words, of names, job-specific terminology, description of sound) and grounded knowledge about readability (e.g. indicating the name of the speaker or a switch of speakers, specifying when someone speaks unclearly or too fast). Finally, they demand autonomy, self-

confidence and methodological mastery in the live situation. The skill with the lowest score, 2.3, was the ability to ‘Deal with different cultural approaches to real-time intralingual subtitling’.

Free-text responses brought to light the different approaches taken towards accuracy (verbatim versus *sensatim*) and to the role of the professional in situations in which communication might be at stake due to poor technique, complex grammar or constraints emerging from the setting or the audiences’ language skills. Furthermore, the respondents’ answers concerning verbatim and *sensatim* show that this variable is mainly dependent on the constraints of the setting and preferences of the end-users. In addition, they pointed out that they might represent two different modalities that require separate training, and are provided by two different types of professionals.

3.2.3 Entrepreneurship and Service competence

This competence area contained ten skills expressed in Learning Outcomes relating to four categories: a) Management skills, b) Interpersonal skills, c) Stress management, and d) Business strategies.

These skills expressed the professionals’ ability to build and manage a business. For this reason, basic knowledge about customer management and portfolio skills were included. The interpersonal and personal skills aimed to develop abilities in interacting with customers and took into consideration other personal aspects such as resilience and stress management.

The skills with the highest score of 2.7 were ‘Respond to a customer’s inquiry or problem in a timely and effective manner’ and ‘Follow up on customer requests to ensure that accessibility service needs are met’. The skills with the lowest score, 2.0, were ‘Explain the role of personal resilience as an own critical skill’ and ‘Prepare himself/herself to conduct professional interviews with customers’.

Free-text answers provided further insight into these topics. Respondents pinpointed the need to follow a code of conduct and dress appropriately for the job. They also propose self-confidence and social competence as necessary personal skills. Some respondents specified that skills in this competence area are mainly relevant for professionals who work as freelancers. This might be due to the fact that this competence area has not been sufficiently addressed in academic training programs (Siegel and Wright 2015). However, in the European Commission (2018) proposal for a council recommendation

on key competences for lifelong learning⁸, Entrepreneurship has recently been included as one of the eight key competences, necessary to ensure resilience and the ability to adapt to change.

3.2.4 IT competence

This competence area included seven skills expressed in Learning Outcomes relating to the categories:

- a) How to set up the working environment;
- b) Input tools;
- c) Output tools.

These skills describe the ability to set up the working station autonomously and to foresee and solve problems instantaneously. This competence area takes into consideration the need to acquire knowledge and skills in both input and output tools.

The skills with the highest scores, 2.7, were ‘Set up the hardware and software for each type of trained setting’ and ‘Use the hardware and software for each type of trained setting and solve problems’. A closer look at both skills reveals that they are intimately related. The verbs used show a difference in proficiency level, since the latter demands that the professional be able not only to understand but also apply the technical know-how to deliver the service. The skills with the lowest scores, 2.3, were ‘Explain the advantages of Text on Top and similar software by the provision of the service’.

Free-text answers provided comments regarding the importance of the mastery of IT skills as a critical point that adds value to the real-time intralingual subtitler. In addition, the comments stressed the challenge of keeping up with technological developments available in the market, the need to understand the advantages and disadvantages of automatically created subtitles and the fact that the software itself

⁸ The European Framework for Key Competences (2018) defines eight ‘key competences’: 1) communication in mother tongue 2) communication in foreign languages 3) mathematical competence and basic competences in science and technology 4) digital competence 5) learning to learn 6) social and civic competences 7) sense of initiative and entrepreneurship 8) cultural awareness and expression. For more information on the European Framework for Key Competences see: <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52018SC0014&from=EN> (last accessed 31/08/2019).

needs to be accessible for people with disabilities. It should be highlighted that training in IT skills is already a key priority, which has also been highlighted in the EQF (2008) and is one of the eight key competences in the European Framework for Key Competences 2018.

3.2.5 Respeaking competence

This competence area displays 15 skills expressed in Learning Outcomes relating to four categories:

- a) Listening while speaking;
- b) Dictating words and punctuation;
- c) Articulatory skills;
- d) Editing while Respeaking;
- e) Factors of high performance.

All categories include high-level abilities that cover the full chain of the real-time intralingual subtitling service, from setup to evaluation. The theoretical knowledge acquired in the other competence areas, such as linguistic competence, is needed to produce subtitles by reformulating, editing and correcting the respoken text during the listening. At the same time, the professional is required to continuously train the software (Arumi and Romero-Fresco, 2008; Rемаel and van der Veer, 2006).

The skill with the highest score, 2.9, is “Communicate with good pronunciation”. The one with the lowest, 2.3, is “Implement non-verbal elements for each work context by applying different techniques” such as changing colours or font-size, or by inserting labels. Respeaking demands special attention to the pronunciation of proper nouns and specific terms, otherwise the software will not recognise them during the Respeaking process, and the Source Text will not be accurately delivered to the end-users. This will cause a major problem in the quality of the service. On the other hand, non-verbal elements will mainly depend on the working context and available technologies.

As mentioned in the free-text responses, participants support the demand for good pronunciation and specify that the type and scope of the reformulation and editing depend on the constraints of the context (TV subtitles, live, other), the speed of speech, the ability to apply exit strategies and the overall goal of avoiding content loss.

Speed rates provided by the respondents range from 106 words per minute (wpm) to 400 wpm; this difference will depend on the speed rate of the speaker, the speed of the respoken subtitles and the amount of editing carried out (Romero-Fresco, 2009). Participants also subordinate this parameter to the quality of the output, the context, and the type of output (verbatim or sensatim). An accuracy of 99% is mentioned twice and some participants established a direct relationship between the accuracy of the subtitles and the delay. Delay remains a major challenge in real-time intralingual subtitles, because 'as a product, it entails the production of non-synchronous subtitles (there is usually a 3-4 second delay)' (Romero-Fresco, 2009: 118). In addition, there are technological issues that pose a major challenge, i.e. scrolling subtitles (more difficult to read) present a delay of 3-4 seconds and block subtitles present a delay of 4-6-seconds (Romero-Fresco, 2009).

3.2.6 Velotyping competence

This competence area displays 15 skills expressed in Learning Outcomes relating to four categories:

- a) Listening while typing;
- b) Typing words and punctuation;
- c) Typing skills;
- d) Editing while typing;
- e) Factors of high performance.

Two skills obtained the highest score of 2.8: 'Remember full sentences while lagging behind because of editing/quick speech/other difficulties', and 'Identify own typing mistakes during typing, decide how relevant these are (minor, major or critical) and correct them, if necessary'. The skill with the lowest score, 2.3, was 'Discipline themselves to practice and improve their skills daily, to reach a higher speed and accuracy'.

Speed, accuracy, delay, and interpreting strategies are also topics arising from the free-text responses. Answers show that they are interconnected. Speed rates ranging from 420 to 500 characters per minute (cpm) should be achieved, and a 99-100% spelling accuracy should be the goal. However; respondents categorised both parameters, typing speed and spelling accuracy, as subordinate to the overall goals of minimising delay and applying higher level interpreting

strategies. As can be observed, Respeaking and Velotype are highly connected in terms of interpreting skills. The main difference remains in the fact that while Respeaking demands a high command of pronunciation, Velotype demands a high command of spelling accuracy.

The free-text answers present two new aspects: first, the need for clear standards to measure the precision of the output and to elect a suitable model for the parameters in use; second, the appropriateness of the technique for specific contexts given that the sound of the keyboard during Velotyping must be considered. Regarding standards to measure precision, it should be mentioned that different models have been developed in recent years to measure real-time subtitling quality, especially in the TV context. Some examples are the WER model (Dumouchel *et al.*, 2011), which has been used as the base for developing the NER model (Romero-Fresco and Martínez 2015), asking for 98% accuracy; and the IRA model (Eugeni, 2017), where the smallest unit of analysis is not the word but the concept. The minimum accuracy required here is 95%.

While the NER model has been criticized, it is still the most widely known, especially in TV work contexts because it can be automatically applied through a specific software called NERstar (Matamala *et al.*, 2017). Even if some survey respondents claim that it is not always easy to apply, it still meets the demand for quantitative data when live, verbatim subtitling is required (Eugeni, 2017). In general, the overall quality of real-time intralingual subtitling is improving, thanks especially to the technological improvements in speech recognition and steno-captioning technology. Still, there is no single model to measure quality, which would serve all settings and work contexts. Moreover, as mentioned in the white paper “Caption Quality: International approaches to standards and measurement” by Media Access Australia, presenting different models used in different countries to measure quality, “not all errors are of the same magnitude. Some are negligible, while a single incorrect word can make an entire news story incomprehensible or, worse, change its meaning completely” (Media Access Australia, 2014: 9). From this point of view, the IRA model seems to solve this issue by focusing on concepts instead of focusing on words. However, it is quite difficult to apply because the segmentation of the source text into units of analysis and its assessment are completely subjective to the evaluator, though some indications are provided (Romero-Fresco and Eugeni, forthcoming). Clearly, much work remains to be done in this area.

4. Conclusions

As we have seen, mediation skills in the real-time intralingual subtitling process are transversal, and consequently, go beyond message transfer. The LTA project aims to fill a gap in the training of real-time intralingual subtitling professionals who are capable of taking on this emerging professional activity. The proposed competence-based training areas serve to foster not only the acquisition of skills related to the linguistic and cultural aspects of the mediation process; but also other relevant competences, some of which are traditionally linked to vocational practices, such as entrepreneurship and service competence, which have long been overlooked in traditional training programs and are now being considered as key competences in the European Framework for Lifelong Learning (2018). In addition, these professionals need to acquire and develop competences related to accessibility, technology and Respeaking and/or Velotyping techniques.

Worth mentioning are also the continuous technological developments that are improving automatic and computer-assisted transcriptions. These developments could be considered a threat to those working in the industry, given that part of the human task is being replaced. However, editing remains human, and the real-time intralingual subtitler's competences are still fundamental in many contexts of both verbatim subtitling (especially when the quality of the source text is poor) and sensatim subtitling, where Artificial Intelligence has not yet been developed for the production of accessible subtitles.

Professionals in the field have been able to bridge the gap between the Deaf and Hard-of-Hearing and the audiovisual content in different settings such as cultural events, parliamentary assemblies, media broadcast, workplace, and education to name a few. The debate over the role they perform in each setting and the traditional role assigned to translators/interpreters and mediators is still open. There is no official recognition for the high demand for the work of real-time intralingual subtitlers and professionals in the Media Accessibility field, nor for the fact that training is still lacking.

Acknowledgments

This work has been co-funded by the Erasmus+ Programme of the European Union, under agreement n° 2018-1-DE01-KA203-004218 (LTA project) and TransMedia Catalonia, a research group funded by “Secretariad Universitats i Recerca del Departament d’Empresa i Coneixement de la Generalitat de Catalunya” (2017SGR113).

References

- Arumí Ribas, M. and Romero-Fresco, P. 2008. “A Practical Proposal for the Training of Respeakers”, *JoSTrans - The Journal of Specialised Translation*, 10.
- Dendrinos, B. 2014. “Testing and teaching mediation: Input from KPG exams in English”. In *Directions in Language Teaching and Testing*, 1, 142-179.
- Eugeni, C. and Mack, G. 2006. (eds) *Intralinea, Special Issue: Respeaking*. Retrieved from <http://www.intralinea.org/specials/respeaking> (last accessed 10/12/2019).
- Eugeni, C. 2008a. “A sociolinguistic approach to real-time subtitling. Respeaking vs. shadowing and simultaneous interpreting”. In C. J. Kellett Bidoli and E. Ochse (Eds.), *English in International Deaf Communication*, Bern: Peter Lang, 357-382.
- Eugeni, C. 2008b. “Respeaking the TV for the Deaf: For a Real Special Needs-Oriented Subtitling”. *Studies in English Language and Literature*, 21
- Eugeni, C. 2017. “La sottotitolazione intralinguistica automatica”. *CoMe*, 2, 1, 102-113
- Eugeni, C. and Bernabé, R. forthcoming. “LTA project - Bridging the gap between training and the profession in real-time intralingual subtitling”. *Linguistica Antverpiensia, New Series: Themes in Translation Studies*, 14.
- Gile, D. 1985. “Le Modèle d’Efforts et l’équilibre d’interprétation en interprétationsimultanée”. *Meta* 30, 1.
- Katan, D. 2009. “Translation as intercultural communication”. *The Routledge companion to translation studies*. London: Routledge, 74-92.
- Lambourne, A.; Hewitt, J.; Lyon, C. and Warren, S. 2004. “Speech-Based Real-Time Subtitling Services”. *International Journal of Speech Technology*, 7, 4, 269-280.

-
- Lambourne, P. 2006. "Subtitle Respeaking". In Eugeni, C. and Mack, G. 2006. (eds.) *Intralinea, Special Issue: Respeaking*. Retrieved from http://www.intralinea.org/specials/article/Subtitle_respeaking (last accessed 10/12/2019).
- Marchionne, F. 2013. "Aspetti teorici del rispeakeraggio televisivo". In Eugeni, C. and L. Zambelli (eds.) *Respeaking*, Specializzazione online, pp. 74-81. Retrieved from <http://www.respeakingonair.org/it/component/banners/click/15> (last accessed 10/12/2019).
- Marsh, A. 2004. *Simultaneous interpreting and respeaking: a comparison*. Unpublished MA thesis. University of Westminster.
- Matamala, A., Romero-Fresco, P. and Daniluk, L. 2017. "The Use of Respeaking for the Transcription of Non-Fictional Genres: An Exploratory Study". in *TRAlinea Volumes*. Retrieved from <http://www.intralinea.org/specials/article/2262> (last accessed 10/12/2019).
- Orero, P. 2006. "Real-time subtitling in Spain, an overview". In Eugeni, C. and Mack, G. 2006. (eds.) *Intralinea, Special Issue: Respeaking*. Retrieved from <http://www.intralinea.org/specials/article/1689> (last accessed 10/12/2019).
- Pöchhacker, F. 2008. "Interpreting as Mediation". In C. Valero-Garces and A. Martin (Eds.), *Crossing Borders in Community Interpreting, Definitions and Dilemmas* Amsterdam - Philadelphia: John Benjamins, 9-26.
- Pym, A. 2018. "Introduction: Why Mediation Strategies Are Important". *Language Problems and Language Planning*, 42, 3, 255-266.
- Pym, A. 2011. "Training translators". In Malmkjaer, K. and Windle, K. (eds.), *The Oxford handbook of translation studies*. Oxford: Oxford University Press, 475-489.
- Remael, A. and van der Veer, B. 2006. "Real-time subtitling in Flanders: needs and teaching". In Eugeni, C. and Mack, G. 2006. (eds.) *Intralinea, Special Issue: Respeaking*. Retrieved from <http://www.intralinea.org/specials/article/1702> (last accessed 10/12/2019).
- Remael, A., Reviere, N. and Vandekerckhove, R. 2016. "From Translation Studies and audiovisual translation to media accessibility. Some research trends". *Target: International journal of translation studies., Special Issue: Audiovisual Translation: Theoretical and methodological challenges*, 28, 2, 248-260.

- Romero-Fresco, P. 2009. "More Haste Less Speed: Edited vs. Verbatim Respeaking", *Vigo International Journal of Applied Linguistics (VLAL)*,6, 109–133.
- Romero-Fresco, P. 2011. *Subtitling through speech recognition: Respeaking*. Manchester/Kinderhook: St Jerome.
- Romero-Fresco, P. 2015. *The Reception of Subtitles for the Deaf and Hard of Hearing in Europe*. Bern:Peter Lang.
- Romero-Fresco, P and Martínez, J. 2015."Accuracy rate in live subtitling: The NER model". In J. Díaz Cintas and R. Baños (Eds.), *Audiovisual translation in a global context: Mapping an ever-changing landscape*. Basinkgstoke: Palgrave Macmillan, 28-50.
- Romero-Fresco, P. and Eugeni, C. forthcoming. "Live subtitling through Respeaking", in Bogucki, Ł. and Deckert, M. (Eds.), *Handbook of Audiovisual Translation and Media Accessibility*, Palgrave: London.
- SIGCHI (Group: U.S.). 1992. *ACM SIGCHI curricula for human-computer interaction*. New York: Association for Computing Machinery.
- Siegel, D. S. and Wright, M. 2015. "Academic Entrepreneurship: Time for a Rethink?", *British Journal of Management*, 1-14.