

Romero-Fresco, P., Melchor-Couto, S., Dawson, H., Moores, Z. & Pedregosa, I. (2019). Respeaking certification: Bringing together training, research and practice. *Linguistica Antverpiensia, New Series: Themes in Translation Studies*, 18, 216–236.

Respeaking certification: Bringing together training, research and practice¹

Pablo Romero-Fresco

Universidade de Vigo, University of Roehampton
P.Romero-Fresco@roehampton.ac.uk and promero@uvigo.es
<https://orcid.org/0000-0003-2166-5792>

Sabela Melchor-Couto

University of Roehampton
smelchorcouto@gmail.com
<https://orcid.org/0000-0002-3867-1892>

Hayley Dawson

University of Roehampton
dawsonh@roehampton.ac.uk
<https://orcid.org/0000-0001-7156-1233>

Zoe Moores

University of Roehampton
Z.Moores@roehampton.ac.uk
<https://orcid.org/0000-0001-9876-8795>

Inma Pedregosa

University of Roehampton
Inma.Pedregosa@roehampton.ac.uk

Abstract

Research and training in respeaking are still lagging behind professional practice. One of the consequences of this lack of training opportunities is the UK government's refusal, in 2016, to use the Disabled Students' Allowances (DSA) to provide for respoken subtitles, arguing that respeaking was not a qualified profession. In order to tackle this issue, the Galician Observatory for Media Accessibility set up LiRICS, the Live Respeaking International Certification Standard, which aims to set and maintain high international standards in the respeaking profession. In 2019, after assessing the online certification process proposed by LiRICS, the Department of Education in the UK concluded that it meets their requirements and that LiRICS-certified respeakers are eligible for Disabled Students' Allowances funding. This article outlines, first, the current provision of respeaking training around the world and the assessments of live subtitling quality carried out to date, both of which inform the LiRICS online certification process presented here. The focus is then placed on the actual certification process, including a description of the tests, the platform used and the quality assurance

Romero-Fresco, P., Melchor-Couto, S., Dawson, H., Moores, Z. & Pedregosa, I. (2019). Respeaking certification: Bringing together training, research and practice. *Linguistica Antverpiensia, New Series: Themes in Translation Studies*, 18, 216–236.

process. This is followed by an analysis of the respeakers' performance, which has been shown to be in line with current professional standards.

Keywords: certification, LiRICS, NER model, quality, respeaking, training.

1. Introduction

Widely held as one of the most challenging modalities in media accessibility, intralingual live subtitling (or “real-time captioning”, as it is known in the United States) is defined by the International Telecommunication Union (ITU) (2015) as “the real-time transcription of spoken words, sound effects, relevant musical cues, and other relevant audio information” (p. 2) to enable deaf or hard-of-hearing persons to follow a live audiovisual programme. Live subtitles may be produced through different methods, including standard keyboards, dual keyboards, Velotype and the two most common approaches, namely, stenography and respeaking (Lambourne, 2006). Stenography uses a system of machine shorthand in which letters or groups of letters phonetically represent syllables, words, phrases and punctuation marks. This is the preferred method to produce live subtitles in the United States and Canada, and it is also used in other countries such as Italy and Spain, especially for transcriptions and subtitles in courtrooms, classrooms, meetings and other settings. Respeaking (known as “real-time voice writing” in the United States) is currently the preferred method for live subtitling around the world, especially for live subtitles on TV (Romero-Fresco & Eugeni, in press). It refers to the production of written text (such as subtitles) by means of speech recognition and may be defined as

a technique in which a respeaker listens to the original sound of a (live) programme or event and respeaks it, including punctuation marks and some specific features for the deaf and hard-of-hearing audience, to a speech-recognition software, which turns the recognised utterances into subtitles displayed on the screen with the shortest possible delay. (Romero-Fresco, 2011, p. 1)

Despite the popularity and widespread use of respeaking in the media accessibility industry and the impact it has on millions of viewers, research in this area is still hard to come by. The translation and interpreting database BITRA shows that only 4% of the academic publications on accessibility and 0.8% of published outputs on AVT, respectively, deal with live subtitling and respeaking. This is in stark contrast with research on audio description, for instance. Although it is much less widespread in the industry than live subtitling (Rossignol-Farjon & Cimino, 2016) and it is received by a smaller number of users, it features twice as many publications in BITRA. One of the reasons for this lack of scholarly activity in respeaking (and live subtitling in general) may lie in the very few respeaking courses available at universities (no more than a handful around the world) and in the very few scholars working in this area (Romero-Fresco, 2018). In addition, respeaking is often taught as a component within larger modules on subtitling for the deaf and hard of hearing (Romero-Fresco, 2012b). The respeaking part is typically small and often taught at the end of the academic year, when students have already chosen the research topic for their dissertations.

Romero-Fresco, P., Melchor-Couto, S., Dawson, H., Moores, Z. & Pedregosa, I. (2019). Respeaking certification: Bringing together training, research and practice. *Linguistica Antverpiensia, New Series: Themes in Translation Studies*, 18, 216–236.

The lack of academic research and training means that scholars have had very little influence on the training and working conditions of professional respeakers and many subtitling companies have no choice but to set up their own in-house respeaking programmes (Robert, Schrijver, & Diels, 2020). As a further consequence, the UK government, which was using the Disabled Students' Allowances (DSA) to fund the provision of live subtitles to provide access to school and university lectures for deaf students, decided in 2016 not to use these funds for respeaking. The argument provided for this decision was that there was little recognized training and, most importantly, no professional certification of respeakers (J. Ward, AiMedia, personal communication, September 18, 2019).

In order to tackle this issue, the Galician Observatory for Media Accessibility (GALMA), a research centre at the Universidade de Vigo concerned with the analysis of media accessibility quality in Galicia and at an international level, set up LiRICS, the Live Respeaking International Certification Standard. As part of its international activity, GALMA provides consultancy services to companies (Netflix, Sub-ti, AiMedia, etc.), broadcasters (Sky, TVE, VRT, etc.) and government regulators in Australia, Canada and the United Kingdom. The aim of LiRICS is to help set and maintain high international standards in the respeaking profession in order to create a pool of respeakers who can provide high-quality access through subtitles for live TV programmes and live events. In September 2019, the Canadian Radio-television and Telecommunications approved LiRICS (jointly with the Canadian company Keeble Media Inc.) as the official certification body to assess live subtitling quality on TV (CRTC, 2019). In the same month, the Department of Education in the United Kingdom assessed the online certification process proposed by LiRICS and concluded that it meets their requirements and that LiRICS-certified respeakers are now eligible for Disabled Students' Allowances funding:

We fully support this venture and can confirm that respeakers who pass the LiRICS certification have the required skill set to provide live remote captioning to students with hearing impairment. The tests, the platform and the quality assurance process used and undertaken have been devised thoroughly by an international team of expert researchers and professionals that have pioneered the development of live respeaking. This certification is the product of years of action research involving testing, user's feedback and continuously evolving modelling. The Department of Education's DSA funding is thus available for LiRICS-certified respeakers. (P. Higgs, Student Finance Directorate, Department for Education, United Kingdom, personal communication, September 20, 2019)

Before presenting LiRICS, this article outlined the current provision of respeaking training around the world and the assessments of live subtitling quality carried out so far, both of which inform the certification process presented here. The following section deals with the actual certification process, including a description of the tests and the platform used and the quality assurance process. This section is followed by an analysis of the respeakers' views and performance.

2. Respeaking training around the world

Live subtitling was first introduced in Europe in the early 1980s, when the British channel ITV began to subtitle headlines of public events using a standard keyboard (Lambourne, 2006). Since the subtitles produced with this method were not fast enough to keep up with the

Romero-Fresco, P., Melchor-Couto, S., Dawson, H., Moores, Z. & Pedregosa, I. (2019). Respeaking certification: Bringing together training, research and practice. *Linguistica Antverpiensia, New Series: Themes in Translation Studies*, 18, 216–236.

speech rates of many live programmes, other approaches were tested, such as the Velotype (a syllabic keyboard developed in the Netherlands), tandem methods involving from two to five subtitlers who would share the workload in a given programme and, in the 1990s, stenography (Lambourne, 2006; Orero, 2006; Romero-Fresco, 2011). Steno-made subtitles are fast and reliable, but also expensive, as the training required to become a competent live subtitler using this method takes between three and four years (Romero-Fresco, 2018). Respeaking was first tested as an alternative method for producing live subtitles in Europe in 2001, both by VRT (Vlaamse Radio- en Televisieomroeporganisatie), the national public service broadcaster in Flanders, Belgium, and by the BBC in the United Kingdom. Other European countries such as Spain, France and Italy followed suit in 2004, 2007 and 2008, respectively, helping to consolidate respeaking as the prevailing live subtitling method in Europe (Romero-Fresco, 2011).

The provision of academic training and scholarly research in this area did not start until 2006 (Eugeni & Mack, 2006). In 2007, the University of Antwerp created the first postgraduate course for respeakers as part of its MA in Interpreting and, following the first pedagogical proposals on respeaking (Arumí Ribas & Romero-Fresco, 2008; Russello, 2010), the Universitat Autònoma de Barcelona and the University of Roehampton, London, launched their own courses in 2008 as part of their respective MAs in Audiovisual Translation. By 2011, respeaking was taught in the following higher-education institutions (HEIs): University of Antwerp (in Dutch), Leeds University (in English), Universitat Autònoma de Barcelona (in Spanish), Universidade de Vigo (in Spanish), University of Bologna-Forlì (in Italian) and University of Roehampton (in English, Spanish, French, German and Italian) (Romero-Fresco, 2012b). Since then, courses on AVT have mushroomed all over the continent and respeaking has become not only the preferred method to produce live subtitles (Romero-Fresco & Eugeni, in press), but also a professional area that is in constant need of new professionals (Robert et al., 2020). However, training in respeaking has been incorporated in only a handful of other HEIs (SDI Munich, University of Warsaw and the European University of Valencia) and it has been reduced at the Universitat Autònoma de Barcelona.

Several reasons may account for this, but perhaps the two main ones are the high cost of the resources (software and equipment) required for respeaking and especially the limited number of respeaking trainers available. In fact, some of the abovementioned higher-education (HE) respeaking courses are taught by the same trainers, who are also often asked to create other bespoke vocational courses. Some examples of the latter are the courses delivered by Romero-Fresco and Melchor-Couto at MacQuarie University (Australia) on pre-recorded respeaking, the University of Helsinki and the national Finnish broadcaster YLE on live respeaking for TV, the University of Vigo on interlingual English–Spanish respeaking or the Galician TV station TVG for respeaking and post-edition in Galician.

The limited offering of respeaking training in HE is also at the core of two recent EU-funded Erasmus+ projects: LTA² (Live Text Access) and ILSA³ (Interlingual Live Subtitling for Access), which aim to create open source and flexible training materials for intralingual and interlingual respeakers, respectively. The first findings produced by the ILSA project (Robert et al., 2020) are the results of the largest questionnaire on respeaking training and practice disseminated so far among professional respeakers. This questionnaire has proved useful not only in

Romero-Fresco, P., Melchor-Couto, S., Dawson, H., Moores, Z. & Pedregosa, I. (2019). Respeaking certification: Bringing together training, research and practice. *Linguistica Antverpiensia, New Series: Themes in Translation Studies*, 18, 216–236.

providing a clear picture of the current landscape in respeaking training but also in informing the certification of respeakers described and analysed in this article. The questionnaire was filled in by 126 participants from 27 countries, including European countries but also Australia, Brazil, Canada, China, India, Iran, Korea, Malaysia and South Africa. Only a minority of respondents (13%) were trained at university, normally as part of face-to-face postgraduate courses on Audiovisual Translation or Interpreting. Respeaking courses range from one hour per week during eight weeks to two hours a week during 28 weeks. They are eminently practical, although they are often introduced by one or two theoretical units. They also include sessions on familiarization with the speech-recognition software, the creation of a voice profile, dictation practice and respeaking practice covering different audiovisual genres (from slow speeches to sports broadcasts and more challenging programmes such as news reports, interviews and chat shows). In some cases, trainees are also taught how to correct their mistakes live.

The majority of the respondents (87%) were trained (partially or fully) in-house, which differs from HE training in several respects. First, most in-house trainees are asked to take an aptitude test (normally as part of the selection process prior to employment), which may consist of a language test and a respeaking or a dictation test. Depending on the company, the training is organized either as on-the-job training led by colleagues without a real course or structure or as longer courses (from one intensive week to three months) focused on subtitling, speech recognition and respeaking practice. This is similar to the training offered by HEIs, although arguably less structured and more focused on the particular subtitling software used for the production of respoken subtitles. Most of the participants were assessed through continuous assessment using the NER model (Romero-Fresco & Martínez, 2015), which calculates the accuracy rate of the respoken text based on the number of words (N), edition errors (E) and recognition errors (R).

This brief overview of the respeaking training landscape points to a few issues that are relevant for the purposes of this article. First, there is a considerable contrast between the increasing scope and impact of respeaking in the AVT industry and the limited training offering at HEIs, which has barely increased over the past decade. Training is mostly delivered in-house and varies considerably across companies, which explains the existence of the two abovementioned EU-funded projects aimed at producing streamlined respeaking training material to be used in industry and HEIs, the UK government's decision not to use Disabled Students' Allowances unless professional respeakers are recognized/certified and, consequently, the creation of the LiRICS certification. Secondly, the analysis of the different components included in the HE and in-house training courses as described by the respondents of the ILSA questionnaire reveals some of the key competences that professional respeakers acquire through training and that can be tested in the certification process, such as preparation, dictation, edition, respeaking different genres of television and applying corrections. Finally, the questionnaire also shows that professional respeakers are often familiar with the NER model, which makes it a useful tool to use in the assessment of respeakers' performance for LiRICS.

Romero-Fresco, P., Melchor-Couto, S., Dawson, H., Moores, Z. & Pedregosa, I. (2019). Respeaking certification: Bringing together training, research and practice. *Linguistica Antverpiensia, New Series: Themes in Translation Studies*, 18, 216–236.

3. Research on quality assessment in respeaking

Research on respeaking, scarce when compared to the sheer volume of this technique in the industry, has so far focused mainly on the respeaking process (Baaring, 2006; Chen, 2006; de Serriis, 2006; Eugeni & Mack, 2006; Lambourne, 2006; Mack, 2006; Marsh, 2006; Romero-Fresco, 2008), the training of respeakers (Arumí Ribas & Romero-Fresco, 2008; Muzii, 2006; Remael & van der Veer, 2006; Romero-Fresco, 2012; Russello, 2010), the analysis of live or respoken subtitles (Bortone, 2015; Eugeni, 2009; García Romero, 2015; Jensema, McCann, & Ramsey, 1996; Luyck, Delbeke, Van Waes, Leijten, & Remael, 2013; Romero-Fresco, 2009; Romero-Fresco, 2016), their reception by the users (Eugeni, 2008; Muller, 2015; Rajendran, Duchowski, Orero, Martínez, & Romero-Fresco, 2013; Romero-Fresco, 2010; 2011; 2012a) and, finally, the application of respeaking for other purposes, such as transcription (Al-Aynati & Chorneyko, 2003; Bettinson, 2013; Matamala, Romero-Fresco, & Daniluk, 2017; Sperber, Neubig, & Fügen, 2013; Zick & Olsen, 2011). Of those strands, the most relevant one for the purposes of this article is the analysis of live/respoken subtitles and, more specifically, the assessment of quality, given that the certification presented here effectively consists of analysing the quality of the output produced by the different candidates.

Assessing the quality of live subtitles is still one of the most often debated topics in this area, as it has proved to be of interest not only to researchers but also to companies and users. Different assessment methods have been proposed for this purpose. Some are based on subtitling theory (Eugeni, 2012), while others have their origin in the professional market (Dumouchel, Boulianne, & Brousseau, 2011) or in scientific efforts to automate quality assessment (Apone, Brooks, & O’Connell, 2010). In Canada, for instance, the Canadian Radio-television and Telecommunications Commission (CRTC) in 2012 launched a two-year project to analyse the quality of the live captions provided in 265 programmes using the so-called Verbatim Test (English Broadcasters Group [EBG], 2014). Only 19% of the programmes analysed reached the 95% threshold established by the Verbatim Test, which considers accuracy as the extent to which the captions match the audio of a programme verbatim.

Broadcasters criticized this method for its inability to assess accuracy and for pushing captioners to provide verbatim captions that may be too fast for the viewers to read instead of correctly edited captions that retain the meaning of the audio. The NER model (Romero-Fresco & Martínez, 2015) was developed in 2012 in order to allow for the possibility of correct (and incorrect) editing and for the occurrence of different types of error in live subtitling. Over the past few years, it has been widely used by universities, broadcasters, access service providers and regulators in countries such as the Switzerland, Italy, the United Kingdom, Spain, France, South Africa and Australia. The NER model accounts for editing errors (caused by the respeakers’ decisions when they need to edit the original audio content if it is not possible to respeak it verbatim) and recognition errors (caused by the interaction between the respeakers and the speech-recognition software). These errors can in turn be minor (the viewer may notice the error but the main meaning is retained), standard (the main meaning is lost) or serious (the incorrect meaning appears on the screen). In 2013, the UK governmental regulator Ofcom adopted the NER model in order to set up the largest study conducted so far on the quality of live subtitling, which analysed the accuracy, delay, speed and edition rate of

Romero-Fresco, P., Melchor-Couto, S., Dawson, H., Moores, Z. & Pedregosa, I. (2019). Respeaking certification: Bringing together training, research and practice. *Linguistica Antverpiensia, New Series: Themes in Translation Studies*, 18, 216–236.

78,000 subtitles from news programmes, entertainment programmes and chat shows broadcast by all terrestrial TV channels in the country (Romero-Fresco, 2016).

The results of this project have proved very useful to informing the design of the LiRICS certification process presented here and also to providing a benchmark against which to compare the performance of the candidates.

Ofcom's reports on the quality of subtitles show an overall accuracy rate of 98.4%, which is above the 98% threshold set by the NER model and may be considered as acceptable quality. Generally speaking, subtitles with a less than 98% accuracy rate are regarded as substandard, subtitles with a 98–98.49% rate are regarded as acceptable, subtitles with a 98.5–98.99% rate as good, subtitles with a 99–99.49% rate as very good and subtitles with a 99.5–100% rate as excellent. Of the 300 programmes analysed across the two-year period, 23% of the programmes did not reach the required accuracy threshold, whereas 77% did. Of the latter, 21% had acceptable subtitles, 29% had good subtitles, 22% had very good subtitles and 5% had excellent subtitles. Interestingly, 60% of the programmes that did not reach the required accuracy threshold are chat shows (as compared to 26% entertainment shows and 14% news shows), which highlights the extent to which genre can have an impact on subtitle quality. News programmes obtained the highest average accuracy rate, at 98.75%, as they normally feature only one speaker at a time and they tend to combine pre-recorded (and thus 100% accurate) and live subtitles. Entertainment programmes, with the added difficulty of featuring several speakers but also with a combination of live and pre-recorded subtitles, followed with an average accuracy rate of 98.54%. Finally, chat shows had an average accuracy rate just below the threshold (97.9%), which can be explained by the high speech rates, the presence of multiple speakers and the absence of a script or pre-recorded subtitles.

In the samples analysed during the two-year project, 69% of the errors observed in the subtitles were editing errors, that is, those caused by incorrect omissions or additions made by the subtitlers, errors of speaker identification, etc. The remaining 31% were recognition errors, that is, those caused by the interaction between the subtitler and the steno machine or the speech-recognition software. Once again, these percentages vary depending on the genres. In chat shows, which feature very high speech rates that are hard for subtitlers to follow, 75% of the errors were caused by incorrect editions (typically omissions) and 25% by misrecognitions. Entertainment programmes, where speech rates are slower, contain 69% edition errors and 31% recognition errors. Finally, news programmes feature 61% edition errors and 39% recognition errors. This relative increase in recognition errors in the news as compared to entertainment programmes and chat shows may be due to both the effort made by subtitlers to type/respeak fast in order to keep up with the audio without editing too much and to the very content of the news, which is likely to include more specialized terms and unexpected proper nouns than chat shows and entertainment programmes. This is one of the reasons why having access to the script before the programme – when one is available – can help to improve accessibility for the viewers.

As far as the seriousness of the errors is concerned, in general 56% of the errors found in the Ofcom sample were minor (i.e., they do not prevent the viewers from following the content of the programme), 39% were standard (i.e., they trigger confusion or cause full factual

Romero-Fresco, P., Melchor-Couto, S., Dawson, H., Moores, Z. & Pedregosa, I. (2019). Respeaking certification: Bringing together training, research and practice. *Linguistica Antverpiensia, New Series: Themes in Translation Studies*, 18, 216–236.

omissions) and 5% were serious (i.e., they introduce misleading information). These figures vary depending on whether they relate to edition errors (56% of which were minor, 42% were standard and 2% were serious) or recognition errors (62% of which were minor, 31% were standard and 7% were serious). In other words, editing errors tend to be more problematic than recognition errors, since it is more common to have standard edition errors (omissions of full sentences) than standard recognition errors (nonsensical misrecognitions). The different genres also play an important role here. As a result of the effort made by the subtitlers to improve the quality of the subtitles for the news, one-third of the errors in these programmes are standard and two-thirds are minor. In contrast, in chat shows the fast speech rates and the overlapping interventions of the speakers force the subtitlers to rush and to omit more information. As a result, 56% of the edition errors were minor and as many as 43% were standard. In other words, almost one in two errors found in chat shows involves the omission of a full sentence; in the worst cases, this may cause the viewers to lose the thread of the programme.

The large-scale research conducted so far on the quality of live subtitles is essential to informing the creation of a certification for professional respeakers, especially when it comes to grading the audiovisual material according to levels of difficulty (e.g., by genre), setting accuracy thresholds for the different levels and analysing the performance of the candidates compared to current professional standards (accuracy, types and severity of errors, etc.).

4. LiRICS certification

LiRICS is an online certification process, where testing is carried out remotely. This section describes the different steps involved in the creation of the certification process, including the design of the test, the different aspects required to conduct it online and the development of a quality assurance process.

4.1 Designing the test

The LiRICS test assesses the candidates' ability to respeak across two broad contexts: TV, on the one hand, and education and live events, on the other. The genres selected for assessment in the TV subject area are news, sports and entertainment/chat shows, that is, the genres covered in the Ofcom project with the addition of sports, which has been assessed in the abovementioned two-year trial set up in Canada by the CRTC. For education and live events, the genres chosen are a class or a lecture, a conference presentation and an interview, which are the most common genres respoken by AiMedia, the biggest provider of respeaking for live events in Europe and Australia and the company whose subtitlers effectively participated in the pilot certification test.

Within these genres, materials were chosen at three different levels of difficulty in order to assess a candidate's ability. Level 1 is the lowest level of assessment and the entry level for all candidates; once a candidate has passed this assessment, they are able to move on to levels 2 and 3. The pilot study was based on this initial Level 1 certification.

A number of parameters differentiate the difficulty between levels, content being one of them. Level 1 test materials are of a more general nature and not as specialized as those of

Romero-Fresco, P., Melchor-Couto, S., Dawson, H., Moores, Z. & Pedregosa, I. (2019). Respeaking certification: Bringing together training, research and practice. *Linguistica Antverpiensia, New Series: Themes in Translation Studies*, 18, 216–236.

levels 2 and 3. Other aspects are also taken into account, such as sound quality, the number and delivery of the speakers (in terms of pronunciation, accent, speech rate and whether they speak spontaneously or read a written text) or whether there are any visual aids that support the verbal information.

4.1.1 Assessment-setting proformas

Assessment-setting proformas were produced as tools for the assessing team to check that the selected video clips were set at the correct level. For each video clip a proforma was filled out, which included the following information: genre, provenance, title, publication date and synopsis of the video with the time frame that had been selected to be respoken. The image below shows an example of the first section of an assessment-setting proforma.



LiRICS CERTIFICATION ASSESSMENT SETTING PROFORMA

LEVEL OF CERTIFICATION **1**
STRAND **EDUCATION AND LIVE EVENTS**

PROFORMA - YOUR ASSESSOR WORKING TOOL (Please complete 1 per genre)

Option (please delete as appropriate)	Q&A
Source (please add a link if available)	https://www.youtube.com/watch?v=n4xzvDzP-IA
Introductory rubric This text will be available to candidates one hour before the start of the respoking. Please provide: <ul style="list-style-type: none"> • the provenance of the video • title of the video • original publication date • date accessed/downloaded • synopsis or gist of the video • minutes from and to be respoken if the whole video is not to be respoken 	Provenance: Totally Emma Watson YouTube channel Title of the video: Emma Watson Facebook Q&A about HeForShe - International Women's Day 2015 Original publication date: 09/03/2015 Date accessed/downloaded: 18/07/2018 Synopsis of the video: Greg James interviews Emma Watson and discusses the speech that Emma Watson delivered at the UN, which discussed gender equality and how societal issues surrounding feminism. Minutes to be respoken: 00:02:57 – 00:17:51
Reasons for choosing the video: <i>This is for internal purposes, not to be published</i>	This video has limited overlapping speech, only two speakers and non-technical vocabulary with the occasional proper noun.
Any other comments <i>This is for internal purposes, not to be published</i>	

Figure 1 Assessment-setting proforma, section 1

In addition to the aforementioned features, the proforma identified a list of challenges which the respeaker could encounter during the test. Types of challenge included change of speaker, proper nouns, specialized terminology and the need to restructure sentences. For Level 1, each video clip was required to contain a minimum of five challenges. For each challenge, the time code, text and type of challenge were noted on the proforma. The image below shows an example of the second section of an assessment-setting proforma.

Romero-Fresco, P., Melchor-Couto, S., Dawson, H., Moores, Z. & Pedregosa, I. (2019). Respeaking certification: Bringing together training, research and practice. *Linguistica Antverpiensia, New Series: Themes in Translation Studies*, 18, 216–236.



Challenges <i>Please include in this column the timecode where this occurs</i>	Challenge A Level 1 Certification assessment should have a minimum of five minor challenges	Type of challenge E.g. change of speaker, proper noun, specialised terminology, sentence restructuring,
00:05:13	Rapid change of speaker between the two and then it changes back again.	Change of speaker
00:06:59	'Impact 10 by 10 by 10'	Name of a project could be specialised vocabulary and it is spoken very fast
00:07:17	'Unilever, PWC, Barclays, Tupperware, some universities, Oxford university'	A list of proper nouns
00:08:36	'Steve Carell'	Proper noun
00:16:47	Overlapping speech between the two speakers	Overlapping speech

Figure 2 Assessment-setting proforma, section 2

4.1.2 Video clips

The video clips chosen for the LiRICS Level 1 certification test were these: a classroom setting focusing on the science of sound for primary-school children with a duration of 12 minutes 9 seconds and spoken at 166 words per minute (wpm); a conference presentation discussing the science of simplicity with a duration of 15 minutes 13 seconds and spoken at 163 wpm, and a Q&A interview on the topic of feminism with a duration of 14 minutes 54 seconds and spoken at 175 wpm.

4.2 Testing online

Setting up online testing for the LiRICS pilot posed several challenges and involved careful consideration of different aspects such as identity verification and invigilation.

4.2.1 Key considerations in creating the platform

When creating the testing platform, it was therefore essential to incorporate identity verification and to prevent opportunities for cheating during the test within the design of the platform and the procedures surrounding its use (LearningLight, n.d.). The platform also had to be one that enabled the respeakers to replicate their workplace routines to allow them to perform as well as possible. For assessment purposes, the examiners needed to provide a transcript of the respoken text and a recording of the respeakers' voices. After much investigation, it was decided that five different pieces of software would be used to achieve this testing set-up: the respeakers would dictate with their regular respeaking software, DragonNaturallySpeaking; and the Screencast-o-matic screen recorder, Google Drive, Classmarker and Vimeo would be used to create the other elements of the platform. Careful thought was given to the precise process of running the tests so the experience for the candidates would be as smooth and stress-free as possible.

Romero-Fresco, P., Melchor-Couto, S., Dawson, H., Moores, Z. & Pedregosa, I. (2019). Respeaking certification: Bringing together training, research and practice. *Linguistica Antverpiensia, New Series: Themes in Translation Studies*, 18, 216–236.

4.2.2 Identity verification and invigilation

Candidates were asked to provide a screen recording of the duration of the test. They were asked to record both audio and video with Screencast-o-matic and to set their screen up so that the certifiers would be able to see the testing platform, the video being respoken and the window recording them at work. This prevented the candidates from pausing or restarting the test clip, which would have invalidated the test. The audio included the voice of each respeaker over the original video, which served to prevent cheating. On completion, the candidates shared the video with the assessors via Google Drive. The official identity of the candidates had already been confirmed by the candidates' company. As testing expands, identity documents will need to be checked before the test begins.

4.2.3 Test procedures

Classmarker was selected as the testing platform, as it allows remote testing which can be scheduled and timed. This set-up formalized the testing process. Each candidate received a specific time slot for the test, during which an invigilator would be online and available to respond to any queries and deal with any technical issues that might arise. Each test expired ten minutes after the due completion time, ensuring that testing conditions were the same for all candidates.

The clips used for the test were open source, so, during the pilot study, a number of measures were taken to prevent the candidates from accessing and identifying the clips before they respoke them. The candidates accessed the assessment clips through a link to Vimeo, which was cued to appear question-by-question after the preparation slot ended. This meant that respeakers could use the preparation time to research the general theme of the clip but not listen to the content that they were about to respeak. Furthermore, the candidates were requested not to share information about the clips with other candidates yet to take the test and the tests were scheduled across a short period of time to support this. When the certification is rolled out on a larger scale, a bank of video clips from different sources will be used, which will remove the possibility of candidates' identifying and preparing for the test clip in advance.

In order to minimize technical problems on the day, then candidates were provided with very thorough instructions a week before their scheduled test date about how to set up each piece of software correctly and to use them in combination. In addition, the first question in the test comprised a trial run which allowed them to select the desktop set-up that was most comfortable for them and ensured that the different pieces of software required worked simultaneously before the testing proper began.

4.2.4 Technical challenges and future considerations

The pre-test instructions meant that, for the most part, testing ran smoothly and minimal technical difficulties were experienced. The main technical issues experienced involved the respoken text not appearing on screen or disappearing entirely. The video recording guaranteed that this text could be retrieved by assessors so that candidates were not

Romero-Fresco, P., Melchor-Couto, S., Dawson, H., Moores, Z. & Pedregosa, I. (2019). Respeaking certification: Bringing together training, research and practice. *Linguistica Antverpiensia, New Series: Themes in Translation Studies*, 18, 216–236.

disadvantaged. The recording also facilitated quality assurance during the marking process (section 4.3 below).

The recording process posed two key problems: some candidates placed a required window on their second screen, so it was not recorded. This did not compromise invigilation or assessment quality, but it meant that the marking process needed to be adapted. Some homeworkers felt that the recording was an intrusion. However, since recording is integral to the certification process, earlier notification will be given in future tests to allow candidates to make any adjustments required to their working set-up.

4.3 Quality assurance process

The tests were marked with the NER model (Romero-Fresco & Martínez, 2015), following the same process used during Ofcom’s sampling of live television subtitling in 2014–2015 (Romero-Fresco, 2016), and all examiners were experienced in using this model. Procedures for first and second marking were determined to ensure rigorous marking and fair treatment of all the candidates. Since the pilot was the first certification of live respeaking to take place, it also revealed areas for further refinement to the quality assurance process, which will be put in place during subsequent rounds of testing.

4.3.1 Marking process

For effective use of the NER model, a careful comparison of the words spoken in the original source video (the verbatim transcript) and the respoken transcript is required. As mentioned in section 3, differences between the two are classified as either edition (E) errors, where the respeaker has added, omitted or changed something in the original, or recognition (R) errors, where mispronunciations or mishearing means that a word is incorrectly recognized by the speech-recognition software. The severity of these errors is weighted and scored and the value obtained is then deducted from the total number of words spoken (N), allowing the accuracy to be calculated:

$$\text{Accuracy} = \frac{N - E - R}{N} \times 100$$

Figure 3 NER assessment formula

Correct editions, where the respeaker has edited the text without causing loss of information, are also noted.

Internal marking sheets were designed to ensure homogenous marking among examiners and errors were automatically calculated:

Romero-Fresco, P., Melchor-Couto, S., Dawson, H., Moores, Z. & Pedregosa, I. (2019). Respeaking certification: Bringing together training, research and practice. *Linguistica Antverpiensia, New Series: Themes in Translation Studies*, 18, 216–236.

Candidate's Name: Assessment Date: Level 1 Q&A Emma Watson 1		NER SCORING								
VERBATIM TRANSCRIPT	RESPOKEN TRANSCRIPT	1ST EXAMINER COMMENTS	2ND EXAMINER COMMENTS	NER Misinformation	No. of CE	No. of SE	No. of ME	No. of SR	No. of STR	No. of MR
INTERVIEWER: We are live at the Facebook offices today (beginning and we will get some fan questions later, which will have time for.				0.00						
But I want to talk about that speech that you did on the 20th of				0.00						
Now obviously talking at the UN is an incredibly nerve-wracking thing to do I imagine.				0.00						
I mean, I can only imagine				0.00						
I have not actually spoken at the UN, yet,				0.00						
they have not asked me yet.				0.00						
but what was the mission statement?				0.00						
what was going through your head in the morning of that speech?				0.00						
What was the mission statement you wanted to get across?				0.00						
EMMA: I guess I really want to communicate that gender equality historically has been predominantly a women's movement for				0.00						
But I think the impact of gender inequality and how it has actually been affecting men has not really been addressed.				0.00						
So, I really wanted to make that one of the clearest messages in my and I think also our society in general devalues the she				0.00						
and you know when I say that I mean qualities that are associated with the fememore that are found in all of us.				0.00						
And as result there is kind of this imbalance and this distortion, and it's just hindering our progress,				0.00						
it's raising discrimination and violence, rain and fear the world over				0.00						

Figure 4 Internal marking sheet

The marking sheet allowed the full transcript to be divided into smaller chunks of text, usually a sentence long each (or independent idea unit, as per NER terminology). This made the process of marking and reviewing far clearer for all the examiners involved and allowed for easy comparison of responses across candidates, ensuring consistency in marking. Comment boxes allowed each marker to explain the error seen and second markers to add their responses, where necessary, to the first marker's decision. The coloured columns allowed each marker to record the types of error noted. The second marker worked with a copy of the first markers' sheet and adjusted any discrepancies according to their own marking. In accordance with the NER model, recognition and edition errors were divided into three categories: serious, standard and minor errors. Serious errors changed the meaning of the original and carried a penalty of one mark; standard errors resulted in the omission of an information unit or disrupted the flow of meaning (Romero-Fresco, 2016) and had a penalty of 0.5 marks; minor errors still allowed the viewer to follow the meaning of the original and had a penalty of 0.25 marks. Correct editions were totalled but not weighted. As the markers enter the types of error, the Excel spreadsheet automatically calculates the total number of points deducted.

Score and Feedback					
Scores	1st marker	2nd marker	Legend	Ded	
Correct Editions	65	75	CE Correct Edition	0	
Serious Edition Errors	2	1	SE Serious Edition Error Misinformation	1	
Standard Edition Errors	2	5	STE Standard Edition Error Omission of independent idea unit	0.50	
Minor Edition Errors	29	23	ME Minor Edition Error Omission of dependent idea unit	0.25	
Serious Recognition Error	6	5	SR Serious Recognition Error Misinformation	1	
Standard Recognition Error	12	9	STR Standard Recognition Error Affects comprehension - nons	0.5	
Minor Recognition Errors	12	15	MR Minor Recognition Error Does not affect comprehension	0.25	
Total Deductions		22.75			
Total Words + Commands	2499				
Accuracy					
1st marker score	98.99		Number of words: all words in the subtitles including labels (i.e. speaker, inaudible) + punctuation		
2nd marker - revised score		99.10%	2499 - 22.5 / 2499 = 2476.5 / 2499 = 99.10%		
Stars		***			
Pass or Fail		PASS	Number of words in the subtitles		
			2274		
			Number of words + punctuation		
			2499		

Figure 5 Excel spreadsheets used for marking purposes

At the end, in addition to completing the NER calculation, each marker wrote a detailed comment to capture the overall quality of the respeaking.

Where a candidate did not achieve the expected level of accuracy, a third marker reviewed the assessment for confirmation.

Romero-Fresco, P., Melchor-Couto, S., Dawson, H., Moores, Z. & Pedregosa, I. (2019). Respeaking certification: Bringing together training, research and practice. *Linguistica Antverpiensia, New Series: Themes in Translation Studies*, 18, 216–236.

4.3.2 Feedback to candidates

Each candidate received an assessment report with feedback on their performance. In addition to finding out whether they had passed or failed and receiving their overall accuracy score across all three tests, for each clip they received detailed comments on their performance and the level of accuracy attained and a grid indicating the total number of errors in each category.

4.3.3 Interrater disagreement

The interrater disagreement seen across first and second marking for LiRICS is 0.24%. This figure was calculated by averaging the difference between first and second marking for the 27 respoken texts produced by nine respeakers. Although this interrater disagreement is higher than the one obtained in the Ofcom study, which was reported at 0.09% (Romero-Fresco, 2016), it is still negligible. It is equivalent to 0.5 in a 1 to 10 scale and it means that all 27 respoken texts assessed by two markers were placed in the same band (i.e., pass or fail).

4.3.4 Modifications to marking process

During the pilot test, it became clear that, despite the rigorous marking of both first and second examiners, a formalized moderation process or discussion forum which allowed responses to be compared across candidates was also required. To further support the first and second markers, moderation sessions will be held to agree on expected standards for each genre and co-mark respoken passages, in the case of those which might be particularly problematic. In the pilot test, the first and second markers were randomly assigned, but as a way forward a more rationalized method of marking would be to assign the same first marker to the same genre. This, we expect, would reduce the extent of interrater disagreement.

5. Analysis of the LiRICS pilot

The first phase of the LiRICS project, completed in December 2018, can be regarded as a pilot study. It involved 27 tests by nine candidates, all of them professional respeakers. Six of them were awarded a Level 1 LiRICS certificate and three of them will need to resit all or some of the tests. In total, the LiRICS examiners analysed 6,000 subtitles and 40,000 words, which is far from the 78,000 subtitles and 546,000 words analysed in the Ofcom project (Romero-Fresco, 2016). However, the unexpected similarity between the results obtained in the two projects makes it possible to draw more conclusions from this pilot than may have been expected.

Tables 1–6 compare the results of the two projects and, more specifically, the accuracy rates, types (edition vs recognition) and severity (minor, standard and serious) of errors in general and per genre.

Romero-Fresco, P., Melchor-Couto, S., Dawson, H., Moores, Z. & Pedregosa, I. (2019). Respeaking certification: Bringing together training, research and practice. *Linguistica Antverpiensia, New Series: Themes in Translation Studies*, 18, 216–236.

Table 1 Summary of results for the Ofcom project and the LiRICS pilot

	Ofcom project	LiRICS pilot
Average accuracy rate	98.4%	98.5%
Excellent subtitles	5%	7%
Very good subtitles	22%	30%
Good subtitles	29%	33%
Acceptable subtitles	21%	7%
Substandard subtitles	23%	23%

The average accuracy rate obtained by all candidates for all videos is 98.5%, practically the same as the rate obtained by the dozens of respeakers analysed over two years for the Ofcom project (98.4%). In the LiRICS pilot, 23% of the videos analysed did not attain the minimum NER threshold of 98% and 77% did – exactly the same percentages found in the Ofcom project. A slight difference may be found in the distribution between acceptable, good, very good and excellent subtitles, with better results in the LiRICS pilot, that is, a higher percentage of good, very good and excellent subtitles than in the Ofcom study. This makes sense if we consider that candidates were being tested for Level 1 certification and that there are two more levels of increased difficulty.

As shown in Table 2, once again, the difference between genres has proved to have a considerable impact on the respeakers' performance. In the Ofcom project, the most challenging genre was the chat shows, 60% of which did not attain the required quality threshold (as compared to 26% for entertainment shows and 14% for news shows). In the LiRICS test, it is the classroom clip that proved to be particularly difficult, with 66% of the programmes not attaining the threshold (as compared to 17% for both the Q&A and the conference clips). Whereas in the chat shows the main problems were the lack of a script, the high speech rates and the overlapping speech, the classroom clip proved particularly challenging because of the interaction between teacher and students. Not all of the pupils were always fully audible and, even when they were, respeakers were not always sure whether or not to include them in the subtitles.

Table 2 Programmes below the accuracy threshold per genre

Ofcom project		
Chat shows	Entertainment	News
60%	26%	14%
LiRICS pilot		
Classroom	Q&A	Conference
66%	17%	17%

As a result, the average accuracy rate of the classroom clip for all candidates was 97.84%, considerably lower than the 98.93% obtained for the Q&A clip and the 98.74% obtained for

Romero-Fresco, P., Melchor-Couto, S., Dawson, H., Moores, Z. & Pedregosa, I. (2019). Respeaking certification: Bringing together training, research and practice. *Linguistica Antverpiensia, New Series: Themes in Translation Studies*, 18, 216–236.

the conference clip. On the whole, these are slightly better results per genre than those obtained in the Ofcom project (97.9% for chat shows, 98.54% for entertainment shows and 98.75% for news programmes); this is shown by the fact that 22% of the candidates produced very good or excellent subtitles (over 99% accuracy rate) for the Q&A clip and as many as 55% produced good subtitles (between 98.5% and 99%) for the conference clip. See Table 3 below.

Table 3 Accuracy rate per genre

Ofcom project		
Chat shows	Entertainment	News
97.9%	98.54%	98.75%
LiRICS project		
Classroom	Q&A	Conference
97.84%	98.93% (22% very good or excellent)	98.74% (55% good)

The findings relating to the types (edition vs recognition) and severity of errors (minor, standard and serious), presented in Tables 4, 5 and 6, provide further data with which to analyse the respeakers' performance and, more specifically, the low scores obtained in the classroom clip. The LiRICS pilot confirms that incorrect editions, that is, instances in which meaning is lost because respeakers cannot keep up with the original audio and omit information, are the most common type of error in respeaking. Edition errors are approximately twice as recurrent as recognition errors (65% edition errors vs 35% recognition errors in LiRICS, as compared to 69% edition errors vs 31% recognition errors in the Ofcom project). It is also evident that the more challenging the genre, the higher the percentage of edition errors: 76.5% edition errors vs 23.5% recognition errors in the classroom clip and, in the Ofcom project, 75% edition errors vs 25% recognition errors in chat shows.

Table 4 Total edition and recognition errors

	Ofcom project	LiRICS pilot
Total edition errors	69%	65%
Total recognition errors	31%	35%

Romero-Fresco, P., Melchor-Couto, S., Dawson, H., Moores, Z. & Pedregosa, I. (2019). Respeaking certification: Bringing together training, research and practice. *Linguistica Antverpiensia, New Series: Themes in Translation Studies*, 18, 216–236.

Table 5 Edition and recognition errors per genre

	Ofcom project			LiRICS pilot		
	Chat shows	Entertainment	News	Classroom	Q&A	Conference
Edition errors	75%	69%	61%	76.5%	60%	55%
Recognition errors	25%	31%	39%	23.5%	40%	45%

The analysis of error severity reveals further parallels between the LiRICS pilot and the Ofcom project: 56% minor errors, 36% standard errors and 8% serious errors in the former and 56% minor errors, 39% standard errors and 5% serious errors in the latter. Once again, edition errors tend to be more serious than recognition errors and this is particularly true of the most challenging genres, in this case the classroom clip, where one in two edition errors is standard – that is, one in two errors involves the omission of a full sentence – as opposed to one in five in the case of the conference and the Q&A clips. Those omitted sentences were often comments by the students.

Table 6 Total minor, standard and serious errors

	Ofcom project	LiRICS pilot
Minor errors	56%	56%
Standard errors	39%	36%
Serious errors	5%	8%

The breakdown of the type and severity of errors in the LiRICS study shows once again that the key to achieving high-quality subtitles is not only to have fewer errors but also to control their severity and to ensure that most of them are minor rather than standard or serious.

6. Conclusions

The considerable impact that live subtitling and, more specifically, respeaking has had on both the audiovisual translation market and society as a whole has not been matched by an equally significant body of publications or training courses. Academic research and training in respeaking had a slow start, which meant that many companies had no choice but to develop their own in-house training programmes. Although by 2012 six European HEIs were offering respeaking courses, seven years later the offering has not increased as much as could have been expected. This may be due to the lack of respeaking trainers, which is related to the low number of researchers working in this area. This landscape might help explain why the UK government had decided not to use Disabled Students' Allowances for respeaking until professional respeakers are recognized/certified and, most importantly, it justifies the need for an official certification such as LiRICS.

As shown in this article, the creation of a respeaking certification is a complex and time-consuming process that involves careful consideration of the choice and grading of material,

Romero-Fresco, P., Melchor-Couto, S., Dawson, H., Moores, Z. & Pedregosa, I. (2019). Respeaking certification: Bringing together training, research and practice. *Linguistica Antverpiensia, New Series: Themes in Translation Studies*, 18, 216–236.

technical set-up, online testing, peer-reviewed assessment, etc. The procedure chosen for LiRICS and tested in a pilot scheme is by no means perfect, but it has been shown to be effective enough to achieve the goal of certifying professional respeakers, while also revealing aspects that are still in need of improvement.

The analysis of the data obtained from the candidates has also proved extremely useful. Their striking resemblance to the results of the much larger Ofcom study (Romero-Fresco, 2016) reveals a few interesting lessons: first, despite the low number of participants in the pilot, their results can probably be extrapolated and are representative of the current provision of respokey subtitles in the United Kingdom; secondly, we have probably succeeded in aligning the certification with current professional standards; and, thirdly, the decision to draw on the Ofcom project as an inspiration for designing the assessment included in LiRICS has been useful. Some limitations have also been revealed, not least of them the excessive difficulty of the classroom clip. Whereas it made sense for some genres such as chat shows to be more difficult than others in the Ofcom project, this should not apply to LiRICS, where all the clips should be equally difficult at a level commensurate with the level that is being tested. Another limitation is the absence of some key live subtitling dimensions such as delay and subtitling speed, which have not been tested in LiRICS.

Finally, the data obtained in the LiRICS pilot have also helped to provide a clearer picture of respeaking, highlighting the importance of edition errors over recognition errors and the need to control error severity in order to obtain a satisfactory accuracy rate. This should be fed back into training, which does not always focus on these elements. This last point helps to end this article on a positive note. Despite the limited number of trainers and researchers working on respeaking, the vibrant work produced in this area is an example of fruitful cross-fertilization between teaching, research and professional practice. The first training programmes in respeaking at HEIs introduced the NER model as a method with which to rate students' work; this was then used in research projects to assess the quality of the live TV subtitles, such as the Ofcom study, which has in turn informed the design and implementation of LiRICS. The circle is now closed, as the findings of the LiRICS project feed back into respeaking training programmes.

As new challenges appear in the professional field of live subtitling (for instance, interlingual live subtitling and live subtitles produced by automatic speech recognition), it is essential to maintain this tight connection between training, research and practice in order to ensure that the rapid development of live subtitling is accompanied by the required quality standards that can guarantee full access for all viewers.

References

- Al-Aynati, M. M., & Chorneyko, K. A. (2003). Comparison of voice-automated transcription and human transcription in generating pathology reports. *Archives of Pathology and Laboratory Medicine*, 127(6), 721–725.
- Arumí Ribas, M., & Romero-Fresco, P. (2008). A practical proposal for the training of respeakers. *The Journal of Specialised Translation*, 10, 106–127.
- Baaring, I. (2006). Respeaking-based online subtitling in Denmark. In C. Eugeni & G. Mack (Eds.), *iTRAlinea special issue: Respeaking*. Retrieved from <http://www.intralinea.org/specials/arti%0Ac>

- Romero-Fresco, P., Melchor-Couto, S., Dawson, H., Moores, Z. & Pedregosa, I. (2019). Respeaking certification: Bringing together training, research and practice. *Linguistica Antverpiensia, New Series: Themes in Translation Studies*, 18, 216–236.
- le/Respeaking-based_online_subtitling_in_Denmark
- Bettinson, M. (2013). *The effect of respoking on transcription accuracy* (Unpublished honour's thesis). University of Melbourne, Melbourne.
- Bortone, M. (2015). *Quality of chat shows in Italy: A comparative analysis of respoken and steno-typed subtitles* (Unpublished master's thesis). University of Roehampton, London.
- CRTC (Canadian Radio-television and Telecommunications Commission). (2019). *Broadcasting regulatory policy CRTC 2019-308*. Retrieved from <https://crtc.gc.ca/eng/archive/2019/2019-308.htm>
- Chen, S.-J. (2006). Real-time subtitling in Taiwan. In C. Eugeni & G. Mack (Eds.), *inTRAlinea special issue: Respeaking*. Retrieved from <http://www.intralinea.org/specials/article/1693>
- De Serriis, L. (2006). Il servizio sottotitoli RAI: Televideo per i non udenti. In C. Eugeni & G. Mack (Eds.), *inTRAlinea special issue: Respeaking*. Retrieved from http://www.intralinea.org/specials/article/II_Servizio_Sottotitoli_RAI
- Dumouchel, P., Boulianne, G., & Brousseau, J. (2011). Measures for quality of closed captioning. In A. Serban, A. Matamala, & J.-M. Lavour (Eds.), *Audiovisual translation in close-up: Practical and theoretical approaches* (pp. 161–172). Bern: Peter Lang.
- The English Language Broadcasters Group. (2014). *Report on efforts to improve the quality of closed captioning*. Retrieved from http://www.crtc.gc.ca/fra/BCASTING/ann_rep/bmt_cbc_rm_sm.pdf
- Eugeni, C. (2008). Respeaking the TV for the Deaf: For a real special needs-oriented subtitling. *Studies in English Language and Literature*, 21, 37–47.
- Eugeni, C. (2008). *La sottotitolazione in diretta TV: Analisi strategica del rispeaking verbatim di BBC News* (Unpublished doctoral dissertation). Università degli Studi di Napoli Federico II, Italy. Retrieved from http://www.fedoa.unina.it/3271/1/Carlo_Eugeni.pdf
- Eugeni, C. (2012). A strategic model for the analysis of respoken TV subtitles. *US-China Foreign Language*, 10(6), 1276-1286.
- Eugeni, C., & Mack, G. (2006). Special Issue on New Technologies in Real Time Intralingual Subtitling. *InTRAlinea, Special Issue on Respeaking*. Retrieved from <http://www.intralinea.org/specials/respeaking>
- García Romero, A. J. (2015). *Measuring accuracy, delay, errors and speed in live subtitling: Revisiting the application of the NER model in the Spanish television* (Unpublished master's thesis). University of Roehampton, London.
- International Telecommunication Union. (2015). *Series F: Non-telephone telecommunication services: Audiovisual services: Accessibility terms and definitions*. Retrieved from <https://www.itu.int/ITU-T/recommendations/rec.aspx?rec=12624&lang=en>
- Jensem, C., McCann, R., & Ramsey, S. (1996). Closed-captioned television presentation speed and vocabulary. *American Annals of the Deaf*, 141(4), 284–292. doi:10.1353/aad.2012.0377
- Lambourne, A. (2006). Subtitle respoking: A new skill for a new age. In C. Eugeni & G. Mack (Eds.), *inTRAlinea special issue: Respeaking*. Retrieved from <http://www.intralinea.org/specials/article/1686>
- Luyckx, B., Delbeke, T., Van Waes, L., Leijten, M., & Remael, A. (2010). Live subtitling with speech recognition: Causes and consequences of text reduction. In *Artesis VT Working Papers in Translation Studies*. Retrieved from: <https://repository.uantwerpen.be/docman/irua/7418cf/963a308c.pdf>
- Mack, G. (2006). Detto scritto: Un fenomeno, tanti nomi. In C. Eugeni & G. Mack (Eds.), *inTRAlinea special issue: Respeaking*. Retrieved from <http://www.intralinea.org/specials/article/1695>
- Marsh, A. (2006). Respeaking for the BBC. In C. Eugeni & G. Mack (Eds.), *inTRAlinea special issue: Respeaking*. Retrieved from <http://www.intralinea.org/specials/article/1700>
- Matamala, A., Romero-Fresco, P., & Daniluk, L. (2017). The use of respoking for the transcription of non-fictional genres: An exploratory study. *inTRAlinea*, 19. Retrieved from <http://www.intralinea.org/archive/article/2262>

- Romero-Fresco, P., Melchor-Couto, S., Dawson, H., Moores, Z. & Pedregosa, I. (2019). Respeaking certification: Bringing together training, research and practice. *Linguistica Antverpiensia, New Series: Themes in Translation Studies*, 18, 216–236.
- Muller, T. (2015). Long questionnaire in France: The viewer's opinion. In P. Romero-Fresco (Ed.), *The reception of subtitles for the deaf and hard of hearing in Europe: UK, Spain, Italy, Poland, Denmark, France and Germany* (pp. 163–187). Bern: Peter Lang.
- Muzii, L. (2006). Respeaking e localizzazione. In C. Eugeni & G. Mack (Eds.), *inTRAlinea special issue: Respeaking*. Retrieved from <http://www.intralinea.org/specials/article/1688>
- Apone, T., Brooks, M., & O'Connell, T. (2010). *Caption accuracy metrics project: Caption viewer survey: Error ranking of real-time captions in live television news programs*. Retrieved from WGBH National Center for Accessible Media old website: http://ncamftp.wgbh.org/ncam-old-site/file_download/CCM_survey_report_final_Dec_17_2010.pdf
- Orero, P. (2006). Real-time subtitling in Spain: An overview. In C. Eugeni & G. Mack (Eds.), *inTRAlinea special issue: Respeaking*. Retrieved from <http://www.intralinea.org/specials/article/1689>
- LearningLight (n.d.) *Online proctoring/remote invigilation: Soon a multibillion dollar market within eLearning & assessment*. Retrieved December 5, 2018, from <https://www.learninglight.com/remote-proctoring-invigilation-market/>
- Rajendran, D. J., Duchowski, A. T., Orero, P., Martínez, J., & Romero-Fresco, P. (2013). Effects of text chunking on subtitling: A quantitative and qualitative examination. *Perspectives: Studies in Translation Theory and Practice*, 21(1), 5–21. doi:10.1080/0907676X.2012.722651
- Remael, A., & van der Veer, B. (2006). Real-time subtitling in Flanders: Needs and teaching. In C. Eugeni & G. Mack (Eds.), *inTRAlinea special issue: Respeaking*. Retrieved from <http://www.intralinea.org/specials/article/1702>
- Robert, I., Schrijver, I., & Diels, E. (2019). Live subtitlers: Who are they? *Linguistica Antverpiensia: New Series: Themes in Translation Studies*, 18, 101–129.
- Romero-Fresco, P. (2008). La subtitulación rehablada: palabras que no se lleva el viento. In Á. Pérez-Ugena & R. Vizcaíno-Laorga (Eds.), *ULISES: Hacia el desarrollo de tecnologías comunicativas para la igualdad de oportunidades* (pp. 49–73). Madrid: Observatorio de las Realidades Sociales y de la Comunicación.
- Romero-Fresco, P. (2009). More haste less speed: Edited versus verbatim respoken subtitles. *Vigo International Journal of Applied Linguistics*, 6(1), 109–133.
- Romero-Fresco, P. (2010). Standing on quicksand: Viewers' comprehension and reading patterns of respoken subtitles for the news. In J. Díaz Cintas, A., Matamala, & J. Neves (Eds.), *New insights into audiovisual translation and media accessibility* (pp. 175–194). Leiden: Brill. doi:10.1163/9789042031814_014
- Romero-Fresco, P. (2011). *Subtitling through speech recognition: Respeaking*. Manchester: Routledge.
- Romero-Fresco, P. (2012a). Quality in live subtitling: The reception of respoken subtitles in the UK. In A. Remael, P. Orero, & M. Carroll (Eds.), *Audiovisual translation and media accessibility at the crossroads* (pp. 109–131). Leiden: Brill.
- Romero-Fresco, P. (2012b). Respeaking in translator training curricula: Present and future prospects. *The Interpreter and Translator Trainer*, 6(1), 91–112. doi:10.1080/13556509.2012.10798831
- Romero-Fresco, P. (2016). Accessing communication: The quality of live subtitles in the UK. *Language & Communication*, 49, 56–69. doi:10.1016/j.langcom.2016.06.001
- Romero-Fresco, P. (2018). Respeaking: Subtitling through speech recognition. In L. Pérez-González (Ed.), *The Routledge handbook of audiovisual translation* (pp. 96–113). doi:10.4324/9781315717166-7
- Romero-Fresco, P., & Martínez, J. (2015). Accuracy rate in live subtitling: The NER model. In R. Baños Piñero & J. Díaz-Cintas (Eds.), *Audiovisual translation in a global context: Mapping an ever-changing landscape* (pp. 28–50). doi:10.1057/9781137552891_3
- Romero-Fresco, P., & Eugeni, C. (in press). Live subtitling through respeaking. In Ł. Bogucki & M. Deckert (Eds.), *The Handbook of audiovisual translation and media accessibility*. London: Palgrave Macmillan.
- Rossignol-Farjon, A., & Cimino, F. (2016). Access services pan European survey 2016. Retrieved from

Romero-Fresco, P., Melchor-Couto, S., Dawson, H., Moores, Z. & Pedregosa, I. (2019). Respeaking certification: Bringing together training, research and practice. *Linguistica Antverpiensia, New Series: Themes in Translation Studies*, 18, 216–236.

<https://gvadata.ch/access-services-pan-european-survey-2016>

Russello, C. (2010). *Teaching respeaking to conference interpreters*. Retrieved from Intersteno Education Committee Archive <https://www.intersteno.it/materiale/ComitScientifico/EducationCommittee/Russello2010Teaching%20Respeaking%20to%20Conference%20Interpreters.pdf>

Sperber, M., Neubig, G., & Fügen, C., Nakamura, S., & Waibel, A. H. (2013). Efficient speech transcription through respeaking. *InterSpeech* (14th Annual conference of the International Speech Communication Association), 1087–1091.

Zick, R. G., & Olsen, J. (2011). Voice recognition software versus a traditional transcription service for physician charting in the ED. *The American Journal of Emergency Medicine*, 19(4), 295–298. doi:10.1053/ajem.2001.24487

¹ This research has been conducted within the frameworks and with the support of the EU-funded projects ILSA: Interlingual Live Subtitling for Access (2017-1-ES01-KA203-037948) and EASIT: Easy Access for Social Inclusion Training (2018-1-ES01-KA203-050275), as well as the Spanish-government funded projects Inclusión Social, Traducción Audiovisual y Comunicación Audiovisual (FFI2016-76054-P), EU-VOS. Intangible Cultural Heritage. For a European Programme of Subtitling in Non-hegemonic Languages' (Agencia Estatal de Investigación, ref. CSO2016-76014-R) and the Galician-government funded project Proxecto de Excelencia 2017 Observatorio Galego de Accesibilidade aos Medios (GALMA).

² <https://ltaproject.eu/>