Using evaluation criteria and rubrics as learning tools in subtitling for the D/deaf and the hard of hearing

Irene De Higes Andino and Beatriz Cerezo Merchán

ABSTRACT
This paper focuses on the use of evaluation criteria and rubrics as tools for training in an audiovisual translation mode: subtitling for the D/deaf and the hard of hearing (SDH). Following an introduction to the current conception of and research carried out on quality assurance and assessment both in the translation industry and academia, a brief overview of the assessment processes and instruments used in SDH university training and in the professional sector is provided. This is followed by a case study of SDH assessment practices in Spain. Two questionnaires were filled in by five service providers and 12 university trainers, and the data gathered has enabled a better understanding of how assessment is currently being performed. These results, together with the Spanish SDH standards and a competence-based training approach, informed the design of two assessment instruments: a set of evaluation criteria and a rubric, which are both presented here. We also present a series of graduated tasks which are designed to enable students to gain an understanding of and be able to use the instruments to assess SDH.

1. Introduction
The study of translation training inevitably involves the study of assessment, as it is an essential element in any teaching and learning process. To date, most of the proposals regarding assessment in Translation Studies (TS) focus on summative and product assessment (Waddington 2000 and Galán-Mañas and Hurtado Albir 2015 provide excellent overviews). This focus on summative assessment has led to diagnostic and formative assessment being under researched and limiting the understanding of how this type of assessment manages correcting translations. The analysis of translation errors has been a long-established practice in TS with many assessment proposals being limited to either translation error categories (e.g. Delisle 1993) or to correction categories and weighting criteria (e.g. Hurtado Albir 1999; Adab 2000).

In the last few decades there has been a shift in how assessment is perceived in translation training to focus on the processes involved and not just on the end product. Current pedagogical approaches advocate competence-based training, in which assessment
is an essential element of curricular design that is directly linked to learning objectives, competences and tasks.

Under this approach, assessment becomes an essential tool for trainers, who can evaluate how students progressively develop competences through a variety of tasks and assessment instruments – including diagnostic questionnaires, surveys of translation knowledge, reflective diaries, translations with commentaries, portfolios, rubrics, recordings of professionals or students translating (Galán Mañas and Hurtado Albir 2015, 70–72; Huertas Barros and Vine 2015, 24).

Assessment is also a useful tool for students, who can play an active role in their own learning, perform assessment themselves (self- and peer-assessment) and be more aware of their learning process, their strengths and weaknesses. Although both academia and industry are investing much effort in designing models to assess and assure the quality of the translation process and product, assessment is a broad and complex area in which much more remains to be done.

On the one hand, more proposals for assessments which are specifically designed for training settings are needed. Teaching-oriented assessment proposals ‘are few and far between’ (Galán-Mañas and Hurtado Albir 2015, 67), which given the complexity of measuring translation competence (Beeby 2000; Hurtado Albir 2017) might be expected.

Also, more empirical studies of assessment instruments and evaluation criteria are needed. There is still much debate about how subjective the application of evaluation criteria to assess translations can be, and ‘finding valid assessment criteria and making more objective judgements about what a good or bad translation is remains a complex matter’ (Huertas Barros and Vine 2015, 23).

Within the translation industry, assessment mainly concentrates on the end product and revolves around the concept of ‘quality assurance’ (QA), which consists in ‘the correction and/or amendment of a professional translation that is carried out by someone other than the translator before the translation is delivered to the customer’ (Rasmussen and Schjoldager 2011, 92). National and international standards (e.g. ISO 2015) and best practice guides outline the processes for producing quality-assured translations. Metrics and tools to assess the quality of translations can also be found in the translation sector (Sánchez-Gijón 2014).

However, conceptions of quality often differ between academia and industry, leading to debates about how to ‘bridge the gap’ between the two conceptions. We need more specific studies on quality criteria and assessment processes from both sectors to be more informed about how these criteria and processes are performed in both settings. This paper investigates assessment practices in subtitling for the D/deaf and the hard of hearing (SDH), which is a form of subtitling that includes dialogue and text on screen and that integrates those elements that the D/deaf or hearing-impaired audience cannot easily perceive (see Section 2). Our research on assessment practices both in the industry and academia is a first step to design a set of instruments that could be implemented in both sectors.

To this end, four objectives are set out:

(1) To investigate the quality assurance processes and instruments in the SDH industry in Spain.
(2) To investigate the assessment processes and instruments for SDH training in Spanish universities.
(3) To design a set of evaluation criteria and a rubric which could be used in academia and industry.
(4) To elaborate interrelated and graduated tasks in which these instruments may be actively used by students.

Following this introduction to the current conception of and research carried out on assessment in translation, sections 2 and 3 focus specifically on SDH and on assessment processes and instruments used in SDH university training and in the professional sector. In section 3, a case study carried out to gain a better understanding of assessment practices in academia and industry in Spain is presented. In section 4, two assessment instruments are proposed on the basis of the data gathered, the Spanish SDH standards and a competence-based training approach. These assessment instruments include a set of evaluation criteria (i.e. a list of correction criteria and the description of the potential errors that can be categorised under each criterion) and a rubric (i.e. a scoring guide with elements to be appraised and performance levels). Section 5 suggests a number of interrelated and graduated tasks for students to actively engage with them in their own learning process. The tasks outlined in this section are aimed at helping the students fully understand the assessment criteria and the marks given and being able to apply these to their own work or that of their peers. Finally, conclusions are drawn, and further research avenues are suggested.

2. Subtitling for the D/deaf and the hard of hearing

Subtitling for the D/deaf and the hard of hearing is a form of subtitling that integrates those elements that the D/deaf or hearing-impaired audience cannot easily perceive, such as contextual information, sound effects, and music with the dialogue and text on screen (Díaz Cintas 2006, 6).

Over the last two decades the increasing awareness of issues of equality for all citizens, and the efforts of governments to ensure equal access to information, along with the greater flexibility of audiovisual media and formats, have boosted the visibility and the professional and academic development of SDH.

Since its appearance in 1973 on American television, SDH has continued to evolve and grow, and many countries already offer SDH services not only on television but also on DVD, in cinemas, theatres or opera houses. Therefore, according to their mode of production, subtitles can be categorised as:

- Live, if they ‘are created (sometimes using speech recognition technology) and cued, with a delay, during transmission’ (Gottlieb 2015, 19).
- Semi-live, if ‘created by the subtitler before broadcast and cued during transmission’ (Gottlieb 2015, 19).
- Pre-recorded, when they are created and/or cued by the subtitler before broadcast.

Although this communication support service has traditionally been intralingual (the first subtitles for the deaf and the hard of hearing on TV were intralingual, Zárate 2014, 41) and much SDH continues to be so, it is also possible to find interlingual SDH, as pointed out by Neves (2008, 185).
Legislation in European countries, encouraged by Directive 2007/65/EC of the European Parliament and of the Council of 11 December 2007 (European Union 2007), recommends a certain percentage of TV emissions be subtitled for the D/deaf and the hard-of-hearing, along with a minimum number of hours of TV transmission to be audio described and signed per week.\(^2\)

In recent years, the number of subtitled programmes has steadily risen. For example, since 2008, the BBC subtitles 100% of its programmes on its main channels.\(^3\) And even in those countries with no specific quota, like Germany, an increase in the quantity of subtitles is evident. German ARD channel, for example, subtitled 16.3% of its output in 2005, this figure rising to 92.2% in 2015.\(^4\) In Spain, legislation established a target of SDH provision for 90% of public service broadcast programs and 75% of commercially broadcast programs by the end of 2013. Spanish TV channels have not yet achieved their target, but the average percentage of subtitled programmes on public and private TV channels fluctuated around 75% between January to June 2017.\(^5\)

The academic world has responded to this exponential growth of SDH provision and in recent decades training in SDH has rapidly been introduced in universities. The urgent need to train future professionals in the field of SDH has been felt in many countries, and universities around the world, in countries such as Spain, the United Kingdom, France, Belgium, Portugal or Italy have begun to offer SDH training both at undergraduate and postgraduate level on their translation and interpreting courses.

The focus of SDH academic research has been on a range of topics including the history and evolution of SDH in different countries, the description of SDH practices and conventions or the study of viewers’ preferences (Díaz Cintas, Orero, and Remael 2007; Jiménez Hurtado 2007; Matamala and Orero 2010; Romero-Fresco 2015). In addition, the growing interest in the study of SDH has led to a number of PhD dissertations being published on SDH, ranging, for instance, from studies of SDH parameters (Neves 2005; Arnáiz 2012) to reception studies with children (Zárate 2014; Tamayo Masero 2015).

The European Union has funded research on media accessibility, encouraging academic partners, broadcasters, research institutes and SMEs to work together. SDH has been explored in the following research projects: DigitalTV4All (a project led by Brunel University to facilitate the provision of access services on digital television, 2008–2010) and HBB4ALL (a project led by Universitat Autònoma de Barcelona on media accessibility for all citizens in the connected TV environment, 2013–2016).

3. Assessment in SDH

In order to ensure not only quantity but also quality in SDH for pre-recorded, semi-live and live TV programmes, sets of guidelines and quality standards have been developed in different countries by a range of institutions including D/deaf associations, broadcasters, regulators, companies, universities. In Spain, the basic standards for the provision of SDH have been set out in the standard UNE 153010:2012 Subtitulado para personas sordas y personas con discapacidad auditiva (AENOR 2012). In the UK the Code on Television Access Services (Ofcom 2015) and the BBC Online Subtitling Editorial Guidelines (BBC 2009) have been developed; in France quality standards have been established by the Charte relative à la qualité du sous-titrage à destination
des personnes sourdes ou malentendantes (CSA 2011); and in the German-speaking countries public service broadcasters have recently agreed on a set of subtitling standards (ARD, ORF, SRF and ZDF 2015).

There is no doubt that quality assurance in SDH, as with QA in subtitling for hearing audiences, is necessary to ensure the good reception of subtitles by the audience. Despite time pressure, which is common in the television industry, QA is also important for subtitling companies not only due to additional costs if the subtitled product is returned, but also because of a potential loss of professional credibility (James 2001, 155).

However, despite the exponential growth of the profession, the drafting of SDH standards and the increase of training courses and research, there is not much information to date on how QA takes place in the SDH industry or how SDH is assessed in graduate and postgraduate courses.

Drawing on the results of the previously mentioned project DTV4ALL, Romero-Fresco (2015) has assessed the quality of SDH in Europe as a combination of three factors: what viewers think about SDH, how they understand the subtitles and how they view them. Research on QA in SDH has mainly focused on live (and semi-live) SDH, carried out by different methods such as stenotype or re-speaking. Initially the WER (Word Error Rate) model – used to evaluate automatic speech recognition systems – was applied. In 2010, the Carl and Ruth Shapiro Family National Center for Accessible Media (NCAM) started the Caption Accuracy Metrics Project to quantify errors in stenocaptions (Media Access Australia 2014). And from 2015 the NER model has been used by regulators, broadcasters, companies and training institutions in countries such as Spain, France, Italy, Switzerland, Germany, Belgium and Australia to evaluate re-spoken subtitles after they have been broadcast (Romero Fresco and Martínez 2015).

3.1. Assessment in SDH in Spain: a case study

As previously stated, despite the growth of the profession, the increase of training and research in the field of SDH and the drafting of quality standards, there is still little information on how assessment practices are carried out. Therefore, to be able to design a meaningful assessment proposal, we undertook a research project in Spanish professional and academic settings to gain a better understanding on how assessment is being carried out.

Two different online questionnaires were designed and distributed among SDH service providers and SDH university trainers in Spain in October 2016. Both questionnaires were anonymous and included 12–15 open-ended and closed-ended questions. The questionnaires incorporated some general questions about the company or the training course, and also some specific questions about assessment procedures and instruments within both these settings.

Respondents were recruited by email. We contacted universities and companies working in the field of SDH in Spain, and we gathered answers from five SDH service providers and 12 universities offering specific training in SDH. Although this may seem a small sample size, the total population of possible respondents was not significantly larger. We gathered data from 12 universities offering SDH training and five
of the nine most representative providers of SDH in Spain, i.e. providers with the higher annual work volume.

In the following sections, we present and interpret the quantitative and qualitative data gathered in the two questionnaires.

3.1.1. Quality assurance in Spanish SDH providers

Five SDH service providers based in Spain answered the questionnaire. To contextualise their relevance in the Spanish SDH industry, they were asked about the percentage of SDH in their annual work volume; three of them indicated that SDH commissions exceeded 75% of their work volume. One service provider stated that SDH was less than 25% of their work volume and another one estimated it to be between 25 and 49%.

Two service providers have no in-house subtitlers and always work with freelancers (1 to 5 in one case and 11 to 20 in another case). The other three providers have between 1 to 5 in-house subtitlers and also work with freelancers (two have between 11 to 20 collaborators and the other provider uses more than 20 freelance subtitlers). As shown in Table 1, most service providers (A, B, C & D) offer pre-recorded SDH for TV, but only one (B) also subtitles live and semi-live.

All five service providers state that quality assurance is part of the process of SDH. The people in charge of carrying out QA vary. Two service providers indicate that QA is performed in-house (either by one in-house subtitler or by the project managers). Four service providers usually request QA from freelancers.

In relation to the percentage of texts normally reviewed in the QA process, three SDH service providers state that they always review 100% of the intralingual translation, while the other two providers stated that they do not always review the entire translation. When asked about the amount of text that they normally review, one of these two providers did not know or preferred not to answer, while the other responded that it reviewed 50–75% of the translation.\footnote{All five service providers state that quality assurance is part of the process of SDH. The people in charge of carrying out QA vary. Two service providers indicate that QA is performed in-house (either by one in-house subtitler or by the project managers). Four service providers usually request QA from freelancers.}

Only three responses were gathered for the open question on how the QA process is carried out in the agency, including phases, timing, materials, etc. The fact that two providers did not respond might either suggest that the question was too broad and, therefore, difficult to answer, or that they did not know or preferred not to answer. In line with the conclusions gathered by Rasmussen and Schjoldager (2011), the specific guidelines or protocols used by translation service providers are difficult to access, even for research purposes. Likewise, our respondents only described the process in general terms. Two SDH service providers focus on the QA process of subtitles made for pre-
recorded audiovisual products before distribution. One of these companies just reviews different elements such as spelling or reading speed. The other provider explains that a third person reviews the subtitler’s work following two steps: firstly, he/she checks some parameters using an automatic checker, and secondly, he/she exhaustively reviews all the subtitles together with the audiovisual material. The other service provider explains the QA process for live and semi-live subtitles, which takes place after broadcasting. This answer was discarded as our proposal focuses on pre-recorded subtitles.

The following two questions from the survey referred to the use of specific programmes or instruments to systematise QA and which elements these included. None of the five SDH service providers uses specific tools (apart from subtitling software) to systematise the QA process. The five providers confirmed subtitlers usually receive feedback after the review process.

### 3.1.2. Assessment in Spanish university SDH courses

A total of 17 questionnaires were gathered from 12 Spanish universities. To contextualise the courses on SDH, trainers were asked about what level their courses are offered at, the forms of instructional interaction and the content of courses followed by more specific questions about SDH tasks and assessment.

According to respondents, 47.1% (8) of the courses are taught at undergraduate level and 52.9% (9) at postgraduate. Face-to-face learning is the preferred mode (76.5%, 13) of teaching at both levels, whereas blended and virtual training modes were also used (17.6%, 3 and 5.9%, 1, respectively) but only at postgraduate level. With regards to the content of these courses, approximately half of them (52.9%, 9) focus on SDH as well as on other audiovisual translation modes. Just over a third of courses only focus on SDH (35.3%, 6), while fewer courses focus on SDH and AD (11.8%, 2) and these are mainly taught at postgraduate level. Due to this heterogeneity, the number of hours devoted to SDH range from 2 to 100 hours per course.

In addition, we asked about the tasks usually undertaken in SDH training, which are shown in Table 2. Two postgraduate courses also included additional tasks, such as interlingual SDH, transformation of conventional subtitles into SDH and creation of SDH following standards from other countries or considering different audiences (having different grades of deafness, different ages), different genres or different media.

| Table 2. Tasks undertaken in SDH training in Spain. |
|-----------------------------------------------------|
| **Tasks**                                           | **No. of courses** (Percentage) |
| Reading of supporting material (UNE standard, international standards, research papers, etc.) | 16 (94.1%) |
| Full subtitling (the student applies all SDH conventions) | 16 (94.1%) |
| Revision of SDH (to identify mistakes) | 14 (82.4%) |
| Partial subtitling (the student focuses on one characteristic of SDH) | 14 (82.4%) |
| Contrastive exercises (critical analysis of professional SDH) | 12 (70.6%) |
| Synthesis and reformulation tasks | 12 (70.6%) |
| Discussions | 9 (52.9%) |
| Analysis of original texts | 8 (47.1%) |
| SDH reports (SDH with commentaries) | 5 (29.4%) |
| Reading of information sheets | 3 (17.6%) |
| Others | 2 (11.8%) |
As regards assessment practices, most trainers carry out initial (i.e. diagnostic), continuous and final assessment – i.e. students are assessed before, throughout and at the end of the learning process. One trainer (5.9%) opts for initial and final assessment, two (11.8%) use all three types of assessments and seven (41.3%) use continuous and final assessment. One trainer (5.9%) only uses final assessment and six (35.4%) only use continuous assessment.

When asked about the person assessing the different tasks, although five indicate that only the trainer assesses students, the majority involve other assessors in the evaluation process:

- Teacher-, peer-, self- and expert-assessment: 1 (5.9%)
- Teacher- and peer-assessment: 2 (11.8%)
- Teacher-, peer- and self-assessment: 4 (23.6%)
- Teacher- and self-assessment: 5 (29.5%)
- Teacher-assessment: 5 (29.5%)

Trainers were also asked to describe their assessment tool(s). Even those following the Spanish standards used differing criteria and ratings. Eleven (68.8%) use a set of evaluation criteria, two (12.5%) use a rubric and three (18.8%) do not make use of any instrument for evaluation. The assessment tools have been mainly designed by the teachers themselves but, on some occasions, they are designed in collaboration with other teachers or they are based on the criteria of a research group or the degree board.

Of the trainers who use assessment tools, 92.3% explain them to the students before using them in assessment. However, only 58.3% (7) prepare tasks working with assessment tools as part of the training (formative assessment). Those tasks may consist of discussing standards, assessing SDHs presented in class by other students, analysing pieces of professional SDHs, or self-assessing their own tasks.

The participants’ responses reveal that both university trainers and service providers attach considerable importance to assessment and quality assurance in SDH. All service providers confirmed that QA is part of the process of SDH and is carried out by someone other than the subtitler, and that the subtitler always receives feedback. However, none of these companies use specific instruments to measure quality in SDH, there are no agreed standards or protocols on how to measure quality and they tend to focus on the concept of translation error.

SDH trainers seem to be using innovative student-centred assessment practices involving a variety of instruments, assessors and assessment. Most trainers stated that they use specific instruments when assessing translations as products; sets of evaluation criteria are the most frequently used assessment tools compare to rubrics which are only used by two trainers. Although all the respondents stated their instruments complied with the Spanish SDH standards, the analysis of the few instruments that were shared with us revealed that there is no consensus on the scales and categories of translation errors they include.

Our study demonstrates that there is a shortage of instruments to assess SDH. Besides, there is no consensus on the procedures used to assess quality either within academia or among service providers. SDH would benefit from sharing assessment
instruments and procedures to reach more consensus. The proposal presented in the following sections aims to be a contribution in this direction.

4. A proposal of assessment instruments for SDH training

Despite the increase in training and research in SDH in recent years, there are virtually no studies focusing on the curricular aspects of training in this area, with a few exceptions such as Díaz Cintas (2006) on the description of the subtitler’s professional skills or Pereira and Lorenzo (2007), Neves (2008), Romero-Fresco (2011), Cerezo Merchán and De Higes Andino (2013) and Tamayo (2016) on teaching methodologies for SDH training.

Translator trainers agree that students should be trained to produce market-ready audiovisual translations and that training should focus on evaluating the quality of the end product (i.e. the subtitles). Nevertheless, it is paradoxical that most research efforts have been related to quality of the subtitles and there is little research about how to train future SDH professionals, especially with regards to teaching methodologies and QA.

In the following sections, we propose two assessment tools (i.e. a set of evaluation criteria and a rubric). These two instruments and the tasks presented in section 5 are directly related to the general and specific competences taught on the two SDH courses at undergraduate level in Spain (see competences and tasks in section 5).

These instruments aim to help trainers assess both the students’ learning processes and the products of this process throughout the course. Students can use them as a guide in translation tasks or as a ‘critical framework’ (Adab 2000, 220) to explain their translation strategies and decisions.

During the academic year 2015/2016 students from Universitat Jaume I and Universitat de València (Spain) were asked their opinion on using a set of evaluation criteria in the classroom. Most students completely agreed that the evaluation criteria were useful in helping them become more aware of their difficulties and also as a guide to their subtitling practice. Besides, students had the impression that assessment is more objective with a set of criteria.

4.1. Evaluation criteria

The set of evaluation criteria we present in this section (see Table 3) includes a list of correction criteria (SDH-specific conventions, as well as some general translation issues), together with their abbreviations and a description of the potential errors that can be categorised under each criterion. This list is grounded in SDH standards (mainly the Spanish UNE 153010:2012) and descriptive research on SDH (such as Arnáiz [2012] and Tamayo Masero [2015]) and can be adapted by trainers depending on the learning objectives.

This evaluation criteria tool is based on previous proposals such as Hurtado Albir’s (1999, 120) who focuses on errors related to comprehension of the source text, errors related to reformulation in the target language and errors of pragmatics; and Mossop’s (2014, 134–135) who suggests four parameters: transfer, content, language and presentation. Our criteria also draw on the approach taken by the research group TRAMA (Universitat Jaume I) in assessing other audiovisual translation modes.

This set of evaluation criteria has been used by the authors in their SDH classes at Universitat Jaume I and Universitat de València for formative and summative
| Abbreviation | Type of mistakes | Specific Items                                                                 |
|--------------|------------------|--------------------------------------------------------------------------------|
| F            | Form             | - Unnecessary spaces                                                            |
|              |                  | - Line over 37 characters                                                        |
|              |                  | - Incorrect line or subtitle segmentation                                        |
|              |                  | - Empty subtitle                                                                |
|              |                  | - Exceeding number of lines per subtitle                                         |
|              |                  | - Poor distribution of information (i.e.: a long first line and a short second line, or two lines when all the information would perfectly fit in one line) |
| POS          | Positioning      | Error in the positioning on the screen                                           |
| RS           | Reading speed    | Exceeding reading speed (recommended 15 characters per second [CpS] for adults and 12 CpS for children) |
| D            | Duration         | - Not respecting the minimum duration of subtitle (1 second)                   |
|              |                  | - Excessive duration of subtitles conveying music and sound effects             |
| SYN          | Synchronization  | 17 - Subtitle is cued in before or after the beginning of the subtitled element |
|              |                  | 17 - Subtitle is cued out before or after the end of the subtitled element       |
| OVER         | Overlapping      | No minimum pause between subtitles (200 ms)                                     |
| ID           | Character identification | - Error in the use of colours (one-time, misprint)                           |
|              |                  | - Error in the use or form of dash                                             |
|              |                  | - Error in the use or form of tag                                              |
|              |                  | - Incorrect colour from the limited range which can be used                    |
|              |                  | - Error in the selection of colours according to character role                |
| SFX          | Sound effects    | - Unnecessary sound effect (minor or major mistake)                            |
|              |                  | - Error in form                                                                |
|              |                  | - Error in lexical choice to describe the sound effect                          |
| CONTEX       | Contextual information | - Unnecessary contextual information (minor or major mistake)               |
|              |                  | - Error in form                                                               |
|              |                  | - Error in lexical choice to describe the contextual information               |
| MUS          | Music and songs  | - Error in the symbols to mark plot music                                      |
|              |                  | - Error in the form of plot or background music                               |
|              |                  | - Error in the lexical choice to describe plot or background music             |
|              |                  | - Unnecessary background or plot music                                          |
|              |                  | - Omission of necessary plot or background music                              |
| TYP          | Typography       | - Error in the use of italics                                                  |
|              |                  | - Triple dots to mark incomplete speech                                       |
|              |                  | - General typographic mistakes                                                 |

General items

| Abbreviation | Type of mistakes | Potential mistakes |
|--------------|------------------|--------------------|
| PRA          | Pragmatics       | Not adequate to subtitling brief and addressee                               |
| NSM          | Not same meaning | Ambiguous phrasing, omission of nuances, exaggeration                        |
| FM           | False meaning    | Incorrect meaning                                                   |
| NMS          | Not making sense | Difficult to understand, unclear phrasing, not making sense at all          |
| CT           | Contradiction    | Contradictory meaning                                               |
| EX           | Excess           | Incorporation of unnecessary or inexistent information                   |
| OM           | Omission         | - Unnecessary omission of linguistic code                               |
|              |                  | - Clip not fully subtitled                                              |

(Continued)
assessment. The set was used during the academic year 2016/2017 in combination with the rubric presented in section 4.2, to carry out final assessment.

4.2. Evaluation rubric

Assessment rubrics are scoring guides in which elements to be assessed and performance levels are described with a view to evaluating students’ performance. Rubrics are useful tools for any assessment task and, when required, they may also include numerical ratings (Hurtado Albir 2015, 72). Constructing a rubric can be time-consuming, since articulating criteria and grading descriptors is not an easy task. However, this difficulty is compensated for by the fact that rubrics provide both learners and teachers with specific feedback about the learning process (Mertler 2001, n.d.). Therefore, placing the rubric at the students’ disposal is essential.

The rubric presented in Table 4 is based on the set of evaluation criteria given above and incorporates aspects of Angelelli’s (2009, 40–41), Hurtado Albir’s (2015, 21) and Pavani’s (2016, 423–424) work in this area. In line with the idea of ‘minimal’ and ‘optimal’ quality by Gummerus and Paro (2001, 138), this rubric includes five performance levels (i.e. excellent, good, fair, poor and unacceptable) and describes the elements to be assessed in each of them.

Following Pavani’s (2016, 423–424) suggestion, the four sub-components i.e. formulation of the source text sense; writing; adequacy of translation mode (mainly relating to the
| GLOBAL SCORE | RATING SCALE | FORMULATION OF THE SOURCE TEXT SENSE (15%) | WRITING (25%) | ADEQUACY OF TRANSLATION MODE (30%) | PRESENTATION OF SUBTITLE ON SCREEN (30%) |
|--------------|--------------|--------------------------------------------|---------------|-----------------------------------|----------------------------------------|
|              |              | - Meaning                                  | - Spelling    | - Compliance with translation brief and addressee |
|              |              | - Information load                         | - Typography  | - Compliance with SDH conventions (based on UNE standard 2012:153010) |
|              |              | - Semiotics                                | - Vocabulary  | - Character identification         |
|              |              | - Linguistic variety                       | - Grammar     | - Sound effects                    |
|              |              |                                           | - Textual (cohesion and coherence) | - Contextual information            |
|              |              |                                           | - Style       | - Plot and background music        |
|              |              |                                           |               |                                   |
|              |              |                                           |               |                                   |
|              |              |                                           |               |                                   |
|              |              |                                           |               |                                   |
conventions of SDH regarding the elements to be subtitled); and representation of subtitle on screen have each been assigned a different percentage. These separate percentages are used to calculate the final score. Given the fact that our students are in the last year of their degree, more importance is given to SDH-specific methodological, strategic and professional competences, i.e. ‘adequacy of translation mode’ and ‘representation of subtitle on screen’. The sub-component ‘formulation of the source text’ has been given a lower weighting since the SDH taught in our course is mainly intralingual and there are therefore few problems of transfer.21

This rubric is still at an early stage of development and was tested, together with the set of evaluation criteria presented above, in class and improved during the academic year 2016–2017 at Universitat Jaume I and Universitat de València.

Due to space constraints, the description of each performance level is not presented in this article.22

5. A comprehensive assessment proposal for SDH training: using assessment instruments as learning and teaching tools

The pedagogical approach to SDH training presented in this article suggests that both teachers and students actively engage with the two assessment instruments described above. In the survey of teachers, we found that these types of tools are mainly used for summative assessment. However, our experience has shown that it is also possible to use them for formative assessment and that they are useful instruments in the development of competences at different stages of the learning process. Therefore, we believe that they are not only essential tools for the teacher to conduct summative assessment, but are also effective guides for clarifying what is expected of students, helping to provide feedback and rating the students’ work. Moreover, we consider that students can also actively engage with them in their own learning process.

Our proposal is part of a competence-based training framework that goes beyond the appraisal of translations and encompasses the evaluation of the learning process. These competences are:

General competences (Hurtado Albir 2015, 11):

(1) Learning strategically, autonomously and continuously (GC1).
(2) Working in teams (GC2).
(3) Analysing and synthesising (GC3).
(4) Thinking critically (GC4).
(5) Taking decisions (GC5).

Specific competences (adapted from Hurtado Albir 2015, 11):

(1) Applying the methodological principles and strategies necessary to use SDH conventions appropriately (methodological and strategic competence, SC1).
(2) Activating encyclopaedic, cultural and thematic knowledge to solve SDH-related problems (extralinguistic competence, SC2).
(3) Managing/applying basic knowledge about the subtitler’s profession (professional competence, SC3).
(4) Managing basic documentary resources to solve SDH-related problems (instrumental competence, SC4).
(5) Using appropriate strategies to solve intralingual and intersemiotic translation problems in different text genres (integration of competences, SC5).

The ‘holistic translation assessment method’ (Waddington 2000) we propose links assessment tasks to competences and performance levels, and uses different instruments and tasks to assess the learning process and the general and specific competences developed throughout such a process. Following Hurtado Albir’s approach (Hurtado Albir 2015), these interrelated graduated tasks prepare students for a complex final task which shows students have acquired the relevant competences and have met the learning objectives.

The tasks described below are not only subtitling tasks, but also tasks related to acquiring knowledge about and preparing students for SDH. The tasks are designed to use evaluation criteria and a rubric not only as assessment, but also as learning tools.

**Task 1. Familiarising students with the Spanish standards UNE 153010:2012 Subtitulado para personas sordas y personas con discapacidad auditiva**

- Linked to GC1 and GC3; SC3 and SC4
- Formative assessment
- Self-assessment

Students are asked to read the Spanish SDH standards and fill in Table 5. To assess this task, students are given an answer key which is discussed in class.

**Task 2. How to assess SDH**

- Linked to GC1, GC3 and GC4; SC2, SC3 and SC4

| Table 5. Familiarising students with the Spanish UNE standards. |
|---------------------------------------------------------------|
| **Content of subtitles**                                      |
| Elements included in the UNE standard | When should this element be subtitled? | How should it be subtitled? |
| **Timing of subtitles**                                      |
| Elements included in the UNE standard | When should this element be subtitled? | How should it be subtitled? |
| **Form of subtitles**                                        |
| Elements included in the UNE standard | When should this element be subtitled? | How should it be subtitled? |
Once they are acquainted with the Spanish UNE standards, the students are asked to consider how to assess SDH quality. They should first discuss in groups which SDH criteria they would include in a QA tool and then design a draft version of evaluation criteria. The proposals are discussed in class.

Students are then presented with a set of evaluation criteria prepared by the teacher, which may be altered in line with some of the students’ suggestions. Where it is the case that there is a QA instrument from the SDH industry, students would benefit from analysing this instrument before being introduced to the criteria to be used in class.

**Task 3. Getting to know other standards**

- Linked to GC1 and GC4; SC3 and SC4
- Formative assessment
- Teacher- and peer-assessment

As the Spanish UNE standards are rather vague on when and how to describe some elements, such as lyrics, students are asked to read other European standards and in pairs answer the following questions:

- Are there any differences in comparison with Spanish UNE standards?
- Which information is more complete regarding...?
- Would it be appropriate to elaborate European or even international standards? What advantages or disadvantages can you think of?
- Would you add some information from the international standards to the set of evaluation criteria presented in task 2?

Answers are compared in class. The trainer together with the students would decide whether the set of evaluation criteria should be modified.

**Task 4. Partial analysis of professional subtitles**

- Linked to GC1, GC2 and GC3; SC1 and SC2
- Formative assessment
- Peer- and self-assessment

Students are asked to analyse how one specific source text element (e.g. background music) has been translated in subtitles shown on TV or distributed on DVD. The material given could be a video clip or some stills. If possible, they may compare more than one set of professional subtitles for the same video clip. After the analysis, students discuss with peers or they could be given an answer key to self-evaluate their task. If it the case that students have been asked to analyse only one specific source text element, this task may be repeated during the course to analyse other elements of SDH separately (e.g. plot music, sound effects, etc.).
**Task 5. Partial subtitling**

- Linked to GC2 and GC5; SC1 and SC5
- Formative or summative assessment
- Peer-assessment

Students are asked to subtitle a video clip focusing specifically on one element, for example, music. This video clip would be an original audiovisual production in Spanish or a foreign production dubbed into Spanish.

Once they have completed their own subtitling, students are then asked to evaluate a peer’s subtitling in two stages. First, they evaluate the subtitles as a written text and then, they evaluate the subtitles in conjunction with the audio-visual material. The evaluation criteria already presented in class would be used to assess quality. The subtitler and proof-reader (peer assessor) would then discuss problems and suggest solutions. This task may be used more than once during the course, focusing on a different element each time.

**Task 6. Comprehensive subtitling**

- Linked to GC4 and GC5; SC1, SC3 and SC5
- Formative or summative assessment
- Self- and teacher-assessment

Students are asked to subtitle a more complex video clip containing all elements to be included in SDH to demonstrate whether the necessary translation competences have been developed and the learning objectives achieved. This final task could be used as a summative assessment, but it could also be used as a self-evaluation task, as the translation should be accompanied by a translation commentary. This commentary would describe and evaluate the processes followed while subtitling, thus demonstrating the students’ awareness of their own translation processes and also their learning process. The students would use the rubric designed by the teacher to assess SDH, which they have already been introduced to, to carry out the self-evaluation, before submitting the translation for teacher assessment.

**Task 7. What did I learn?**

- Linked to GC4 and SC1
- Formative assessment
- Self-assessment

Students are asked to reflect on their learning process and the use of rubrics and evaluation criteria to assess SDH quality by answering the questions listed in Table 6 (based on Hurtado Albir 2015, 235–236).
Table 6. What did I learn?

- What conclusions did you draw on the application of SDH standards in Spain?
- What kind of problems did you find when subtitling and using these standards?
- What conclusions did you draw on the use of the set of evaluation criteria and the rubric?
- What new strategies did you learn to try to solve the problems you encountered?
- What aspects did you improve?
- What aspects do you still need to improve?
- How do you think you can improve them?

Rate from 1 to 10. I am able to:

- Identify elements to subtitle:
- Apply Spanish SDH standards:
- Identify problems when subtitling:
- Solve basic subtitling problems:
- Apply appropriate strategies to subtitle:
- Assess SDH quality with the set of evaluation criteria and the rubric:

6. Conclusions

This research presents a set of evaluation criteria and a rubric as tools for training in subtitling for the D/deaf and the hard of hearing. Data from the industry and trainers were gathered to examine which assessment processes and instruments are used in the SDH professional sector and university training in Spain.

Our case study reveals that SDH service providers do monitor quality but have not created any specific instrument for this purpose. They consider some criteria and use subtitling programmes to automatically check technical elements when undertaking QA. Most SDH trainers in Spanish universities use assessment tools, especially sets of evaluation criteria. However, only some of them include tasks to work with these tools as part of formative assessment.

Based on the data gathered, the Spanish SDH standards and a competence-based training approach, this paper proposes two assessment instruments and a number of interrelated and graduated tasks for students to actively engage with them in their own learning process. This assessment methodology is a work in progress. During the academic year 2016/2017, assessment combined the set of evaluation criteria described in section 4.1 with the rubric presented in section 4.2. We believe that this assessment methodology is not only useful for trainers as a guide to assess SDH and to promote self- and peer-assessment, but also for students to assimilate SDH conventions and to serve as a framework to explain their subtitling strategies and decisions.

During the academic year 2017/2018 we will continue validating and improving the assessment tools and tasks (i.e. making changes to the timing, sequencing and required previous knowledge of the tasks). At the end of the course students will also be asked about the usefulness of the tasks and tools and their level of satisfaction.

Results from this research will be shared with SDH service providers and university trainers, and they will be asked about the usefulness and potential application of the tools in their professional settings.

To conclude, this research corroborates the need for more empirical studies which aim to validate similar assessment instruments used in academia and industry by trainers, students, professionals and also, as has been suggested by CNMC (Comisión...
Nacional de los Mercados y la Competencia) (2016), potential institutions in charge of quality control. In line with Huertas Barros and Vine (2015; 2017; 2018), we believe that academia can help industry to ensure quality and, conversely, industry can help academia to assess more effectively. If the gap is bridged, SDH quality will be improved and, as a consequence, this improved quality will also be of benefit to the end users, i.e. the D/deaf and the hard of hearing.

Notes

1. We regard the task as ‘the unit on the basis of which the learning process is organised’ and distinguish between ‘preparatory tasks and final tasks, with the former laying the groundwork for the performance of the latter’ (Hurtado Albir 2015, 263).
2. Audio description and sign language are accessibility modes outside the remit of this paper.
3. http://www.bbc.co.uk/pressoffice/pressreleases/stories/2008/05_may/07/subtitling.shtml. Hyperlinks are not accompanied by access date. All URL addresses included in this paper are available in June 2017.
4. http://www.schwerhoerigen-netz.de/RATGEBER/RUNDFUNK-FERNSEHEN/PDF/ut-stastistik-2015.pdf.
5. http://www.cesyas.de/television/seguiamiento.
6. Re-speaking is ‘a technique in which a re-speaker listens to the original sound of a (live) program or event and re-speaks it, including punctuation marks and some specific features for the deaf and hard-of-hearing audience, into a speech recognition software, which turns the recognized utterances into subtitles displayed on the screen with the shortest possible delay’ (Romero-Fresco 2011, 1).
7. ‘The NER model was introduced initially in Romero-Fresco (2011) and further developed in Romero Fresco and Martínez (2015). (...) N is the number of words in the subtitles. E stands for edition errors[.] (...) R stands for recognition errors’ (Romero-Fresco 2016, 59).
8. Questionnaires can be retrieved in Spanish from: https://goo.gl/forms/51h809WJBoSLHXS2 (survey of SDH teachers) and https://goo.gl/forms/Y6Yi5v95l3mjzZAx1 (survey of SDH service providers).
9. Trainers were asked to fill in as many questionnaires as courses in SDH they taught. Therefore, some of them filled in more than one questionnaire. Hence, the number of questionnaires completed by trainers (17) is higher than the number of universities (12).
10. At the moment of preparation of this paper (November 2016), 13 degrees offered SDH training at undergraduate level and five universities at postgraduate level. We gathered data from eight undergraduate degrees and four at postgraduate level.
11. According to Rasmussen and Schjoldager (2011, 101), the main reasons for not submitting all translations to revision are increasing pressures on delivery deadlines, the existence of QA protocols in subcontracted translation companies and their reliance on the professionals in charge of the translation.
12. These steps are similar to the ones identified by James (2001) regarding subtitles for hearers: (1) spellcheck, (2), preliminary reading of the subtitled texts only, without a viewing of the videotape, and (3) full viewing of the subtitled programme.
13. One response has not been considered as the trainer explained that no assessment tool is used because students are not asked to subtitle for the D/deaf and the hard of hearing but only to analyse a piece of professional SDH.
14. The questionnaires addressed to students are available for consultation at: https://goo.gl/forms/T7yVULauXzigiMH12 (in Spanish).
15. As an exception, one student questioned its real validity, arguing that he/she felt translation solutions are subjective. This issue is also raised by some students in specific case studies on assessment practices (see Huertas Barros and Vine 2018).

16. The Spanish standards recommend a reading speed of 15 CpS with no distinction between addressees (adults or children). Based on research on reading and writing skills of D/deaf children, Arnáiz (2012), Zárate (2014), Tamayo Maser (2015) and Ofcom (Office of Communications) (2015) recommend more edited subtitles and a slower reading speed (approximately 12 CpS).

17. Subtitles should match the pace or the duration of the element to be subtitled. For the sake of synchronisation, subtitles appear with (i.e. are cued in to match) its onset or disappear (i.e. are cued out) with the end of the corresponding element.

18. Sound effects should be typed within brackets in white with an initial capital letter.

19. Contextual information should be typed within brackets in upper case in the colour assigned to identify that character.

20. Background music should be typed in the same form as sound effects. Plot music should be typed in white or in the colour used to identify the character singing.

21. Weightings assigned could vary depending on the stage of the learning process. Our proposal aims to evaluate a final subtitling task, such as the one suggested in Task 6 in Section 5.

22. The complete rubric is available at: https://goo.gl/eXDvqk.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Irene De Higes Andino http://orcid.org/0000-0002-2295-7909
Beatriz Cerezo Merchán http://orcid.org/0000-0002-2955-5590

References

Adab, B. 2000. “Evaluating Translation Competence.” In Developing Translation Competence, edited by C. Schäffner and B. Adab, 215–228. Amsterdam and Philadelphia: John Benjamins. AENOR (Asociación Española de Normalización y Certificación). 2012. Norma UNE 153010: Subtitulado para personas sordas y personas con discapacidad auditiva. Madrid: AENOR.

Angelelli, C. V. 2009. “Using a Rubric to Assess Translation Ability. Defining the Construct.” In Testing and Assessment in Translation and Interpreting Studies, edited by C. V. Angelelli and H. E. Jacobson, 13–47. Amsterdam and Philadelphia: John Benjamins.

ARD, ORF, SRF and ZDF. 2015. Untertitel-Standards Von ARD, ORF, SRF, ZDF. https://www.zdf.de/barrierefreiheit-im-zdf/untertitel-standards-von-ard-orf-srf-und-zdf-100.html

Arnáiz, V. 2012. “Subtitling for the Deaf and the Hard-of-hearing. Some Parameters and their Evaluations.” PhD diss., Universitat Autònoma de Barcelona.

BBC (British Broadcasting Corporation). 2009. “Online Subtitling Editorial Guidelines V1.1.” http://www.bbc.co.uk/guidelines/futuremedia/accessibility/subtitling_guides/online_sub_editorial_guidelines_vs1_1.pdf

Beeby, A. 2000. “Evaluating the Development of Translation Competence.” In Developing Translation Competence, edited by C. Schäffner and B. Adab, 185–198. Amsterdam and Philadelphia: John Benjamins.

Cerezo Merchán, B., and I. de Higes Andino. 2013. “Trabajo colaborativo y desempeño profesional: Un caso práctico en la clase de accesibilidad audiovisual.” Hikma 12: 65–85. doi:10.21071/hikma.v12i.
CNMC (Comisión Nacional de los Mercados y la Competencia). 2016. “Informe sobre el grado de cumplimiento de las obligaciones impuestas en materia de accesibilidad y presencia de las personas con discapacidad en los medios audiovisuales (2014–2015).” Madrid: CNMC. http://www.cesya.es/sites/default/files/articulos/Informe%20CNMC%20sobre%20grado%20cumplimiento%20accesibilidad%2C%202015oct16.PDF

CSA (Conseil Supérieur de l’Audiovisuel). 2011. Charte relative à la qualité du sous-titrage à destination des personnes sourdes ou malentendantes. Paris: CSA. www.csa.fr/content/download/20043/334122/file/Chartesoustitrage122011.pdf.

Delisle, J. 1993. La traduction raisonnée. Manuel d’initiation à la traduction professionnelle de l’anglais vers le français. Ottawa: Université d’Ottawa.

Díaz Cintas, J. 2006. Competencias profesionales del subtitulador y el audiodescriptor. Getafe: CESyA. http://www.cesya.es/sites/default/files/documentos/informe_formacion.pdf.

Díaz Cintas, J., P. Orero, and A. Remael, eds. 2007. Media for All. Subtitling for the Deaf, Audio Description and Sign Language. Amsterdam: Rodopi.

European Union. 2007. Directive 2007/65/EC of the European Parliament and of the Council of 11 December 2007 Amending Council Directive 89/552/EEC on the Coordination of Certain Provisions Laid Down by Law, Regulation or Administrative Action in Member States concerning the Pursuit of Television Broadcasting Activities. Strasbourg: European Parliament.

Galán-Mañas, A., and A. Hurtado Albir. 2015. “Competence Assessment Procedures in Translator Training.” The Interpreter and Translator Trainer 9 (1): 63–82. doi:10.1080/1750399X.2015.1010358.

Gottlieb, H. 2015. “Different Viewers, Different Needs: Personal Subtitles for Danish TV?” In The Reception of Subtitles for the Deaf and Hard of Hearing in Europe, edited by P. Romero-Fresco, 17–44. Bern: Peter Lang.

Gummerus, E., and C. Paro. 2001. “Translation Quality. An Organizational Overview.” In (Multi) Media Translation: Concepts, Practices, and Research, edited by Y. Gambier and H. Gottlieb. Philadelphia: John Benjamins.

Huertas Barros, E., and J. Vine. 2015. “Assessing the Situation: Closing the Gap in Translation Assessment.” The Linguist 54 (4): 22–24.

Huertas Barros, E., and J. Vine. 2017. “Current Trends on MA Translation Courses in the UK: Changing Assessment Practices on Core Translation Modules.” The Interpreter and Translator Trainer 12 (1). doi:10.1080/1750399X.2017.1400365.

ISO (International Organization for Standardization). 2015. ISO 17100:2015. Translation Services - Requirements for Translation Services. Geneva: International Organization for Standardization.

James, H. 2001. “Quality Control of Subtitles: Review or Preview?” In (Multi) Media Translation: Concepts, Practices and Research, edited by Y. Gambier and H. Gottlieb, 151–160. Philadelphia: John Benjamins.

Jiménez Hurtado, C., ed. 2007. Traducción y accesibilidad. Subtitulación para sordos y audiodescripción para ciegos: Nuevas modalidades de traducción audiovisual. Frankfurt am Main: Peter Lang.
Matamala, A., and P. Orero, eds. 2010. Listening to Subtitles. Subtitles for the Deaf and Hard of Hearing. Bern: Peter Lang.

Media Access Australia. 2014. “Caption Quality: International Approaches to Standards and Measurement.” http://www.mediaaccess.org.au/sites/default/files/files/MAA_CaptionQuality-Whitepaper.pdf

Mertler, C. A. 2001. “Designing Scoring Rubrics for Your Classroom.” Practical Assessment, Research & Evaluation 7 (25), n.d.

Mossop, B. 2014. Revising and Editing for Translators. Oxon and New York: Routledge.

Neves, J. 2005. “Audiovisual Translation: Subtitling for the Deaf and Hard-of-Hearing.” PhD diss., Roehampton University.

Neves, J. 2008. “Training in Subtitling for the d/Deaf and the Hard of Hearing.” In The Didactics of Audiovisual Translation, edited by J. Díaz Cintas, 171–189. Philadelphia and Amsterdam: John Benjamins.

Ofcom (Office of Communications). 2015. “Code on Television Access Services.” https://www.ofcom.org.uk/__data/assets/pdf_file/0016/40273/tv-access-services-2015.pdf

Pavani, S. 2016. “La evaluación sumativa en la didáctica de la traducción. Un estudio empírico en la enseñanza de la traducción español-italiano.” PhD diss., Universitat Autònoma de Barcelona.

Pereira, A., and I. Lorenzo. 2007. “Teaching Proposals for the Unit ‘Subtitling for the Deaf and Hard of Hearing’ within the Subject Audiovisual Translation (English>Spanish).” Translation Watch Quarterly 3 (2): 26–36.

Rasmussen, K. W., and A. Schjoldager. 2011. “Revising Translations. A Survey of Revision Policies in Danish Translation Companies.” JoSTrans, the Journal of Specialised Translation 15: 87–120.

Romero Fresco, P., and J. Martínez. 2015. “Accuracy Rate in Live Subtitling: The NER Model.” In Audiovisual Translation in a Global Context. Mapping an Ever-Changing Landscape, edited by R. Baños Piñero and J. Díaz Cintas, 28–50. New York: Palgrave MacMillan.

Romero-Fresco, P. 2011. Subtitling Through Speech Recognition: Respeaking. Manchester: St Jerome.

Romero-Fresco, P., ed. 2015. The Reception of Subtitles for the Deaf and Hard of Hearing in Europe. Berlin: Peter Lang.

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.

Sánchez-Gijón, P. 2014. “Traducció i qualitat: Catàleg d’eines i recursos.” Tradumàtica 12: 493–507. doi:10.5565/rev/tradumatica.79.

Tamayo, A. 2016. “Subtitulación para personas sordas: Metodología de aprendizaje-servicio en el aula de traducción audiovisual.” íkala, Revista De Lenguaje Y Cultura 21 (3): 327–342. doi:10.17533/udea.ikala.v21n03a06.

Tamayo Masero, A. 2015. “Estudio descriptivo y experimental de la subtitulación en TV para niños sordos. Una propuesta alternativa.” PhD diss., Universitat Jaume I.

Waddington, C. 2000. Estudio comparativo de diferentes métodos de evaluación de traducción general (inglés-español). Madrid: Universidad Pontificia Comillas de Madrid.

Zárate, S. 2014. “Subtitling for Deaf Children Granting Accessibility to Audiovisual Programmes in an Educational Way.” PhD diss., University College London.