EDICIONES
COMPLUTENSE

Complutense Journal of English Studies
ISSN-e: 2386-3935

ESTUDIOS

# Automatic closed captions and subtitles in academic video presentations: possibilities and shortcomings[1]

**Mª Azahara Veroz-González**
Department of Language Sciences, University of Córdoba

**Mª Pilar Castillo Bernal**
Department of Social Sciences, Philosophy, Geography and Translation and Interpreting, University of Córdoba

https://dx.doi.org/10.5209/cjes.94649

Recibido: 19/02/2024 • Aceptado: 17/06/2024

ENG **Abstract:** In light of the increasing number of academic events being recorded or held online since the onset of the COVID-19 pandemic, the present work combines automation processes in audiovisual translation and academic texts–more specifically, video presentations. The research questions are whether the automatic generation of captions is functional to ensure accessibility in academic events and how much post-editing effort would such content require in case a machine translation of the subtitles is to be applied. The research method comprises several phases. First, in a corpus of video presentations of specialised content in English, captions were generated automatically using YouTube Studio to ascertain the general quality and the type of errors generated in the automatically generated closed captions[2] according to Multidimensional Quality Metrics (MQM) framework. These auto-generated captions were corrected and annotated by considering the following parameters: a) pre-editing time, b) type of error according to MQM framework, and c) severity of the error. Second, the auto-generated captions and corrected were machine translated into Spanish. Furthermore, errors detected in the machine translation of the subtitles (English-Spanish) were post-edited and errors were analysed following the MQM. Reception by a potential audience was also studied, as evaluated by academics from the same field of expertise. The main conclusion is that most errors in machine-translated subtitles stem from incorrect caption segmentation and lack of context awareness, making it essential to correct the closed captions before translation. This thesis is supported by the reception study in which the level of comprehension was higher when the transcription was pre-edited, as most of the problems arise from the closed captions rather than from the translation itself.

**Keywords:** Auto-generated captions; subtitles; machine translation; Quality Assessment; academic presentations

## ESP *Closed captions* automáticos y subtítulos en comunicaciones académicas en vídeo: posibilidades y limitaciones

ES **Resumen:** Dado el creciente número de eventos académicos celebrados en línea o grabados desde el inicio de la pandemia por COVID-19, el presente trabajo combina los procesos automáticos de traducción audiovisual y los textos académicos, en concreto las comunicaciones en vídeo. Las preguntas de investigación son si la generación automática de la transcripción es funcional para garantizar la accesibilidad de eventos académicos y qué esfuerzo de posedición requerirían estos contenidos si se desea una traducción de los subtítulos. La metodología investigadora consta de varias fases. Primero, en un corpus de presentaciones en video de contenido especializado en inglés se generó la transcripción automáticamente empleando YouTube Studio para determinar la calidad general y el tipo de errores. Dichas transcripciones se corrigieron y anotaron con base en los siguientes parámetros: a) tiempo de posedición, b) tipo de error según el Multidimensional Quality Metrics (MQM) y c) gravedad del error. En un segundo paso, se generó una traducción automática al español de la transcripción generada por YouTube Studio y corregida. Asimismo, los errores detectados en la traducción automática de los subtítulos (inglés-español) se analizaron según el MQM y se estudió la recepción por parte de espectadores potenciales:

1 The research presented in this study has been (partially) carried out in the framework of research project "Training app for post-editing neural machine translation using gamification in professional settings" (GAMETRAPP) (reference number TED2021-129789B-I00).

2 Automatically generated closed captions are commonly referred to as 'automatic captions' or 'auto-generated captions.' These captions are produced using speech recognition technology and are often available on platforms like YouTube. (Youtube Help, s.d.)

académicos del mismo campo de especialidad. La principal conclusión es que la mayoría de los errores de los subtítulos traducidos automáticamente se deben a una segmentación incorrecta la transcripción y la falta de contexto, por lo que es esencial corregir la transcripción antes de traducirlos. Esta tesis se ve respaldada por el estudio de recepción, en el que el nivel de comprensión fue mayor cuando la transcripción se editó previamente, ya que la mayoría de los problemas surgen esta y no de la propia traducción.
**Palabras clave:** transcripción automática, subtítulos, traducción automática, evaluación de la calidad, comunicaciones académicas.

**Contents:** 

**How to cite:** Veroz-González, M.ª A.; Mª Pilar Castillo Bernal (2024). Automatic closed captions and subtitles in academic video presentations: possibilities and shortcomings, en *Complutense Journal of English Studies* 32, e94649. https://dx.doi.org/10.5209/cjes.94649

## 1. Introduction

Machine translation (MT) is a frequently used tool by researchers in daily work. Since the first MT systems were implemented in the 1940s (Bennet and Slocum 1985), MT has undergone several phases and has improved dramatically in its implementation by users, researchers, and professional translators.

In recent years, the inclusion of MT in the workflow of companies and researchers has resulted in increased productivity and decreased production costs (Moorkens 2017). The reason is a change of paradigm whereby, on the one hand, the need for immediate communicative exchange has become global, leading to the use of MT (Hernández Mercedes 2002); on the other hand, the improved quality of machine translations as compared to a few years ago has caused a rise in the number of MT users. In this context, some researchers reckon that MT quality is starting to equal human translation (Wu, *et al.* 2016).

Along these lines, Briva-Iglesias (2021) claims that translation companies are witnessing an increase in MT. This is also a finding in the 2020 report by the European Union of Associations of Translation Companies (EUATC), which claims that in 2017, 45% of companies offering language services used MT, and in just three years, this figure had gone up to almost 70%. This reality could be extrapolated to knowledge transfer, a context in which an increasing number of researchers are using MT systems, either for understanding the texts they need in their work or to translate short texts such as abstracts or conference papers (Gaspari et al. 2015).

Bearing in mind the liability an error in interpretation of research texts may carry, it seems necessary to assess the quality of machine translations to put in place improvement measures that facilitate communication accurately. For this reason, the general research question in this article is whether auto-generated captions (YouTube Help s.d.) and MT are reliable for translating oral communications of specialised content from English into Spanish from the point of view of technical, linguistic and textual correctness and for promoting understanding among the scientific community.

Therefore, the main objectives of this study are as follows:

1. To demonstrate the usefulness of auto-generated captions and MT in oral scientific communications in the EN>ES language combination.
2. To investigate the reception of machine-translated oral communications.

The specific research questions are whether the auto-generated captions are functional to ensure accessibility in academic events and how much post-editing effort would such content require in case a translation of the subtitles is to be applied. Several video recordings of scientific communications were selected on YouTube, the auto-generated captions were pre-edited[3] and their errors were classified using Multidimensional Quality Metrics (MQM), following customised error classification systems. Furthermore, both the non-pre-edited and the pre-edited auto-generated captions were machine-translated and their errors were analysed again according to MQM. Finally, a survey was conducted among potential viewers of the communications using auto-generated and machine-translated captions. we worked with two subcorpora: subcorpus 1 automatically generated closed captions by YouTube, and subcorpus 2 automatically generated closed captions by YouTube that were subsequently pre-edited. Both corpora were subjected to automatic translation, and the errors in the translated texts were categorized following the MQM framework

### 1.1. The concept of machine translation (MT)

Since the implementation of machine translation in the 1940s (Bennet and Slocum 1985), an increasing number of authors has studied the concept of MT and its implications (Hutchins and Somers,1992; Berner 2003; Llisterri 2009; Sánchez Ramos and Rico Pérez 2020, to name but a few).

[3] Pre-edition: Pre-editing is the process of correcting CCs before they are translated.

Berner (2003) defines MT as the use of a computer program to translate a text from one natural language into another. This idea revolves around computer programs and the language translation process. Subsequently, Sánchez Ramos and Rico Pérez (2020, 2) have taken up this concept, emphasising that MT entails the automatic translation of text by a computer program, without requiring any human intervention (the translator). In other words, MT systems can translate from one language to another without human intervention.

This combination of computer programs and translation without humans has always been at the core of MT; however, it should be mentioned that other authors and even companies include other aspects that are as necessary as a subsequent review, quality, and visibility from an economic point of view or understanding the target text. This is the case for companies such as Systran (s.d.) and authors such as Hutchins and Somers (1992), who refer to the ability of MT systems to produce raw translations in a well-defined field. However, the same authors highlight that these translations should be reviewed at an economically viable pace to achieve good quality and be understood by specialists. It is interesting that they introduced the concept of economic viability because if the MT had poor quality, it would not be viable from an economic point of view, since the reviewer would need to retranslate the text to make it understandable by a specialist.

Therefore, regardless of the precision introduced by each author, MT is no more than translating a text or speech from one language to another using a computer program without the intervention of a human being, whereby factors as important as the quality and reliability of the target text should be considered, together with human intervention in the reviewing process and the cost of this intervention.

### 1.1.1. MT systems and automatic subtitles

Before explaining the materials and methods used in the present work, it is relevant to mention the MT system used, as this determines the quality of the target text.

Parra Escartín (2018) claims that nowadays three MT types coexist: a) rule-based MT, using linguistic information (RBMT); b) statistical machine translation or SMT, and c) neural machine translation or NMT. Neural MT is used by DeepL and Google Translate[4] and its purpose is to mimic the way human brain cells work. According to Parra Escartín (2018), language components connect with underlying information to establish associations and create translations, similar to how neurones in the brain process and integrate information. Therefore, in this MT system, the computer automatically learns how to translate both linguistic and intralinguistic information from a large corpus of parallel texts.

The MT technology investigated in this study is NMT, applied to the field of audiovisual translation (AVT), as described by Karakanta et al. (2020a, 1):

> (…) recent developments in neural machine translation (NMT) and speech translation (ST) are paving the way for viable and usable (semi-)automatic solutions for subtitling. Compared to solutions providing MT for subtitling, automatic subtitling tools do not simply translate human-generated source language subtitles but incorporate automatic transcription of the speech, MT, automatic synchronisation (spotting), and segmentation of the translated speech into subtitles. Altogether, these technologies come with the promise of reducing the human effort in the subtitling process, but, to date, automatic subtitling has still to be put to test by the actual users.

NMT systems have improved the quality of machine translation dramatically; however, as described by Karakanta et al. above, additional technical steps play a role in automatically generated subtitles before they are machine-translated: speech recognition and transcription, spotting (synchronisation with the speech) and segmentation (separation of the text in subtitles and lines). To simplify the process and reduce errors, Papi et al. (2023) propose a direct speech translation model for automatic subtitling that generates subtitles in the target language along with their timestamps. Although the results are promising, the effort and scope of developing such a system are not applicable to the type of academic subtitling (i.e., using automatic subtitles for sporadic video presentations) which is the subject of the present work.

The use of automatic subtitles in academia is not new: Valor Miró et al. (2015) describe a lecture video repository with automatically generated transcriptions and translations in Spanish, Catalan and English at the Universitat Politècnica de València (Spain) and evaluate the efficiency of the manual review process from automatic subtitles in comparison with the conventional generation of video subtitles from scratch. Despite errors, they report significant savings in time of up to almost 75 % when reviewing subtitles. In turn, Che et al. (2017) propose an integrated framework of automatic bilingual subtitle generation for lecture videos, especially for MOOCs. The evaluation of the auto-generated subtitles, the manually produced subtitles from scratch, and the auto-generated subtitles with manual modification shows that the auto-generated subtitles in the original language (English) are fairly accurate already but the effectiveness of machine translated subtitles (English to Chinese) is limited. The total working time in preparing bilingual subtitles can be shortened by approximately 1/3, with no decline in quality.

In the field of television, de Higes (2023) analyses automatic bilingual subtitles used in València TV news broadcast. From the perspective of user comprehension, the automatic bilingual subtitles are not yet acceptable since they lack coherence, subtitle delay is high and subtitle speed exceeds standard recommendations, dual speakers in a subtitle are not indicated, and text segmentation is poor. Along the same topic, Hagström & Pedersen (2022) conducted a diachronic study of subtitles before and after machine translation by comparing a corpus of Swedish subtitles of Anglophone TV programmes produced after machine translation was introduced

[4] It should be noted that Google Translate uses both SMT and NMT.

to a corpus of subtitles from before that period. The post-edited subtitles produced in the 2020s were found to be faster, more oral, less cohesive, less complete and with less meticulous punctation and line-breaks than those produced in the 2010s. They were also of significantly lower quality in all areas investigated.

Other efforts to reduce errors in automatic subtitles focus on improving segmentation by training customised models (Wan et al. 2020) or customising neural machine translation for subtitling (Matusov at al. 2019). However, these extensive studies require large data set collections, technology and time usually not available to most academics who may sporadically use automatic and machine-translated subtitles for communication purposes. Therefore, the investigation focuses on MT quality applied to automatic subtitling, as well as potential users' opinions on said quality, in order to ascertain which types of errors are most frequent and where post-editing efforts should focus to achieve acceptable quality.

### 1.1.2. MT quality assessment

Quality assessment of machine translation has become an important focus of research in recent years, ever since the introduction of neural MT and the significant improvement of raw MT output. The MT quality can be assessed manually or automatically (Moorkens et al. 2018, 25). Automatic machine translation evaluation metrics measure either the number of corrections needed by the raw MT (such as the metrics TER and PER) or the order of words and words groups (metrics BLEU or METEOR), using a reference translation as a point of comparison. This is rather limited, since there is no single possible translation for each text (Briva-Iglesias 2021, 579). For this reason, Läubli et al. (2020, 1-2) indicate that human evaluation of MT is still the standard despite its high cost in time and resources, and the fact that it can be compromised if not applied rigorously.

For the purpose of the present study, the study by Karakanta et al. (2022a and 2022b) is of special relevance. The authors conducted a survey to collect subtitlers' impressions and feedback on the use of closed captions in workflows. Their results showed that the main issues of closed captions stem from failures in speech recognition and preprocessing, which result in error propagation, translations out of context, inaccuracies in auto-spotting, and suboptimal segmentation. However, the use of closed captions saves time and effort, and is valued by respondents from neutral to positive. In the present work, pre-editing and categorization of errors of closed captions and machine translation were conducted using the MQM categorisation. To complete the evaluation of the translated captions, a survey was conducted among potential viewers of the communications involved, as shall be explained in the following.

## 2. Materials and methods

In this section, the following aspects are explained.

a) Corpus design.
b) Analysis method: MQM.
c) Survey to measure reception.

### 2.1. Corpus design

At their most recent conference in Oslo in 2022, the European Society for Translation Studies (EST 2022) advised virtual presenters to record their papers and use automatic captioning to ensure accessibility. Recently, several virtual platforms have offered users the possibility of autogenerated closed captions generated during meetings or conferences (e.g. Cisco Webex), or stored videos (for example, YouTube). The term 'closed captions' indicates that the captions on the screen can be turned off and on and become visible on the screen only when activated by the viewer (Díaz Cintas & Remael 2021, 26-27), as opposed to open captions, which cannot be turned off by the user. It should be noted that by 'captions' platforms mean an automatic transcription of the speech which appears on screen in the form of a subtitle, but without complying with subtitling standards in terms of text length, visual conventions to avoid confusion, minimum time onscreen to ensure readability, etc.

For these reasons, we aimed to explore whether this practice of using auto-generated captions can ensure accessibility in academic presentations, especially if the captions are machine-translated and automatically generated. Given these developments and the increasing number of academic events being recorded or held online since the onset of the COVID-19 pandemic, the present work combines automation processes in audiovisual translation (speech recognition software, automatic and machine-translated subtitles, and captions) and academic texts–more specifically, video presentations.

In this study, a series of videos was selected from a virtual event organised by the European Commission in 2022, titled 'AVT workflows and the role of automation. From Translation To Accessibility.' These videos were available on YouTube, and specifically, Session 2 and the round table were chosen for analysis, comprising a total duration of 2 hours and 23 minutes, equivalent to approximately 29,670 words.

Closed captions for these videos were automatically generated by the platform hosting the event. To improve the quality of the captions, they were first pre-edited using the SDL Trados Studio 2022. The pre-editing process involved aligning two versions of the auto-generated captions within SDL Trados Studio. This alignment allowed for a direct comparison between the initial and revised versions, making it easier to identify and correct errors. We then used the 'Revision' and 'Quality Control (QC)' functions in SDL Trados Studio to meticulously correct and annotate the captions. The 'Revision' function enabled us to systematically review and edit the text, while the Q' function helped identify issues such as spelling errors, grammar mistakes, and inconsistencies. This thorough pre-editing ensured that the captions were accurate and met quality standards. Subsequently, the errors present in the pre-edited captions were categorized according to the MQM framework, which provides a standardised method for evaluating translation quality.

Once the pre-editing process was completed, the auto-generated captions were machine translated into Spanish. This study aimed to compare the quality of machine-translated text with and without pre-editing of the source text. The researchers sought to establish the difference in machine translation quality between the two cases by examining the number of errors in both versions of the translated text.

In summary, we worked with two subcorpora: subcorpus 1 automatically generated closed captions by YouTube, and subcorpus 2 automatically generated closed captions by YouTube that were subsequently pre-edited. Both corpora were subjected to automatic translation, and the errors in the translated texts were categorized following the MQM framework.

## 2.2. Analysis method: MQM

In this section, the analysis parameters used to determine the quality of the machine-generated and machine-translated captions are explained. The analysis method chosen was MQM, a framework for analytic Translation Quality Evaluation (TQE) that can be applied to both human translation and machine translation, developed by Lommel (2018) and colleagues at the German Research Centre for Artificial Intelligence (DFKI) in Berlin. However, MQM can also be applied to the automatic transcription process. This makes sense because automatic transcriptions, like translations, can contain various types of errors that affect their overall quality and usability. Applying MQM to automatic transcriptions allows for a comprehensive evaluation of the transcription quality by identifying and categorizing errors, which can then inform improvements in both transcription and subsequent translation processes. MQM can be used to identify quality issues in translation products; classify them against a shared, open, and standardised error typology; and generate quality measures that can be used to gauge how well the translation product meets the quality requirements. MQM divides errors into seven main categories: *terminology, accuracy, linguistic conventions, style, locale conventions, audience appropriateness,* and *design and markup.* They are further divided into subcategories, and the classification can be customised (see Group Multilingual Technologies, 2024). The classification used in this study were as follows:

- Accuracy: Errors occurring when the target text does not accurately correspond to the propositional content of the source text introduced by distorting, omitting, or adding to the message.
    - Mistranslation: Target content that does not accurately represent the source content.
    - Misrecognition: Errors in speech recognition.
    - Overtranslation: Target text that is inappropriately more specific than the source text.
    - Undertranslation: The target text is inappropriately less specific than the source text.
    - Addition: Target content that includes content not present in the source.
    - Omission: Errors where content is missing from the translation present in the source.
    - Do not translate: Errors occurring when a text segment or even a whole section of a text marked in the specifications as “Do not translate!” were translated into the target text.
    - Untranslated: Errors occur when a text segment intended for translation is left untranslated in the target content.
- Audience appropriateness: Culture-specific reference. Errors arising from the use of content in the translation product are invalid or inappropriate for the target locale or audience.
- Design and markup: Errors related to the physical design or presentation of a translation product, including character, paragraph, and UI element formatting and markup; integration of text with graphical elements; and overall page or window layout.
    - Character formatting: Inappropriate application of any glyph variation to a character or string of characters, such as font, font style, font colour, or font size.
    - Layout: Inappropriate presentation format of paragraphs, headings, graphical elements, and user interface elements and their arrangement on a form, page, website, or application screen.
    - Markup tag: Incorrect markup tag or tag component.
    - Truncation/text expansion: Target text that is longer or shorter than allowed or where there is a significant and inappropriate discrepancy between the source and target text lengths.
    - Missing text: Existing text missing in the final laid-out version.
    - Link/cross reference: Incorrect or invalid (no longer active) links or URL.
- Linguistic conventions: Errors related to the linguistic well-formedness of the text, including problems with grammaticality and mechanical correctness.
    - Grammar: An error that occurs when a text string (sentence, phrase, or other) in the translation violates the grammatical rules of the target language.
    - Punctuation: Punctuation incorrect for locale or style.
    - Spelling: Errors occur when letters in a word in an alphabetic language are not arranged in a normally specified order.
    - Unintelligible: Text garbled or incomprehensible, perhaps due to conversion or other processing errors.
    - Character encoding: Error occurs when characters garble because of incorrect application of encoding.

- Locale conventions: Errors occur when a translation product violates locale-specific content or formatting requirements for data elements.

  - Number format: Inappropriate number format for its locale.
  - Currency format: incorrect currency format for locale.
  - Measurement format: incorrect measurement format for locale.
  - Time format: incorrect time format for locale.
  - Date format: Incorrect date format for locale.
  - Address format: Incorrect address format for its locale
  - Telephone format: Incorrect telephone format for its locale
  - Shortcut key: Shortcuts in translated software products that do not conform to local expectations or make no sense for the locale.

- Terminology: Errors arising when a term does not conform to normative domain or organizational terminology standards, or when a term in the target text is not the correct normative equivalent of the corresponding term in the source text.

  - Inconsistent with terminology resource: Use of a term that differs from term usage required by a specified term base or other resources.
  - Inconsistent use of terminology: Use of multiple terms for the same concept in cases where consistency is desirable.

  - Wrong term: Use of a term that it is not the term a domain expert would use or because it gives rise to a conceptual mismatch.
- Style: Errors occurring in a text that can be grammatical but inappropriate because they deviate from organizational style guides or exhibit an inappropriate language style.

  - Organizational style: Errors occurring when the text violates company/organization-specific style guidelines.
  - Third-party style: Errors occurring when the text violates a third-party style guide.
  - Inconsistent with external reference: Errors occurring when text fails to conform to an external resource.
  - Register: Errors occurring when a text uses a level of formality higher or lower than that required by the specifications or common language conventions.
  - Awkward style: Style involving excessive wordiness or overly embedded clauses, often due to inappropriate retention of the source text style in the target text.
  - Unidiomatic style: A style that is grammatical but unnatural, often due to interference from the source language.
  - Inconsistent style: A style that varies inconsistently throughout the text, often due to multiple translators contributing to the target text.

See MQM (2024) for further information and examples.

The severity of the error was also classified, prioritizing factors that directly impact the correct understanding of the text, while considering linguistic errors on a secondary level. Consequently, the errors were classified as minor, major, or critical. Specifically, a minor error impacts only language correctness but does not affect text comprehension. In contrast, a major error affects both language correctness and text comprehension, although the text can still be understood with some difficulty. Furthermore, a critical error significantly impacts both language correctness and text comprehension, leading to gross mistranslations or misinterpretations.

The computer programs we used include YouTube for auto-generated captions in English and Downsub for downloading the captions as a. srt file, DeepL Translator to translate the auto-generated captions and SDL Trados Studio 2022: to pre-edit the original captions while measuring the time needed, to classify error type and gravity according to the customized MQM classification, and to classify errors in the MT captions from English into Spanish according to MQM. Thus, not only were the errors quantified and classified, but the post-editing time was also recorded. This data is essential for evaluating the efficiency of MT and determining whether it effectively provides accessibility to content.

## 2.3. Survey

To fulfil objective number 2, which refers to the reception of automatic subtitling in oral communications, a survey was designed and validated by 12 experts in translation and interpreting. Potential recipients had to assess the reception of such subtitles from the point of view of comprehension, linguistic expression, and technical aspects.

The survey consisted of two parts. The first part was devoted to the collection of personal data of a statistical nature: age group, gender, native language, and profession. The second part was devoted to the evaluation of the automatic translation of subtitles.

The second part, which evaluates the reception of subtitling, is divided into three parts. First, the evaluation of the quality of the subtitles of a 2 minutes and 30 seconds fragment of the video 'AVT workflows and the role of automation - Round table' after extracting the closed captions generated by Youtube and machine translating them into Spanish with DeepL. Second, the evaluation of the quality of the subtitles of the same 2 minutes and 30 seconds video 'AVT workflows and the role of automation - Round table' after extracting the closed captions generated by YouTube, pre-editing them and machine translating them into Spanish with DeepL. And, finally, the comparison of the reception of both auto-generated captions.

In the first two parts, the following questions were asked:

- Evaluate from 1 to 5 (5 being very good and 1 being very bad) whether the automatic subtitles are understandable (assuming the receiver did not understand English audio).
- Rate the degree of linguistic correctness of the automatic subtitles in Spanish from 1 to 5 (5 being very good and 1 being very bad).
- Evaluate from 1 to 5 (5 being very good and 1 being very bad) whether the subtitles are technically correct (time of appearance on the screen, segmentation, etc.).

And for the third part, the questions were the following:

- Which of the videos did you understand best?
- Which of the videos did you find easier to follow from the point of view of segmentation, subtitle exposure time, etc.?
- Do you think that this automatic subtitling system can be used to understand videos of academic communications (at conferences, etc.)?

## 3. Results and discussion

In this section, the following results shall be explained: pre-editing time, specific and general errors categorised after pre-editing the original language and the machine translation into Spanish, and survey results.

In general terms, a specialised translator can translate up to 3500 words daily (Gámez & Cuñado 2019), that is, an average of 437 words per hour on an 8-hour workday. The corpus in this article reaches 29,670 words, for which a specialised translator would need approximately 2 days, 19 hours and 12 minutes. However, it should be noted that the work of audiovisual translators may greatly vary from this estimate because their tasks are often of a creative nature, have to deal with spatial and other constraints, and involve technical aspects, such as spotting and segmentation. To date, no reliable data can be quoted on the number of words that an AVT translator can translate or post-edit.

To ascertain whether NMT is applicable to our corpus, not only should errors be analysed, but the economic profitability of post-editing should also be taken into account (Hutchins & Somers 1992). In this context, time is an essential factor, since post-editing would not be profitable if it took longer than from-scratch translation. For this reason, in this study, the pre-editing for the corpus in the source language and post-editing for the machine translation was recorded:

1) Post-editing DeepL's MT of the auto-generated captions took one day and six hours to complete.
2) Downloading and pre-editing the closed caption generated by YouTube, machine translation by DeepL, and post-editing took 5 h and 16 min for pre-editing the source text (whereby speech recognition and segmentation errors were corrected), and 7 h and 15 min for post-editing the MT.

Therefore, MT saves time in both translation and post-editing compared to the time required for translating from scratch. However, it is more cost-effective to pre-edit the source text before machine-translating and then post-editing the translation.

### 3.1. Error types and severity

As explained above, errors were categorised according to their type and severity. In this section, errors in the original auto-generated captions downloaded from YouTube (subcorpus 1) and in the two machine translations generated (one from subcorpus 1 and another from the pre-edited captions or subcorpus 2) are outlined.

#### 3.1.1. Error types and severity in the auto-generated captions

Subcorpus 1 presents a total of 6314 errors: 66% are segmentation errors, 17% are punctuation errors, and 16% are spelling errors. The remaining types are not statistically significant, as can be seen in Figure 1.

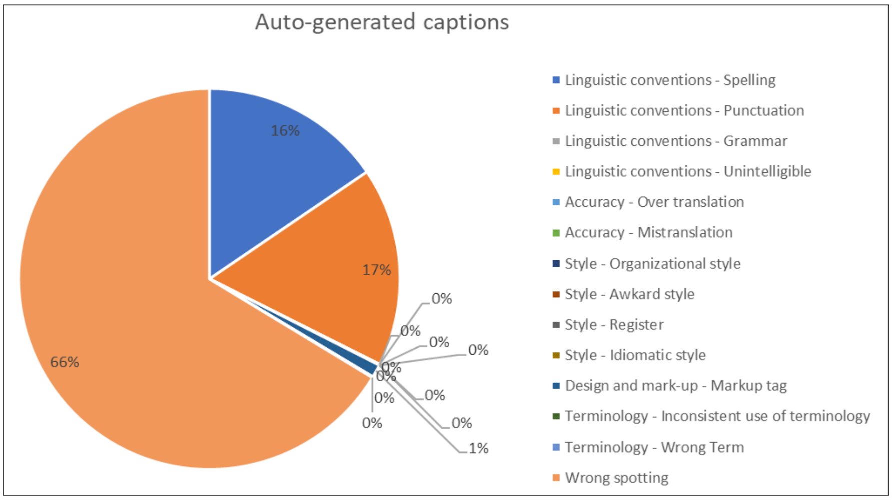


Figure 1. Classification errors by typology in source text or subcorpus 1.

After conducting a qualitative analysis of the source text, ¿it can be concluded that? YouTube transcribed the text directly from speech recognition software: no punctuation marks and no capital letters. It also does not consider the complete units of meaning when segmenting the subtitles, resulting in considerable issues for machine translation.

Table 1 shows how units of meaning and punctuation are not respected during segmentation.

Table 1. Example 1

| AUTO-GENERATED CAPTIONS<br>(Subcorpus 1) | PRE-EDITED AUTO-GENERATED CAPTIONS<br>(Subcorpus 2) |
|---|---|
| 1<br>00:00:00,080 --> 00:00:01,599<br>for the<br><br>2<br>00:00:01,599 --> 00:00:04,400<br>final session today we have a round<br><br>3<br>00:00:04,400 --> 00:00:06,640<br>table with lots of very interesting | 1<br>00:00:00,080 --> ~~00:00:01,599~~<br>00:00:04,400<br>~~f~~For the final session today,<br><br>2<br>00:00:01,599 --> 00:00:04,400<br>~~final session today~~ we have a round table<br><br>3<br>00:00:04,400 --> 00:00:06,640<br>~~table~~ with lots of very interesting people. |

Similarly, a large number of unnecessary repetitions and interjections were found because the software transcribes everything from the original audio.

Interestingly, only 3% of misrecognitions were detected; these were speech recognition errors, and most of them concerned proper names, acronyms, and names of associations, as can be seen in Table 2.

Table 2. Example 2

| AUTO-GENERATED CAPTIONS<br>(Subcorpus 1) | PRE-EDITED AUTO-GENERATED CAPTIONS<br>(Subcorpus 2) |
|---|---|
| 00:01:29,840 --> 00:01:32,159<br>british of texas association satel and | 00:01:29,840 --> 00:01:32,159<br>~~b~~British of ~~texas~~ Subtitlers' ~~a~~Association ~~satel~~<br>SUTLE and |
| 00:02:14,080 --> 00:02:17,599<br>um martin reverse from limecraft and | 00:02:14,080 --> 00:02:17,599<br>~~um martin reverse~~ Maarten Verwaest from<br>~~l~~Limecraft and |

Regarding the severity of errors, it should be mentioned that most of them did not have an impact on the general comprehension of the text, so they were ranked as minor errors (see Figure 2).

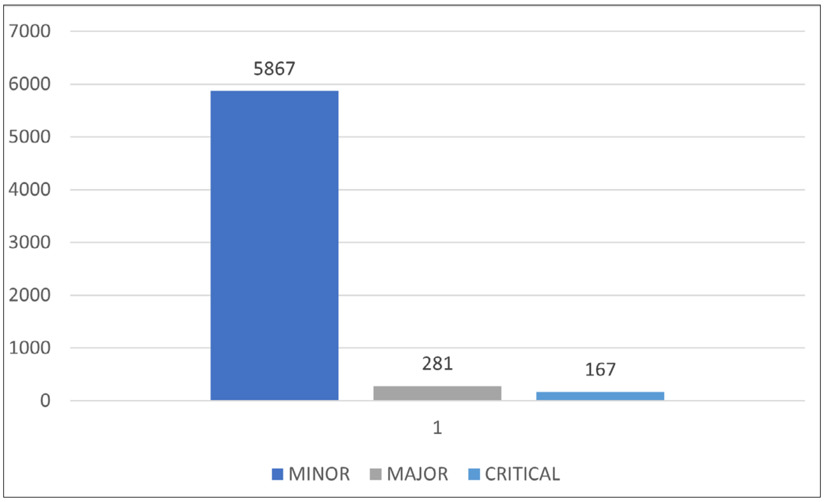


Figure 2. Classification of errors by severity in the subcorpus 1.

### 3.1.2. Error types and severity in the machine translation generated from the non-pre-edited source text

The number of errors found in the machine translation of the non-pre-edited auto-generated captions (subcorpus 1) did not show a quantitatively significant difference from the errors found in subcorpus 1, with a total of 6614 errors (Figure 3).

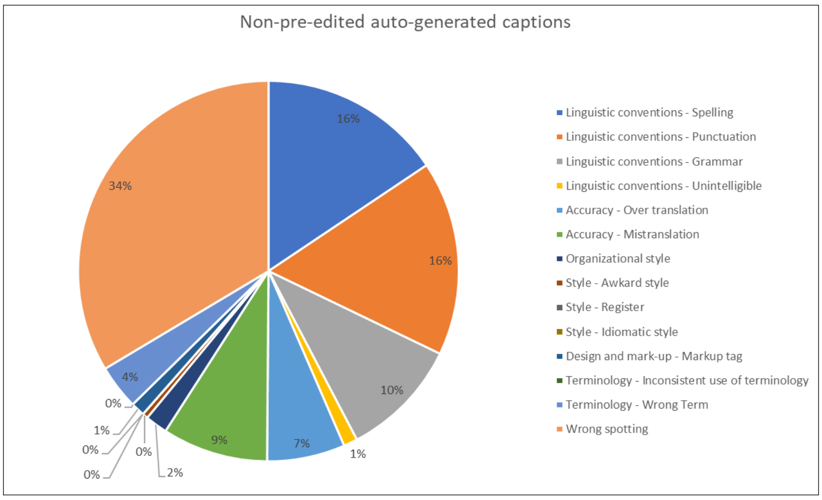


Figure 3. Classification errors by typology in the non-pre-edited subcorpus 1.

As shown in Figure 3, most errors were caused by spotting and linguistic conventions, more specifically punctuation and spelling, amounting to 60% of the errors, although most of them were ranked as minor (same as in the analysis of the source text), since they do not have an impact on comprehension. It should be noted that other types of errors occur that were not present in subcorpus 1: grammatical errors, over-translation, wrong terms, or inconsistencies with external references.

After performing a qualitative analysis of the errors, we found that errors concerning grammar (9%) were caused by the following three aspects:

a) Wrong segmentation of the source text, for example, separation of the subject and verb:

Table 3. Example 3

| AUTO-GENERATED CAPTIONS (Subcorpus 1) | MT results from subcorpus 1 |
|---|---|
| 131<br>00:05:02,479 --> 00:05:05,520<br>um my name is eva and i<br><br>132<br>00:05:05,520 --> 00:05:08,479<br>started a year ago at sue um and i'm | 131<br>00:05:02,479 --> 00:05:05,520<br>um mi nombre es eva y yo<br><br>132<br>00:05:05,520 --> 00:05:08,479<br>comenzó hace un año en sue um y estoy |

As can be observed, subtitle 131 finishes at “yo”, and 132 begins with “comenzó”, instead of “comencé” (I began). This also occurs when other units of meaning are severed.

b) Issues of recognition of male and female nouns with a tendency towards males. For example, Rosario[5], a woman, becomes a male subtitler (“subtitulador”); however, Marteen is considered a female translator or “traductora”, instead of “traductor”.

In the case of over-translation (6%), most errors concerned interjections and unnecessary repetitions introduced by the speaker, mostly when hesitating. Table 3 presents a clear example of the interjections in Subtitles 131 and 132.

The category inconsistencies with external references (7%) and wrong terms (2%) are due to misrecognitions of the source text: either the proper name of a person, company, association, or an acronym. Table 3 presents an example of Zoo Digital translated as Sue (the same term appears in the source text). The speaker Ewa is called Eva in both the source and translation.

Finally, it is relevant to mention the mistranslation errors (8%), where the target content does not accurately represent the source content. As shown in Table 4, the text does not correspond to the source, because the units of meaning are severed in the segmentation. Similarly, interjections in the source and repetitions, as well as backtracking, result in errors in machine translation. However, once the source text is corrected, machine translation does not generate this type or severity of errors.

Table 4. Example 4

| AUTO-GENERATED CAPTIONS (subcorpus 1) | MT from subcorpus 1 |
|---|---|
| 120<br>00:04:35,919 --> 00:04:39,280<br>my focus is uh on | 120<br>00:04:35,919 --> 00:04:39,280<br>mi enfoque es uh en |
| 121<br>00:04:39,280 --> 00:04:40,880<br>the tools | 121<br>00:04:39,280 --> 00:04:40,880<br>las herramientas |
| 122<br>00:04:40,880 --> 00:04:42,479<br>also known as a toolkit which is the | 122<br>00:04:40,880 --> 00:04:42,479<br>también conocido como kit de herramientas que es el |
| 123<br>00:04:42,479 --> 00:04:46,000<br>time text the web-based context product | 123<br>00:04:42,479 --> 00:04:46,000<br>texto de tiempo el producto de contexto basado en la web |
| 124<br>00:04:46,000 --> 00:04:46,960<br>uh from | 124<br>00:04:46,000 --> 00:04:46,960<br>uh de |
| 125<br>00:04:46,960 --> 00:04:50,240<br>also the integration part uh in of of of | 125<br>00:04:46,960 --> 00:04:50,240<br>también la parte de integración uh en de de de |
| 126<br>00:04:50,240 --> 00:04:52,000<br>this product into | 126<br>00:04:50,240 --> 00:04:52,000<br>este producto en |
| 127<br>00:04:52,000 --> 00:04:55,040<br>management solutions | 127<br>00:04:52,000 --> 00:04:55,040<br>soluciones de gestión |

Most errors were categorised as minor because they did not have an impact on the general understanding of the text. However, a higher percentage of major and critical errors was detected compared with the source, mostly grammar errors, inconsistencies with external references, and wrong terms.

5 It should be noted that Rosario is a gender-neutral name.

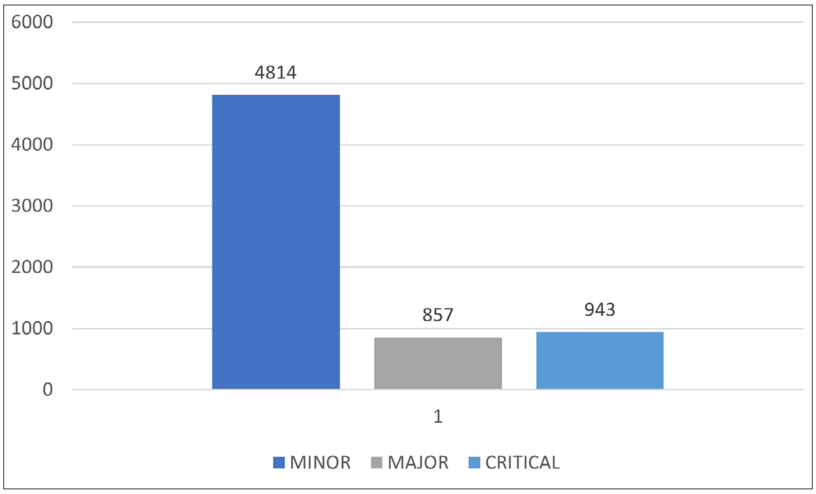


Figure 4. Classification of errors according to severity in the non-pre-edited subcorpus 1.

### 3.1.3. Error types and severity in the machine translation from the corrected source texts (subcorpus 2)

Finally, we outline the data obtained after analysing the machine translation from the post-edited source text. It is worth mentioning that 800 errors were detected in the same fragment, which means 5814 fewer errors than the translation from the non-corrected source text, a significant improvement of 86,24% in the machine translation from the corrected closed captions.

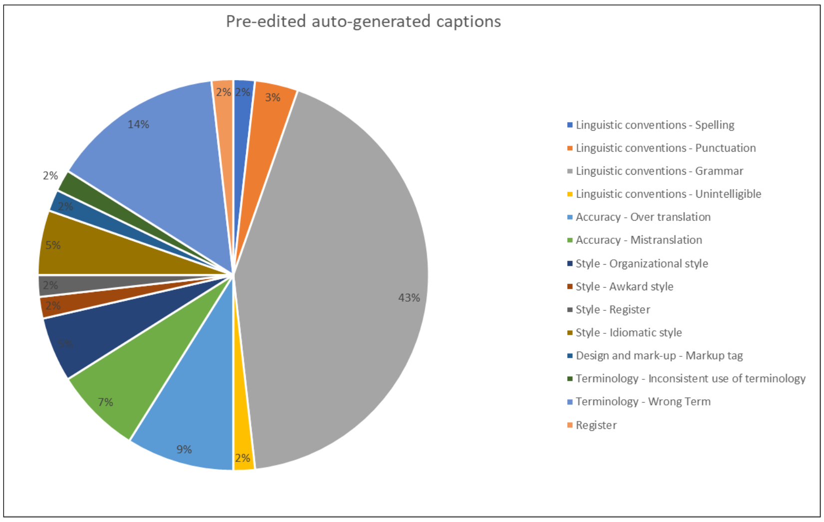


Figure 5. Classification errors by typology in the pre-edited auto-generated captions or subcorpus 2.

Of the 800 errors, 43% were grammar errors. In this case, although the percentage is higher, the number of errors is significantly lower than in the first machine translation: a total of 343 errors compared to 614 in the first translation, i.e. 44,13% fewer grammatical errors. It should also be noted that most errors were due to concordance, mostly because the term in the concordance appeared in the previous subtitle. This does not impede understanding and is, therefore, ranked as minor.

Regarding wrong terms, 14% of the errors belong to this category. However, the number was 114 compared to 229 errors detected in the first translation, indicating an improvement of over 50%. These are again errors which do not impede comprehension of the text:

Table 3. Example 3

| AUTO-GENERATED CLOSED CAPTIONS (subcorpus 1) | MT from subcorpus 1 | PRE-EDITED MT |
|---|---|---|
| 89<br>00:04:35,910 --> 00:04:42,470<br>My focus is on the tools,<br>also known as a toolkit, | 89<br>00:04:35,910 --> 00:04:42,470<br>Mi atención se centra en las herramientas,<br>también conocido como conjunto de herramientas | 89<br>00:04:35,910 --> 00:04:42,470<br>Mi atención se centra en las herramientas,<br>también conocidas como paquete de herramientas, |
| 98<br>00:05:12,080 --> 00:05:18,000<br>It has, obviously, grown out of proportion sales tasks | 98<br>00:05:12,080 --> 00:05:18,000<br>Obviamente, ha crecido desproporcionadamente las tareas de ventas | 98<br>00:05:12,080 --> 00:05:18,000<br>Obviamente, las tareas de ventas han crecido exponencialmente |

Other significant percentages were over-translation (9%) and mistranslation errors (7%).

Regarding severity, most errors do not impact text comprehension and are easily corrected. In the example in Table 4, which shows the corrected source text and its machine translation, we can see no mistranslations, and the text can be easily understood:

Table 6. Example 6

| ORIGINAL CLOSED CAPTIONS (Subcorpus 1) | MT from subcorpus 1 | PRE-EDITED MT |
|---|---|---|
| 89<br>00:04:35,910 --> 00:04:42,470<br>My focus is on the tools,<br>also known as a toolkit,<br><br>90<br>00:04:42,470 --> 00:04:46,000<br>the web-based on text product,<br><br>91<br>00:04:46,000 --> 00:04:50,877<br>from what also the integration part of this product<br><br>92<br>00:04:50,943 --> 00:04:53,817<br>into management solutions. | 89<br>00:04:35,910 --> 00:04:42,470<br>Mi atención se centra en las herramientas,<br>también conocido como conjunto de herramientas,<br><br>90<br>00:04:42,470 --> 00:04:46,000<br>el producto de texto basado en la web,<br><br>91<br>00:04:46,000 --> 00:04:50,877<br>de lo que también la parte de integración<br>de este producto<br><br>92<br>00:04:50,943 --> 00:04:53,817<br>en soluciones de gestión. | 89<br>00:04:35,910 --> 00:04:42,470<br>Mi atención se centra en las herramientas,<br>también conocidas como paquete de herramientas,<br><br>90<br>00:04:42,470 --> 00:04:46,000<br>el producto de texto basado en la web,<br><br>91<br>00:04:46,000 --> 00:04:50,877<br>que forma parte de integración<br>de este producto<br><br>92<br>00:04:50,943 --> 00:04:53,817<br>en soluciones de gestión. |

Finally, with the data collected on the severity of errors, a paired difference test was conducted to determine the significant differences between text A (without correcting the automatic source text) and text B (correcting the source).

To establish a statistically significant variation, the p-value must be lower than 0.05. In this case, the variations for both minor errors (*p-value* = 0.04 < 0.05) and major (*p-value* = 0.03 < 0.05) or critical errors (*p-value* = 0.03 < 0.05) were lower than 0.05. Thus, the differences between pre-edited and non-pre-edited closed captions are statistically significant, even more so for major and critical errors, for which the value is even lower.

The implications of these findings are substantial. The statistical significance indicates that the process of pre-editing has a clear and measurable impact on the quality of auto-generated captions. Specifically, the lower *p-values* for major and critical errors suggest that pre-editing is particularly effective in reducing more severe errors, which are likely to have a greater impact on the comprehensibility and usability of the captions. This reinforces the necessity of pre-editing in ensuring high-quality captions and supports the argument for its integration into workflows where accuracy is paramount. Additionally, it highlights the importance of investing time and resources in the pre-editing process to achieve a significant improvement in caption quality, thereby enhancing accessibility for users who rely on these captions

Table 7. Paired difference test

| | Paired difference test | | | | | t | gl | P-VALUE |
|---|---|---|---|---|---|---|---|---|
| | Average | Standard deviation | Standard Error Average | 95% from confidence interval for the difference | | | | |
| | | | | Lower | Upper | | | |
| MINOR_A - MINOR_B | 370,186335403727000 | 970,080357816360000 | 202,275737398356000 | -49,307868677127100 | 789,680539484580000 | 1,830 | 22 | 0,040 |
| CRITICAL_A - CRITICAL_B | 79,503105590062100 | 204,532413340282000 | 42,647956323328400 | -8,943342430030800 | 167,949553610155000 | 1,864 | 22 | 0,038 |
| MAJOR_A - MAJOR_B | 55,900621118012400 | 136,814612828658000 | 28,527818828413200 | -3,262454036275830 | 115,063696272301000 | 1,960 | 22 | 0,031 |

Thus, we believe that we have achieved objective data demonstrating that the most recurrent errors in the closed captions generated by YouTube are caused by wrong segmentation, punctuation, and spelling mistakes. These errors were replicated in the machine-translated text. In general, the errors generated in closed captions are minor and do not generate comprehension difficulties. Similarly, it should be noted that segmentation errors in closed captions generate, in addition to those already mentioned, major and critical grammar and mistranslation errors in machine-translated text, generally caused by the break-up of units of meaning.

However, in light of the aforementioned results, we have observed that by correcting the closed captions, a higher-quality machine-translated text is generated, finding mostly minor errors. These results are consistent with the results of the reception questionnaire, as we shall see in the next section. Furthermore, this improvement in quality directly relates to the productivity results discussed earlier, demonstrating that the pre-editing process not only enhances the accuracy of the machine-translated text but also contributes to overall efficiency. By reducing the number of major and critical errors, the time required for post-editing is significantly decreased. This aligns with our first objective and confirm that the integration of pre-editing and MT is both time-effective and beneficial for producing accessible and high-quality content in scientific communications.

## 3.2. Results of the survey

In total, 26 responses were collected. 60% of the participants were aged 18-29, 8% were 30-39 years old, 8% were 40-49 and 24% were over 50. 80% were female, 20% were male, and 88% were native Spanish speakers. The following figure shows the distribution by profession.

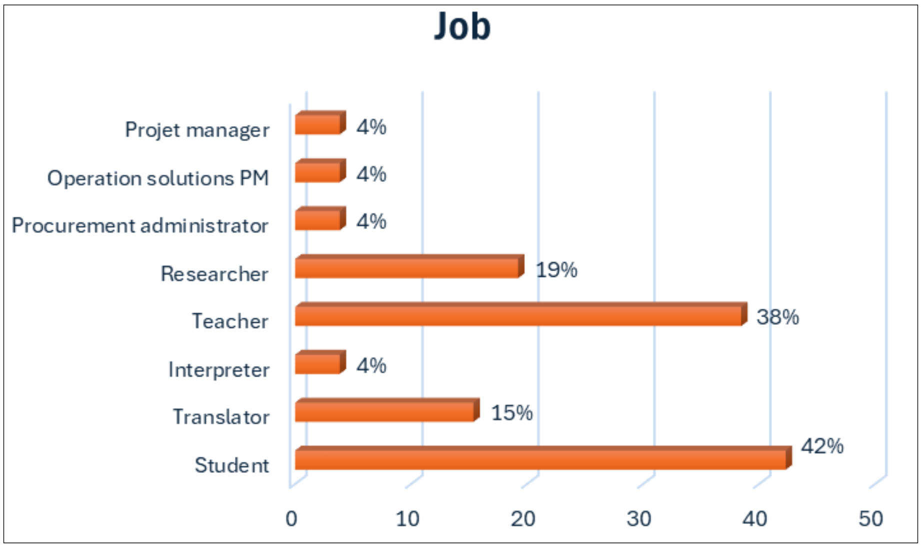


Figure 6. Profession of respondents to the survey

As for the responses, in Video 1, 53.8% of respondents ranked the comprehension level of raw machine-translated subtitles as neutral (3 out of 5), 23.1% as good, and 15.4% as very good, as shown in Figure 7.

Question 1: Video 1: Rate from 1 to 5 (5 being very good and 1 being very bad) whether the auto-generated captions are understandable (assuming that the receiver did not understand the English audio).

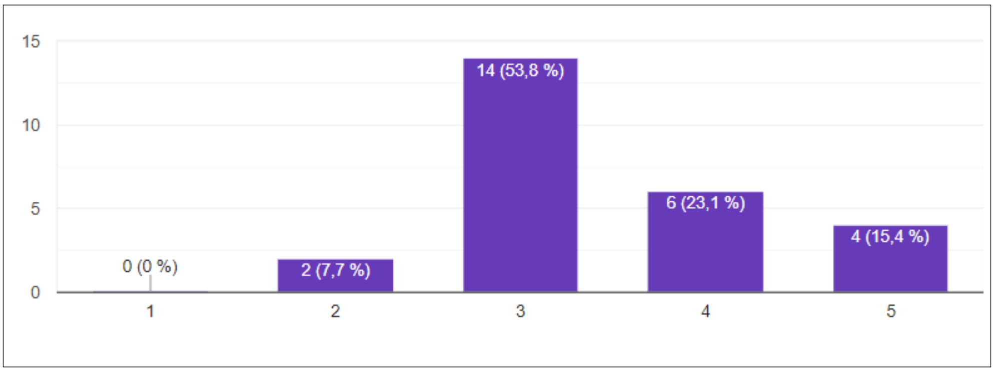


Figure 7. Degree of understanding of machine translation from non-pre-edited captions

In contrast, in Video 2, 57.7% of respondents ranked the comprehension level of pre-edited and machine-translated auto-generated captions as very good and 30.8% as good, while one respondent ranked them as neutral and two as very bad, as shown in Figure 8.

Question: Video 2: Rate from 1 to 5 (5 being very good and 1 being very bad) whether the auto-generated captions are understandable (assuming that the receiver did not understand the English audio).

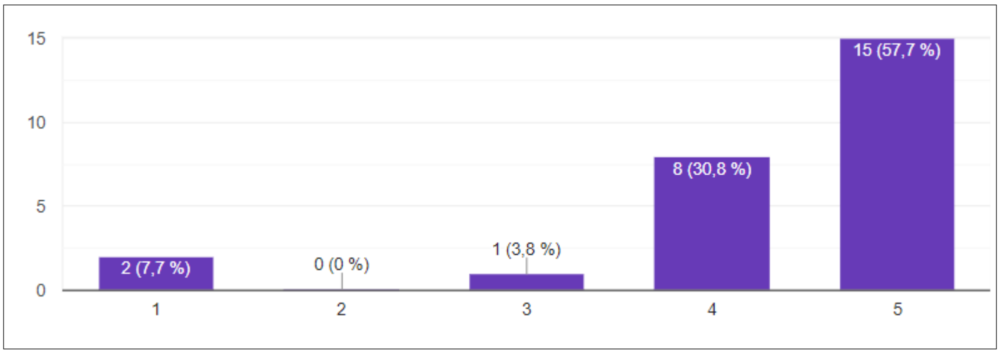


Figure 8. degree of understanding of machine translation from pre-edited auto-generated captions

Regarding linguistic correctness (Figure 9 and 10), for Video 1 (non-pre-edited and machine translation captions), 34.6% of the respondents thought the raw MT was neutral, 11.5% and 30.8% ranked it as very bad or bad, and 15.4% and 7.7% as good or very good, respectively.

Question: Video 1: Rate from 1 to 5 (5 being very good and 1 being very bad) the degree of linguistic correctness of the automatic subtitles in Spanish.

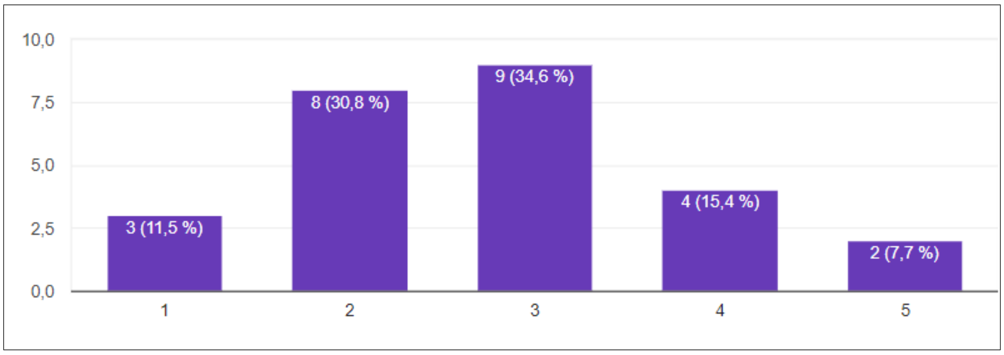


Figure 9. Degree of language correctness of machine translation from non-pre-edited auto-generated captions

Regarding the pre-edited and machine-translated auto-generated captions (Video 2), most respondents (53.8% and 34.5%, respectively) ranked their linguistic correctness as good or very good:

Question: Video 2: Rate from 1 to 5 (5 being very good and 1 being very bad) the degree of linguistic correctness of the automatic subtitles in Spanish.

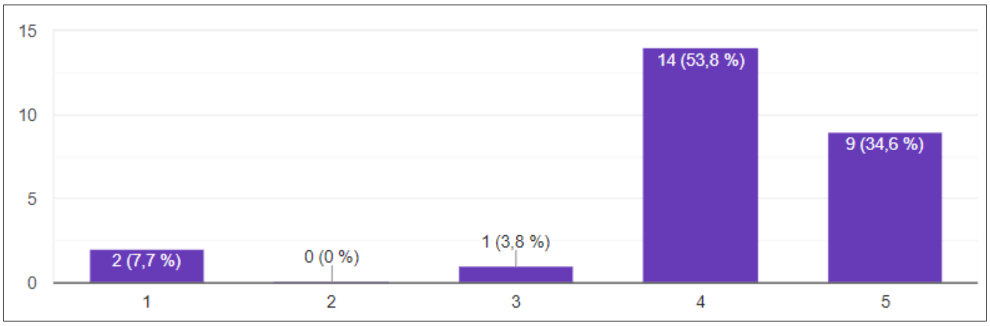


Figure 10. Degree of language correctness of machine translation from pre-edited auto-generated captions

Regarding the technical aspects of subtitles, the non-pre-edited and MT auto-generated captions were evaluated as bad by 34.6% of respondents, as neutral by 26.9%, and the rest were equally distributed between very bad, good, or very good, as shown in Figure 11:

Question: Video 1: Please rate from 1 to 5 (5 being very good and 1 being very bad) whether the subtitles are technically correct (screen time, segmentation, etc.).

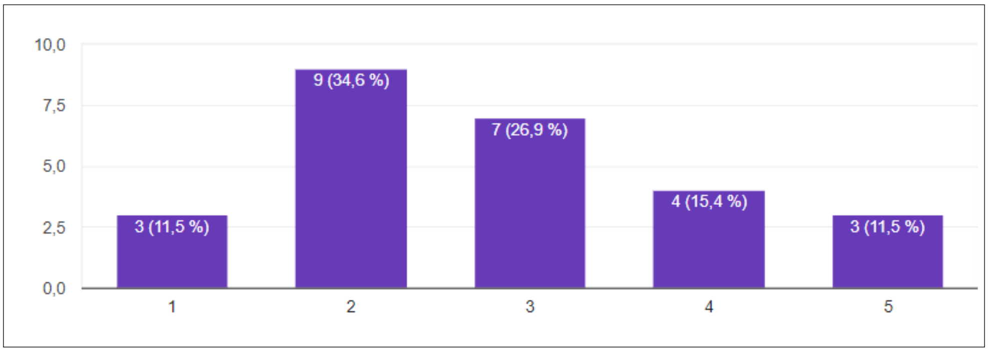


Figure 11. Degree of technical correctness of machine translation from non-pre-edited auto-generated captions

The pre-edited and machine-translated auto-generated captions (Video 2) were considered technically very good by 46.2% of respondents, good by 26.9%, neutral by 15.4%, and bad or very bad by 7.7% and 4%, respectively (Figure 12).

Question: Video 2: Please rate from 1 to 5 (5 being very good and 1 being very bad) whether the subtitles are technically correct (screen time, segmentation, etc.).

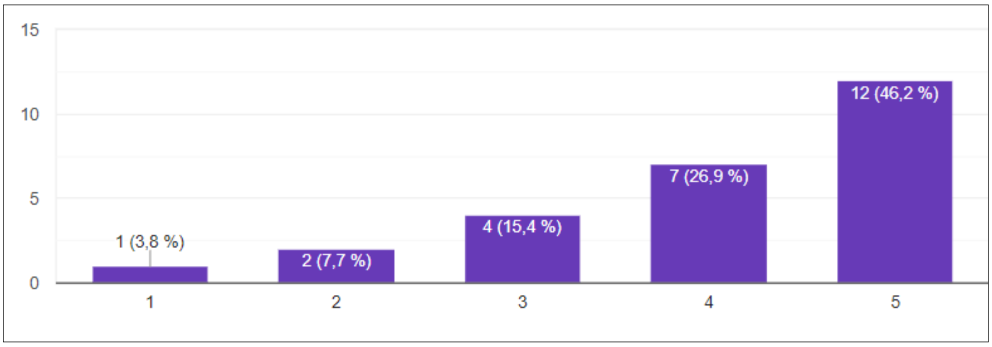


Figure 12. Degree of technical correctness of machine translation from pre-edited auto-generated captions

As for general comprehension, the pre-edited auto-generated captions were logically best understood by 88% of respondents, although 12% considered both videos equally understandable (Figure 13):

Question: Which of the videos did you understand best?

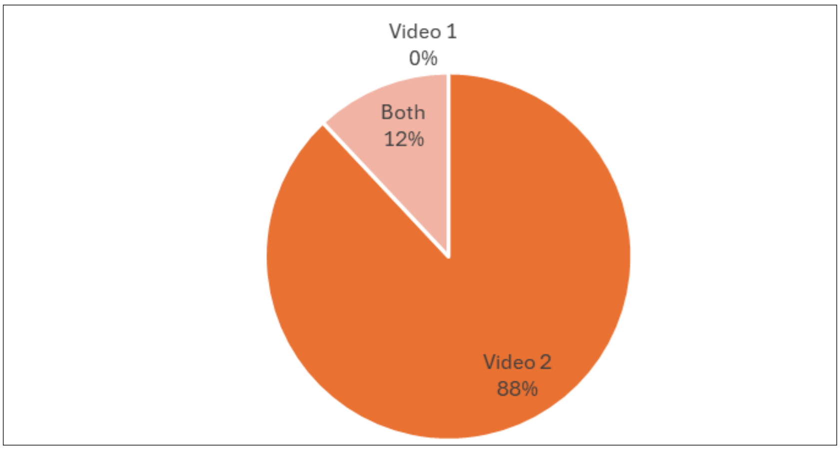


Figure 13. General comprehension of the two videos (with machine-translated subtitles from raw automatic captions and from corrected automatic captions)

Technically, the pre-edited auto-generated captions were easier to understand from the perspective of segmentation and other aspects (Figure 14).

Question: Which of the videos did you find easiest to follow from the point of view of segmentation, subtitle exposure time, etc.?

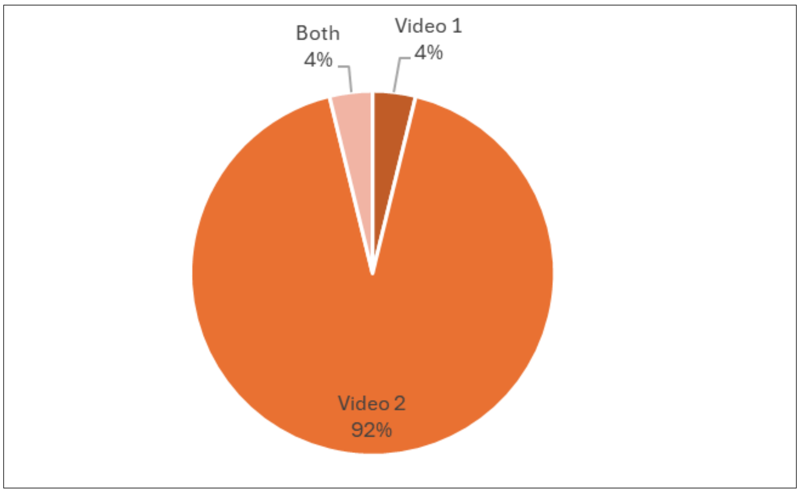


Figure 14. Comprehension of the two videos (machine-translated subtitles from raw automatic captions and from corrected automatic captions) according to technical aspects

Finally, regarding whether auto-generated captions can be used to understand academic presentations, 56% gave a positive answer and 38% thought maybe in the future, while one respondent said no (Figure 15):

Question: Do you think auto-generated captions can be used to understand videos of academic communications (at conferences, etc.)?

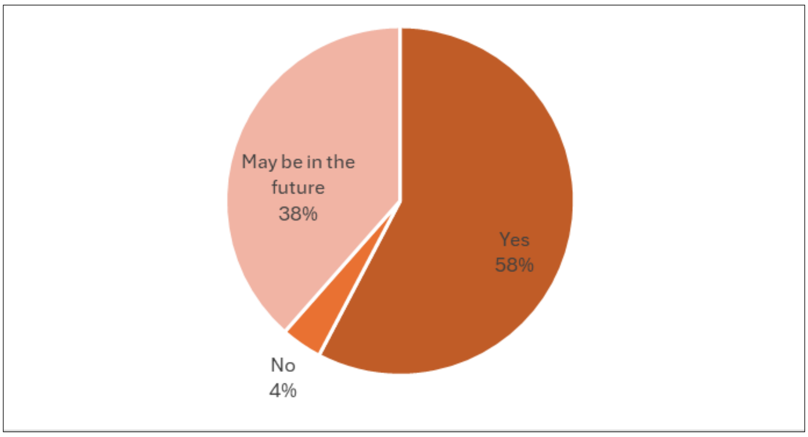


Figure 15. Respondent opinions on whether automatic captions can work in conferences

It is worth mentioning the reasons respondents gave to support this last answer, as they reveal a nuanced perspective. They highlighted several key issues, including syntactic and lexical errors, discourse disruptions due to filler words, and improper segmentation, which collectively hinder comprehension and coherence. Some respondents noted that while the raw subtitles provided a basic level of understanding, especially when the source language was not understood, the presence of grammatical and lexical errors often caused distractions. Additionally, the system's current inability to effectively handle idioms, proper names, and present information in a user-friendly manner was seen as a significant limitation for academic contexts. Despite these shortcomings, some respondents found the auto-generated captions generally understandable and useful for facilitating communication between speakers of different languages. The second video, in particular, was noted for its accuracy and effectiveness in conveying information, unlike the first video, which suffered from omissions and errors. This variability underscores the importance of continuous improvements in speech recognition and translation models to meet academic standards.

The implications of these findings suggest that while auto-generated captions can provide a foundational level of accessibility, their current state requires significant enhancements to achieve the desired quality for academic use. The need for human intervention, in terms of pre-editing and post-editing, remains crucial to ensure accuracy and coherence. These results directly address our second research objective by illustrating both the potential and the limitations of machine-translated captions in academic settings. The conclusions drawn from the reception study indicate that although MT systems are advancing, they are not yet fully reliable for high-stakes academic communications without substantial human oversight. This aligns with our broader aim of evaluating the usefulness and efficiency of these technologies in making academic content accessible, thereby confirming that while MT has potential, it is not yet sufficient on its own.

## 4. Conclusions

This investigation sheds light on the effectiveness of pre-editing in enhancing the quality of machine translation outputs and provides valuable insights into the role of automation in the context of audiovisual translation workflows and accessibility. The findings of this study may have implications for future advancements in machine translation technology and its applications in academic and professional settings.

Regarding the first objective, we aimed to demonstrate the usefulness of MT in oral scientific communication in the EN>ES language combination. The results explained above are consistent with those of previous studies by Karakanta et al. (2022a): most errors in machine-translated subtitles arise from the wrong segmentation of automatically generated captions. Spelling errors (in proper names, acronyms, the pronoun "I"), and punctuation errors (no question or exclamation marks) are replicated in the machine translation and the separation of units of meaning causes further issues. In addition, misrecognition errors (due to sound problems, different languages being spoken in the videos, etc.) result in the software inserting random words or terms. It should be noted that the system does not take into account the context and textual coherence, aspects which are already being considered in other MT software (such as MateCAT). In this context, it is essential and cost-effective to correct the automatic source text before using machine translation.

Errors in the source (repetitions or backtracking by the speakers) are maintained, and technical and linguistic conventions for subtitles are obviously not respected by the software and the machine translation, but the reception survey shows that users have different opinions on the impact and severity this has on their comprehension of academic presentations.

With regard to the second objective, the reception of machine-translated oral communications was investigated. To this end, a survey was conducted to compare the comprehension and quality levels of the two videos. In the first video, the auto-generated captions were pre-edited before being machine-translated; in the second video auto-generated captions were not pre-edited. As can be seen in the results, the level of comprehension and quality was higher when the auto-generated captions in the original language were pre-edited before being automatically translated, which leads us to think that, in most instances, the problem lies in the original transcription, rather than in the translation itself, on the part of the translator. Although there was no complete linguistic correction in the machine-translated subtitles, viewers considered them understandable.

The question remains whether we are headed towards a context in which accessibility is preferred to linguistic and technical correctness and perfect comprehension of academic content, or whether users need to raise their expectations in the same way that machine translation requirements should be higher. As to further research in this field, reception studies of automatic and machine-translated subtitles with different target groups would help understand viewers' expectations, which are likely in correlation with their knowledge of audiovisual translation standards and audiovisual content in general.

## References

Bennet, Winfield S. and Jonathan Slocum. 1985. "The LRC machine translation system". In *Computational Linguistics, 11*, 111-121.

Berner, Sam. 2003. "Lost In Translation: Cross-Lingual Communication, And Virtual Academic Communities". In *CiteSeer*. Retrieved from: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.130.2973>.

Bowker, Lynne and Jennifer Pearson. 2002. *Working with specialized language.* London: Routledge.

Briva-Iglesias, Vicent. 2021. "Traducción humana vs. traducción automática: análisis contrastivo e implicaciones para la aplicación de la traducción automática en traducción jurídica". *Mutatis Mutandis*, Vol. 14, Extra 2, 571-600.

Che, Xiaoyin, Sheng Luo, Haojin Yang and Christoph Meinel. 2017. “Automatic Lecture Subtitle Generation and How It Helps”. 2017 IEEE 17th International Conference on Advanced Learning Technologies (ICALT), Timisoara, Romania, 34-38, doi: 10.1109/ICALT.2017.11. Retrieved from: < https://ieeexplore.ieee.org/abstract/document/8001709>.

Córdoba Mondéjar, Iris, Celia Rico Pérez, María Ortiz Jiménez, Anna Doquin de Saint Preux, Juan José Arevalillo Doval, Manuel Arcedillo Jiménez and Jorge Cabero. (2015). *Estudio de viabilidad para la implantación de la traducción automática en la empresa.* Madrid: Plataforma del español. Retrieved from: <http://hdl.handle.net/11268/4362>

Corpas Pastor, Gloria and Miriam Seghiri. 2007. ReCor. Málaga.

de Higes Andino, Irene. 2023. “Automatic bilingual subtitles in Spanish tv broadcast: the case of l’Informatiu – Comunitat Valenciana”. In Media for all 10, University of Antwerp. Retrieved from: <https://www.uantwerpen.be/en/conferences/media-for-all-10/program/6july/room2/>.

Díaz Cintas, Jorge and Aline Remael. 2021. *Subtitling Concepts and Practices*. London: Routledge.

Díaz Cintas, Jorge and Serenella Massidda. 2019. “Technological advances in audiovisual translation”. In *The Routledge Handbook of Translation and Technology*, edited by O’Hagan, Minako, 255–270. London/New York: Routledge.

EST. 2022. Closed captioning. Retrieved from: <https://www.hf.uio.no/english/research/news-and-events/est22/registration/video-presentations.html>

EUATC. 2017. *European Language Industry Survey*. Retrieved from:<https://ec.europa.eu/info/sites/info/files/2017_language_industry_survey_report.en.pdf>.

EUATC. 2020. *2020 European Language Industry Survey*. Retrieved from: <https://ec.europa.eu/info/sites/default/files/2020_language_industry_survey_report.pdf>.

European Commission. 2022. “AVT workflows and the role of automation. From Translation To Accessibility”. EMT Train The Trainer Summer School, 2nd edition. Retrieved from: <https://www.youtube.com/playlist?list=PLLqIRaiVCGCSD9Z3ek8Y-I2TIi9O9vvWW>

Farkas, Anna and Renáta Németh. 2022. “How to measure gender bias in machine translation: Real-world oriented machine translators, multiple reference points”. In *Social Sciences & Humanities Open*, Volume 5 (1). Retrieved from: <https://doi.org/10.1016/j.ssaho.2021.100239>.

Gámez, Ruth and Fernando Cuñado. 27 February 2019. *Traducción Jurídica*. Retrieved from: <https://traduccionjuridica.es/cual-es-la-capacidad-de-trabajo-de-un-traductor/>.

Gaspari, Federico, Almaghout, Hala, & Doherty, Stephen. 2015. “A survey of machine translation competences: Insights for translation technology educators and practitioners”. *Perspectives*, 23(3), 333-358. Retrieved from: <https://www.tandfonline.com/doi/full/10.1080/0907676X.2014.979842>.

Group Multilingual Technologies. 2015. *Multidimensional Quality Metrics (MQM)*. Retrieved from: <http://www.qt21.eu/quality-metrics/>.

Hagström, Hanna and Jan Pedersen. 2022. “Subtitles in the 2020s: The Influence of Machine Translation”. In *Journal of Audiovisual Translation*, 5(1), 207–225. Retrieved from: <https://jatjournal.org/index.php/jat/article/view/195/84>.

Hernández Mercedes, Pilar. 2002. “En torno a la traducción automática”. In *Cervantes, 2*, 101-117.

Hutchins, William John and Harold L. Somers. 1992. *An Introduction to Machine Translation.* Cambridge: Cambridg University Press.

Karakanta, Alina, M. Luisa Bentivogli, Mauro Cettolo, Marco Gaido, Matteo Negri and Marco Turchi. 2022a. “Post-editing in Automatic Subtitling: A Subtitlers’ Perspective”. *Proceedings of the EAMT 2022*. Retrieved from: <https://aclanthology.org/2022.eamt-1.29.pdf>.

Karakanta, Alina, M. Luisa Bentivogli, Mauro Cettolo, Marco Gaido, Matteo Negri and Marco Turchi. 2022b. “Towards a methodology for evaluating automatic subtitling”. Proceedings of the EAMT 2022. Retrieved from: < https://aclanthology.org/2022.eamt-1.57.pdf>.

Läubli, Samuel, Sheila Castilho, Graham Neubig, Rico Sennrich, Qinlan Shen and Antonio Toral. 2020. “A set of recommendations for assessing human-machine parity in language translation.” In *Journal of Artificial Intelligence Research*, 67, 653-672. https://doi.org/10.1613/jair.1.11371

Llisterri, Joaquim. 2009. *Tecnologías lingüísticas: Las tecnologías del texto.* Retrieved from: <http://www.corpus.unam.mx/cursocorpus/TecnoLing_Lynx02.pdf>.

Lommel, Arle. 2018. “Metrics for Translation Quality Assessment: A Case for Standardising Error Typologies”. In *Translation Quality Asessment. From Principles to Practice*, edited by Moorkens, Joss, Sheila Castilho, Federico Gaspari and Stephen Doherty, 109-127. Springer, 109-127.

Matusov, Evgeny, Patrick Wilken and Yota Georgakopoulou. 2019. “Customizing Neural Machine Translation for Subtitling”. In *Proceedings of the Fourth Conference on Machine Translation (Volume 1: Research Papers)*, 82–93, Florence, Italy. Association for Computational Linguistics. Retrieved from: < https://aclanthology.org/W19-5209.pdf>.

Moorkens, Joss. 2017. “Under Pressure: Translation in Times of Austerity”. In *Perspectives: Studies in Translation Theory and Practice, 25*(3), 464-477.

Moorkens, Joss, Sheila Castilho, Federico Gaspari and Stephen Doherty. 2018. *Translation Quality Assessment. From Principles to Practice*. Springer.

MQM. 2022. Welcome to the MQM Core Typology. Retrieved from: <http://themqm.info/typology/>.

Papi, Sara, Marco Gaido, Alina Karakanta, Mauro Cettolo, Matteo Negri, Marco Turchi. 2023. “Direct Speech Translation for Automatic Subtitling”. In *Transactions of the Association for Computational Linguistics*, 11, 1355–1376. Retrieved from: <https://doi.org/10.1162/tacl_a_00607>.

Parra Escartín, Carla and Marie-Josée Goulet. 2021. “When the post-editor is not a translator: can machine translation be post-edited by academics to prepare their publications in English?”. In *Translation Revision and Post-Editing*, edited by Koponen, Maarit, Mossop, Brian, S. Robert, Isabelle and Giovanna Scocchera, 89-106. London/New York: Routledge.

Parra Escartín, Carla. 2018. “¿Cómo ha evolucionado la traducción automática en los últimos años?”. In *La Linterna del Traductor. La revista multilingüe de ASETRAD*. Retrieved from: <http://www.lalinternadeltraductor.org/n16/traduccion-automatica.html>.

Romero Fresco, Pablo. 2011. *Subtitling Through Speech Recognition: Respeaking*. Routledge.

Savoldi, Beatrice, Marco Gaido, Luisa Bentivogli, Matteo Negri and Marco Turchi. 2021. “Gender Bias in Machine Translation”. In *Transactions of the Association for Computational Linguistics.* Vol. 9, 845–874. doi: https://doi.org/10.1162/tacl_a_00401.

Sánchez Ramos, María del Mar and Celia Rico Pérez. 2020. *Traducción automática: Conceptos clave, procesos de evaluación y técnicas de posedición.* Granada: Comares.

Seghiri, Miriam. 2011. “Metodología protocolizada de compilación de un corpus de seguros de viajes: aspectos de diseño y representatividad”. In *RLA. Revista de lingüística teórica y aplicada*, Vol. 49 (2), 13-30.

Stanovsky, Gabriel Noah Smith and Luke Zettlemoyer. 2019. “Evaluating Gender Bias in Machine Translation”. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 1679–1684, Florence, Italy. Association for Computational Linguistics. Retrieved from: <https://arxiv.org/pdf/1906.00591.pdf?ref=https://githubhelp.com>.

Systran. (2021). *Systran Translate*. Retrieved from: <https://www.systran.net/en/translate/>.

Valor Miró, Juan Daniel, Joan Albert Silvestre-Cerdà, Jorge Civera, Carlos Turró and Alfons Juan. 2015. “Efficient Generation of High-Quality Multilingual Subtitles for Video Lecture Repositories”. In: Conole, G., Klobučar, T., Rensing, C., Konert, J., Lavoué, E. (eds), Design for Teaching and Learning in a Networked World. EC-TEL 2015. Lecture Notes in Computer Science, vol 9307. Springer, Cham. Retrieved from: < https://doi.org/10.1007/978-3-319-24258-3_44>.

Wan, David, Zhengping Jiang, Chris Kedzie, Elsbeth Turcan, Peter Bell and Kathy McKeown. 2020. “Subtitles to Segmentation: Improving Low-Resource Speech-to-TextTranslation Pipelines.” In *Proceedings of the workshop on Cross-Language Search and Summarization of Text and Speech (CLSSTS2020)*, 68-73. Retrieved from: < https://arxiv.org/pdf/2010.09693>.

Wu, Yonghui, Mike Schuster, Zhifeng Chen, Quoc V. Le, Mohammad Norouzi, Wolfgang Macherey, Maxim Krikun, Yuan Cao, Qin Gao, Klaus Macherey, Jeff Klingner, Apurva Shah, Melvin Johnson, Xiaobing Liu, Łukasz Kaiser, Stephan Gouws, Yoshikiyo Kato, Taku Kudo, Hideto Kazawa, Keith Stevens, George Kurian, Nishant Patil, Wei Wang, Cliff Young, Jason Smith, Jason Riesa, Alex Rudnick, Oriol Vinyals, Greg Corrado, Macduff Hughes and Jeffrey Dean. 2016. “Google’s neural machine translation”. In *ArXiv:1609.08144 [Cs]*. Retrieved from: <https://arxiv.org/abs/1609.08144>

YouTube Help. (n.d.). *Add subtitles & captions*. Retrieved from <https://support.google.com/youtube/>.