

Machine translation (MT) and post-editing (PE) are still considered, in many sectors, enemies of the creative freedom traditionally associated with the audiovisual genre and its translation. Its growing presence in the market is unheralded due to the widespread rejection by a large part of the professional community. However, it is necessary to study the benefits of their implementation in all areas, from quality to productivity, considering all the parties involved in the process.

In order to study the differences between the traditional Audiovisual Translation (AVT) process without tools and the process implemented with MTPE, a pilot study was carried out aimed at comparing the results in the English-Spanish translation of a series of 3-minute news and sports clips with different technical characteristics. The results will allow us to obtain an objective comparative description of both processes in order to better know the strengths and weaknesses of each one.

KEY WORDS: Machine Translation Post-editing (MTPE), audiovisual translation (AVT), comparative study, sports and news clips.

Machine Translation and Post-editing in AVT: A Pilot Study on a Rising Practice

VERÓNICA ARNÁIZ-UZQUIZA
Universidad de Valladolid
ORCID: 0000-0003-3703-6480

PAULA IGAREDA GONZÁLEZ
Universitat de Vic-Universitat Central de Catalunya
ORCID: 0000-0001-5412-199X

Traducción Automática y Posedición en TAV: estudio piloto de una práctica en alza

La traducción automática y la posedición (TAPE) siguen considerándose, a menudo, enemigas de la libertad creativa tradicionalmente asociada al género audiovisual y a su traducción. El rechazo que su creciente presencia suscita entre gran parte de la comunidad profesional hace que la traducción audiovisual (TAV) no sea una de las disciplinas donde el uso de la TAPE se aborde abiertamente. Sin embargo, es necesario estudiar los beneficios de su implantación a todos los niveles, considerando todas las partes y a todos los implicados en el proceso.

Con el fin de estudiar las diferencias entre el proceso tradicional de TAV –sin herramientas– y el proceso implementado con TAPE, se ha llevado a cabo un estudio piloto destinado a comparar los resultados en la traducción inglés-español de una serie de vídeos de noticias y deportes de 3 minutos de duración con diferentes características técnicas. Los resultados permitirán obtener una descripción comparativa objetiva de ambos procesos.

PALABRAS CLAVE: Traducción Automática y Posedición (TAPE), traducción audiovisual (TAV), estudio comparativo, noticias y deportes.

198 **1. INTRODUCTION**

The presence of tools in the development of most audiovisual translation modalities has, to date, made it natural for various purposes, from the configuration of a work environment to the systematisation of content integration. In all cases it seeks to systematise translators' work and make it more profitable, guaranteeing the quality of production and increasing—actively (voluntarily) or passively (involuntarily)—productivity.

However, the growing incorporation of some of these tools in professional tasks, and the practices associated with them, has been—and, in some cases, continues to be—polemic. The systematisation of templates, the incorporation and expansion of speech recognition and the automation of subtitle synchronisation are just some of the controversial elements surrounding audiovisual translation (AVT) (Díaz-Cintas & Massidda, 2020; Fernández Moriano, 2019; Athanasiadi, 2021). Although technical improvement has allowed the progressive acceptance and incorporation of some of these tools—not yet generalised—in professional practice, the growing presence of tools that address the increase of productivity is still controversial. Machine translation (MT) and post-editing (PE), widely used in certain areas of translation, are in many sectors still considered enemies of the creative freedom traditionally associated with the audiovisual genre and its translation, often facing wholesale rejection that ignores the possible advantages (ATRAE, 2021; AVTE, 2021) they could bring to the practice. As is often the case, the needs of the market once again take precedence over academic contexts, forcing professionals to face realities that are not yet present in training programmes.

As already experienced in other cases (Bywood et al., 2017), the presence of MTPE in most translation fields, and its significant technological improvement in recent years, makes it necessary to consider the role that these tools may also play in the case of AVT. To this end, it is necessary to study the benefits of their implementation in all areas, from quality to productivity, considering all the parties involved in the process and including future professionals in these realities so that they are aware of the future of the market.

2. FROM AUDIOVISUAL TRANSLATION TO MACHINE TRANSLATION & POST-EDITING**2.1. Audiovisual Translation Practice: Technologies and Controversies**

The practice of audiovisual translation, aimed at processing texts in “less traditional” formats, makes of it a discipline which, almost from its origins, has used different tools to facilitate the treatment of multimedia and audiovisual texts, systematise (automate) and simplify processes, and homogenise (standardise) production. Subtitling programmes on the cloud (Ooona, Amara, SubtitleNext, eCaption), automatic transcription programmes [Dragon Naturally Speaking, Webcaptioner, Microsoft, oTranscribe—audio or live speech only—and AppTek, Trint and Limecraft, created specifically for video transcription and automatic subtitle generation (Agulló García, 2020)] and subtitling editors with MT (MemoQ Video Preview Tool, SDL Trados Studio Subtitling) are among the tools that have completely modified the AVT scene in the last decade. Additionally, there are also companies already working in the cloud (ZOOsubs, Netflix, Iyuno, Plint), generating automatic subtitles

(Media Studio Omniscien) or using advanced automated speech recognition (ASR) technology (AppTek, 2021).

Although most of these tools have been naturally integrated into professional practice, it is also true that the emergence of tools and processes that are not exclusive to this modality, such as MTPE, have generated a growing debate within the profession (Mejías-Climent & de los Reyes Lozano, 2021). But controversy has accompanied many other technological advances in AVT too:

1. The incorporation and expansion of speech recognition technology has also been subject to debate (Díaz-Cintas & Massidda, 2020, Fernández Moriano, 2019, Athanasiadi, 2021). Speech recognition is a tool that automates and streamlines the manual work of transcription into subtitles (Dragon Naturally Speaking, Microsoft), i.e. it converts audio and video into text, albeit with some limitations, as it depends on the quality of the audio and the pronunciation of the speakers, the language and so forth.
2. The systematisation of templates: the use of already timed and spotted template files in the creation of subtitles from the same source audio assets, typically English, was one of the greatest innovations in the industry at the end of the last century (Georgakopoulou, 2019). This practice streamlines processes, eliminates duplication of work, reduces direct costs, improves timelines and facilitates quality control of large volumes of work (Georgakopoulou, 2006). The emergence of software that automatically synchronises video and script.
3. The incorporation and extensive use of AVT workflows in the cloud, especially

subtitling programmes that deviate from the “less flexible desktop-based solutions” (Bolaños et al., 2021, p. 3), such as Ooona and SubtitleNext.

However, in most cases, the debates that arise within the audiovisual translation community with the incorporation and implementation of each new technology fade away as their progressive implementation gains ground in the market, though it may take some time. More than two decades after their introduction and systematisation, template files for interlingual subtitling, for example, “are still a topic of debate among language service providers and subtitlers” (Georgakopoulou, 2019, p.137). The Spanish Audiovisual Translation and Adaptation Association ATRAE has recently issued a statement on the use of the pivot languages and templates, encouraging professionals to reject such practices (2022). ATRAE states that original and pivot language scripts should be an aid for professionals who already know the original language and culture of the product, as they are the ones who will always best convey the original message. The Association argues that, on the grounds of an alleged shortage of professionals for certain language combinations, this practice is ongoing and rising. Time will tell whether this debate also fades away in the future.

In the case of MTPE, however, this process does not seem so evident. As Patrick Cadwell, Sharon O’Brien & Carlos S. C. Teixeira (2017) point, the negative perception around its practice may be explained by the number of errors it can cause, its mandatory application to translation processes by companies and the negative consequences its use entails for human translation. In the end, quality, productivity and rates lie at the heart of the conflict.

200 2.2. Machine Translation and Post-editing Practice

In 2009, Adriane Rinsche & Nadia Portera-Zanotti conducted a study to find out how many language service providers in the EU member States used MTPE and obtained a 3.26% affirmative response. In 2015, the UAB (Autonomous University of Barcelona) (Torres-Hostench, Presas & Cid-Leal, 2016) conducted a similar study in the Spanish market and 47.3% said they used MTPE, albeit for few projects. According to these data, localisation departments of large multinationals and translation service providers are progressively incorporating MTPE into their translation processes (Aranberri, 2014), making of it one of “the main growth-supporting activities” (ELIS, 2023). As some national associations recall (ASETRAD, 2021; AVTE, 2021), MTPE is useful for texts with structured language, so it is usually applied to texts of a technical nature that are characterised by their objectivity and do not contain idiomatic turns of phrase, puns or creative use of language. Originally aimed at translating technical documentation, MTPE has been progressively integrated in other fields, as has already happened with the use of computer-assisted translation tools (CAT) (Sánchez-Gijón, 2016).

If we focus on the development of MT, there are currently three main types of linguistic information-based machine translation in the market, commonly referred to as Rule Based Machine Translation (RBMT), Statistical Machine Translation (SMT) and Neural Machine Translation (NMT) (Parra Escartín, 2018), the latter being the most recent and sophisticated one that has been replacing the previous MT types thanks to the advances in artificial intelligence. Clients can now choose the quality

of their products with the incorporation of MT (Szarkowska et al., 2021), full PE or light PE (TAUS, 2010), depending on their needs. Lastly, companies have also developed their own style guides that the professional must follow according to the instructions previously indicated. This complex combination—MT improvements, different quality perception and production—have made MTPE explode the rules of the game when it comes to translation (Rico & González-Pastor, 2022).

2.2.1. AVT and Machine Translation Quality and Productivity

In terms of quality, in the industry “[it] is often viewed from the perspective of the process rather than the product” (Pedersen, 2017), while in academia quality is viewed more from a perspective of linguistics and intercultural communication (Szarkowska et al., 2021). Focusing on AVT, there is a limited number of widespread initiatives to assess quality in different AVT modalities: the NER model (Number of words, Editions, Recognition model) developed by Pablo Romero-Fresco & Juan Martínez (2015) used in live intralingual subtitling done via respeaking; the NTR model (Number of words, Translation, Recognition model) presented by Pablo Romero-Fresco & Franz Pöchhacker (2017) that assesses quality in live interlingual subtitling done via respeaking; and the FAR model, by Jan Pedersen (2017), which assesses the functional equivalence, the acceptability and the readability in interlingual subtitling quality. If we focus on the field of MTPE, although not specific to AVT, a large number of quality metrics have been developed, including automatic evaluation metrics such as BLEU, NIST, TER and METEOR (Martín-Mor & Sánchez-Gijón, 2016) and other approaches

like MQM (Multidimensional Quality Metrics) designed to focus on the need to objectively describe translation errors (Lommel et al., 2014), meeting the following criteria: accuracy, design, fluency, local convention, style, terminology and audience appropriateness. The combination of both AVT and MTPE in terms of evaluation metrics is still to be considered.

In terms of profitability, and in relation to MTPE, productivity in PE is measured through the cognitive, technical and temporal effort involved (Krings, 2001; Alves et al., 2016), although usually only time is measured (Aranberri, 2014, Kenny, 2022). Partly due to the heterogeneity of source text, it is very difficult to estimate the benefits in terms of money and time with the use of MTPE. According to the Spanish Association of Translators, Copy-editors and Interpreters (ASETRAD), productivity with full post-editing is 600-900 words/hour and is charged at around 66%-75% of the usual rate per word. Nevertheless, the Association admits that the key is to charge per hour the same as with conventional translation (ASETRAD, 2021). That said, no reference is included on the type and genre of texts that achieve these productivity figures. To our knowledge, no AVT research has yet reported on (word) rates / hour so far, as estimations may be difficult to calculate.

But the reluctance of an important part of the AVT community to accept and adopt MTPE into its practices is due to the complexity entailed in processing an audiovisual text for translation, since not only linguistic content but also its visual and acoustic configuration must be considered (Burchardt et al., 2016; Mejías-Climent & de los Reyes Lozano, 2021). Despite this fact, and at the other end of this debate, the industry has openly admitted its interest in increasing productivity, ensuring the quality of its work (and) saving costs, but also admits that MT without human

post-editing does not meet quality standards in almost any field of expertise and that it cannot (yet?) match human understanding of humour, meaning, nuance, tone, sarcasm, among other topics (Lionbridge, 2022).

3. AUDIOVISUAL TRANSLATION, MACHINE TRANSLATION & POST-EDITING TODAY

The MTPE-free market where AVT seemed to continue evolving isolated from this technique and its related processes is no longer MTPE-free. Many AVT practitioners had not yet considered the application of MTPE to audiovisual environments because “it is considered not to be ready yet, in addition to concerns about a lack of definition of quality, a tendency to mass production and worry of de-professionalisation” (Nunes Vieira & Alonso, 2018, p. 17).

In terms of quality, it seems that AVT professional associations agree in their rejection of MTPE (in AVT), as evidenced by the recent statements by ATRAE (2021) and ATAA (2022) or the AVTE manifesto (2021) which defended that “quality of MT output depends on many factors: the way the source text is written, how creative it is, whether the MT engine has been trained with appropriate data, etc.”. Meanwhile, Nora Aranberri (2014) asserts that end quality depends on the quality of the MT and the ability of the post-editor to transform that MT into a correct and fluent text. Regarding PE, AVTE (2021) also points out that “fixing a poor translation can take longer than translating the same text from scratch” and that the “unscrupulous use of MT will increasingly lead to brain drain and talent crunch”. In the same way, some detractors of MT argue that post-edited texts are stiff and lose fluency compared to manual/human translations, although other studies show opposite results (O’Curran,

202 2014). In some cases, evaluators seem unable to distinguish between translations obtained by the two methods and translations obtained by combining MT and PE (Carl et al., 2011; Läubli et al., 2013). In fact, translations resulting from MT with PE are sometimes scored more favourably, but this is not always the case (Guerberof Arenas, 2009). In any case, it seems evident MTPE is a rising practice in industry, even admitted by academia who estimate that, by 2030, will be used in most professional translation work (ibid 2023, p. 37).

AVT practitioners had long considered the application of MTPE in audiovisual environments unrealistic, based on former MT quality results. As more and more practices are now becoming known (Belles-Calvera & Caro Quintana, 2021; Karakanta, 2022), opinions also emerge, whether they are based on personal and professional experience or not. Maarit Koponen et al. (2020) already mention previous work and practitioners' feelings towards MT when they expressed that it limited their creativity, very close to what other studies concluded about literary translation (Moorkens et al., 2018) and localisation (Guerberof Arenas, 2013). There are also translators who express feeling "trapped by MT" (Bundgaard, 2017) and are concerned about the possible detriment to the quality of the final translation (Moorkens et al., 2018; Matusov et al., 2019) and the overall "uberization of translation", often accompanied by a widespread use of cloud-based workflows (Firat, 2021, p. 48).

The scenario shows that, especially when dealing with AV texts and MTPE processes, there are clear discrepancies between companies, associations, freelancers and academia. Companies defend that the process is much faster than human translation alone, and that it reduces the cost of services. In contrast, AVT

associations (ATRAE, 2021; AVTE, 2021; ATAA, 2022) are, as previously mentioned, mostly against MTPE as it is a process that affects the basics of AVT practice: quality. For their part, freelancers, often grouped under the umbrella of associations, face individual economic concerns and moral issues (Sakamoto, 2021). Finally, academia, which should be the first step in the process, always lags behind market needs (Gasparini et al., 2015).

In this regard, it should be noted that most translation and interpreting curricula are now integrating MTPE training for future professionals-to-be (Rico & González-Pastor, 2022), trying to fill the permanent gap there seems to be between education and market needs. As pointed out by Dorothy Kenny & Stephen Doherty (2014), professionals-to-be should be able to evolve and adapt whatever tool they wish to their practice; tools that should no longer be considered part of technological contents isolated in thematic courses but transversal contents in translators' training (González-Pastor & Rico, 2021). Nevertheless, it is clear that they are still reluctant to accept MTPE jobs mainly because they might have inherited preconceptions or negative opinions about MTPE while studying, as well as because of their lack of (or poor) experience during their formative years (Bruno et al., 2021).

In any case, leaving controversies aside, we need to accept that the presence of MTPE in the AVT market is gaining ground. Difficulties in adapting MTPE to different AV genres seem to be easy to address; and the number of AVT modalities that adopt MTPE is progressively increasing, moving from the isolated examples of interlingual subtitling (Burchardt, 2016; Karakanta, 2022; Martínez-Martínez & Vela, 2016; Varga, 2021; among others), to others such as audio description (Fernández-Torné

Table 1. Time length and word count of the videoclips selected for the pilot study per genre (S= Sports; N=News).

	S1	S2	S3	N1	N2	N3
Time Length	2:46	3:35	3:02	4:00	2:25	2:40
Word count	295	448	134	602	389	476

& Matamala, 2015; Vercauteren et al., 2021) and dubbing (Mejías-Climent & de los Reyes-Lozano, 2021). All of these, both genres and modalities, are examples that are still limited to preliminary research initiatives that in some cases try to shed some light onto a professional practice.

4. MACHINE TRANSLATION AND POST-EDITING IN AVT: PILOT STUDY

A multi-phase pilot study was designed to obtain information in comparing translators' productivity when translating and/or post-editing audiovisual texts. Bearing in mind that, to date, most studies have approached professional practices in the translation of "source texts produced with controlled language constraints" (Martín-Mor and Sánchez-Gijón, 2016, p. 173), it seemed necessary to focus on some audiovisual types that have not been addressed by academia, despite also being present in the market. This is the case of the translation of sports events and news programmes, although the latter has already been present in AVT studies often in relation to more technical—rather than linguistic—aspects of subtitling practices such as line-breaks, shot-changes, subtitling speed or live subtitling, among others¹.

¹ Previous research had already reported on the difficulties of applying MTPE to AVT due to domain and genre differences, lack of visual context, oral style or lack of context (Burchardt et al., 2016). On this basis, many research initiatives have focused so far on specific genres, with documen-

4.1. Set-up

4.1.1. Materials

Three different video clips were selected per AVT type, with a different word count per clip [Table 1]².

In the case of news, most clips follow the same scheme and word count. Although partly scripted, speech rate is higher than for spontaneous dialogues or scripted fictions, somehow conditioned by the fact that the visuals do not add extra information viewers need to process simultaneously.

In the case of sports events, however, the word count shows significant differences among the clips selected, conditioned by the varying nature of the broadcasts depending on the sport modality. Given the preliminary nature of the present study, which is aimed at dealing with source texts produced with no controlled language constraints, the sports selected—basketball (S1), cycling (S2) and tennis (S3)—

aries (Martínez-Martínez & Vela, 2016; Ortiz-Boix & Matamala, 2016; Petukhova et al., 2012) being one of the most often employed genres, together with films and TV series. At this point, it is worth noting the variety of genres studied by the SUMAT project (Petukhova et al., 2012), where news programmes were also present, though no data were collected in the SP-EN combination.

² Length and conditions in accordance with recent studies on the same topic [Koponen, Sulubacak, Vitikainen & Tiedemann (2020); Karakanta, Bentivogli, Cettolo, Negri & Turchi (2022)]. Given its pilot nature, only three genres were selected for this preliminary phase of the study.

204 tried to represent different broadcast styles with different language densities and scene configuration. Besides these features, the most significant and challenging aspect in the nature of these AV texts is the disruptive presentation of text content. Given that in all cases—as is the case in sports competitions—text contents are not scripted and so are filled with hesitations, disruptions and other features of oral discourse, fluency, coherence and cohesion are but some of the main limitations that will be present at later steps of the study.

As this was a pilot research study, a simplified layout was selected, avoiding time codes or other symbols that would hinder or distract professionals, such as the subtitle number or TCRs.

Once the clips were selected and trying to replicate the nature of many AVT processes and projects in the market, video transcripts were automatically generated using the automatic speech recognition system of Microsoft 365. At this step, again, as indicated by previous research by Adrià Martín-Mor & Pilar Sánchez-Gijón's (2016) and Laura Mejías-Climent & Julio de los Reyes Lozano (2021), controversy arose when it came to generating the transcript translators would be provided with. On this question, Martín-Mor & Sánchez-Gijón (2016, p. 172) mentioned technical issues that were considered an important obstacle, such as how to specifically transcribe audio for MT, how to deal with the interaction between different characters in a scene, the noise and ambient sound affecting automatic sound recognition or human transcription that could compromise the accuracy of the transcription. In our case, this was the main difficulty due to the ambience sounds that accompanied the speaker's intervention—especially in the case of sports—which also showed the lack of fluency

that characterises this type of speech, filled with interruptions, unfinished sentences and hesitation (Varga, 2021).

After the transcripts had been generated, three different MT engines were used to obtain different translation versions of each clip. The MT systems used in this process were Google Translate, AppTek and DeepL. Here again, trying to reproduce the conditions professional subtitlers work under in the market, the election of the engines was based on their presence and connection to some of the most popular AVT software (See section 2.1). Here, comparing the error rate provided by the three different engines, the versions provided by DeepL were the ones selected due to their lower error rate for this part of the study.

In order to evaluate the quality of the results produced by the three MT engines, we used the MQM metric mentioned in section 2.2.1. and widely detailed in previous literature (Lommel et al., 2014).

In no case did the authors of the study find accuracy problems that caused the source text not to be reflected correctly, cases of addition, deletion or no translation. If there were mistranslations in which the target text did not partially match the source text, these were due to errors already detected in the transcriptions that the professionals were able to solve in their post-editing task.

The three sports texts showed a constant set of problems in Google and AppTek, mainly of fluency: use of a grammatical register and a linguistic variant different from that of peninsular Spanish (errors in pronouns and verb forms), basic spelling errors, incoherence in punctuation, style and terminology, none of which are new and have already been mentioned by Evgeny Matusov et al. (2019). Google presented a lower number of terminology

errors, while DeepL showed some sporadic errors in punctuation and terminology. Of the three sports videos, all three engines presented more fluency errors in the case of S3, which was perhaps more challenging as it contained a greater number of choppy sentences, lack of fluency in the source text and a more specialised terminology.

As for the three news texts, following the steps adopted in the case of sports videos, the examples selected showed a single speaker, addressed international news and were taken from international news channels such as Al Jazeera, Euronews and CBS. In contrast to the results obtained for sports, all three engines delivered results of a much higher quality in the translation of news texts. AppTek once again presented grammatical errors (syntax errors, calques of the grammatical structure of the original text, spelling, punctuation, terminology), while Google maintained calques of the original and DeepL maintained style problems at times.

Regarding the number of words generated by the MT with respect to the original text, there is a big difference between the audiovisual genres: in the case of one of the news texts (N1) there was no significant increase in the number of words, while in the other two texts, it rose to 18%. In the case of sports, the increase was significantly smaller: 5.2% (S1), 4.8% (S3) and, in the case of S2, it decreased by 2.8%. In all cases, the order in terms of the highest number of words returned was AppTek, Google and DeepL.

Once the MT version—DeepL—had been selected, flat texts both for transcripts and MT versions were used with a simple division in word format. This layout makes it possible to focus on the linguistic transfer, leaving aside the technical issues and restrictions for further research.

Additionally, an online questionnaire with 21 questions was designed to compile information

on the background of the professionals taking part in the pilot study. The questionnaire, divided into different sections, was designed as a qualitative tool to ask participants about their education, profile, training in AVT and MTPE, the role of MTPE in their professional practice and on their personal approach.

4.1.2. *Participants*

18 translators were recruited for the pilot study and were asked to complete a two-step task. Different profiles, with over five years' experience, were represented, with six members of academia (AVT trainers), four professional translators and eight mixed profiles (i.e. professional translators collaborating with academia or with different profiles in the language industries—voice talents, proofreaders, etc.). In line with Celia Rico & Diana González-Pastor (2022), all the different profiles were integrated in the pilot study, from professional translators to AVT trainers.

The total number of participants was established according to the possible combinations of videos to be translated (3) / post-edited (3) per genre.

4.1.3. *Methods*

First, all the translators were asked to fill in the online questionnaire with general questions on their education, training and the role of MTPE in their professional practice. They were then provided with the material to complete two tasks: an EN-SP translation for one of the genres (sports / news) and a post-editing task for the other genre, in which participants were asked to fully post-edit the text, i.e. to reach a level of text editing with a human quality standard (high quality MT). Although professional assignments

206 traditionally include the adjustments required by the AV modality, as is the case of subtitling, the study focused, in this preliminary stage, on the implications of MTPE on the translation process, leaving the implications on the specific tasks demanded by the different AV modalities for further research.

As the participants were asked to complete their translation task in their own professional settings, they were also asked to report on the time invested in the completion of each task. No further documents and/or indications were provided, and translators were not intentionally informed on the AV modality they were working on.

4.2. Results

Having presented the set-up and given the large amount of data collected in the pilot study, from quality metrics for translation and post-editing tasks to online questionnaires, among others, this section presents the results in terms of three of the main elements that were measured and analysed during the tasks that participants completed during the pilot study: time, number of words and opinion.

4.2.1. Time

Results show interesting differences in the time invested to complete the tasks participants were assigned. On average, for the sports videos they spent 47 minutes to translate the texts and only 44 minutes to post-edit the texts. However, when it comes to news videos, the time spent to translate was on average 59 minutes, compared to 36 minutes spent on the post-edition. That makes 23 minutes less.

As previously indicated, some genres like sports broadcasts and, in non-technical aspects, news programmes are rarely addressed in AVT research. Therefore, while acknowledging the significant differences existing among the video lengths and text densities, it is worth noting the significant decrease in the time spent in post-editing the news videos, with more representative variations between the two tasks commissioned. [See Figure 1].

On average, translators invested 15% less time in the post-editing process of sports videos than in translating, whereas the difference rose to 36% in the same task for news videos. In this regard, it is not surprising that professional translators

FIGURE 1. Time invested per sport video (S) and news report (N) in translation tasks (W) and post-editing tasks (O).

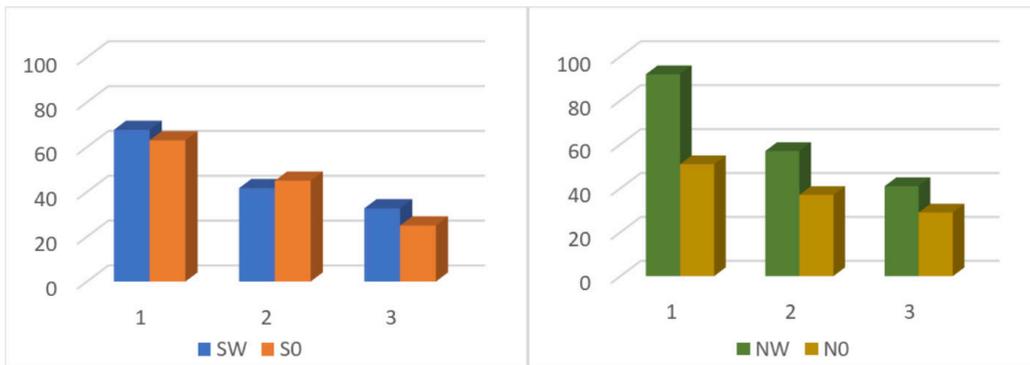
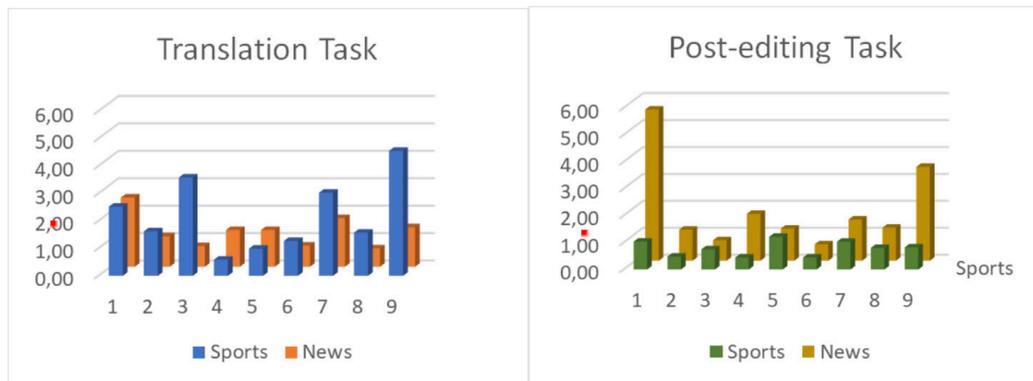


FIGURE 2. Time (seconds / character) invested per participant, video and task assigned.



invested the lowest times in both tasks: translation and post-editing. If the average time spent on the translation task was 1.68 seconds / character, most professional translators worked below 1.5 seconds / character. Similarly, for the post-editing task, the average time invested plummeted to 1.31 seconds / character, with professional translators working far below 1.0.

Although results of the preliminary research already reported on the time reductions when comparing translation and post-editing processes, the specificities of the text genres and professional profiles in the present pilot study provide interesting data that, once again, give ground for further research.

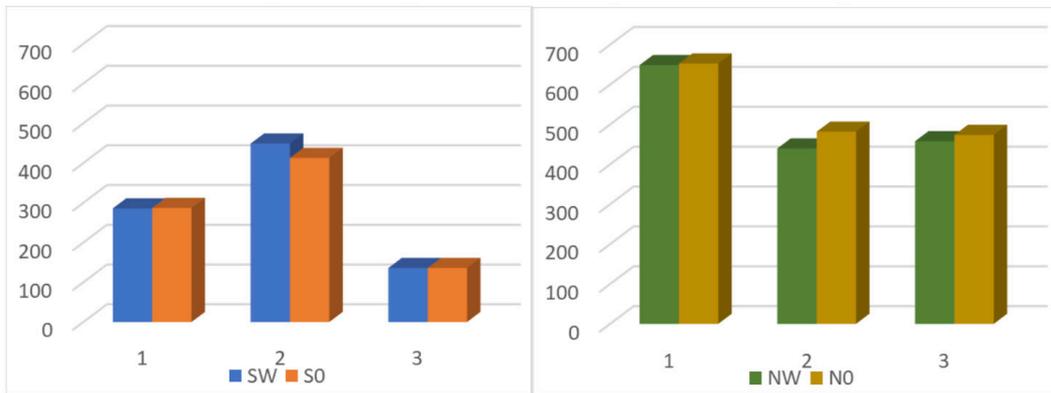
4.2.2. Words

In parallel to the time invested in both tasks and considering that no additional information had been provided regarding the translation modality to be completed, it seemed interesting to analyse the number of words used in each process. Considering the differences underlying the nature of the video materials—the diverse nature of sports and sports broadcasts, in

language density and speech rates, to name but a few—and in contrast to the important variations observed in the analysis of the times invested by the participants, only subtle differences were perceived between the translation and post-editing results for both genres. As expected by the different average word lengths between English (4.5-4.7) and Spanish (5.0-5.4) and the longer extension of Spanish translated texts, most translators provided longer Spanish versions in terms of word count: 10/18 for translations and 11/18 for post-edited versions. With respect to genres, longer texts corresponded to news texts—15 out of the 18 videos—as opposed to only 6 sports videos.

The word increase observed in the final texts brings results which provide a basis for further research. Far from approaching the traditional 20% extra that is assumed for translated texts in the EN-SP combination (Eriksen, 2019; Cantos & Sánchez, 2011; Quesada-Granja, 2009), for the texts in this pilot study, in the case of sports, the final texts are in both cases shorter—1% in the case of translated texts and 4% in the case of post-edited texts. In the case of news, however, the final texts are longer than the

FIGURE 3. Average number of words per video and task assigned.



original versions, with differences between the translated versions (5% longer) and the post-edited versions (9.5% longer). But they are nevertheless also far from the traditional 20%.

4.2.3. Opinion

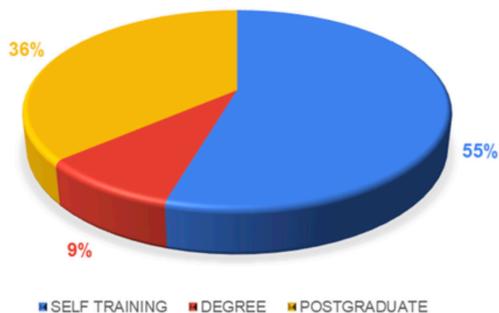
Although last in the presentation of our research results, it is interesting and necessary to focus on the answers provided by the professionals who took part in the pilot study, especially in those aspects related to their perception of and experience with MTPE tools.

As already mentioned in previous sections, the participants in the pilot study represented the three most significant profiles in the AVT professional market—professionals, trainers and mixed profiles that combine both occupations—and all of them completed the online questionnaire previously mentioned.

Even though almost all the participants expressed their opinion on the use of these tools, only 83% of the professional translators and those with a mixed profile had previous experience with MTPE.

Not surprisingly, 16% of the participants—all of them professionals—reject the use of MTPE tools as these impoverish their work and worsen the quality of AVT products, according to their answers, much in line with the results presented in 2023 Elis report (ELIS, 2023, p. 22). However, most participants (66%) consider that these tools increase productivity and/or make their work easier. Nevertheless, although the majority of participants seem to support the boost to productivity with the integration of MTPE, 44%

FIGURE 4. Training modality in MTPE.



also admit “professional practice is pauperised and the quality of the final result reflects on it”.

It is worth noting here that 70% of the participants with previous experience with MTPE acknowledged having self-trained, with no additional training received in 55% of cases, and 36% completed postgraduate training.

In this respect, it is significant that 75% of professional translators admit that they have integrated these tools into their professional practice on a voluntary basis—50% report having self-trained—as opposed to 25% who have been forced to implement their use by the companies for which they provide AVT services.

5. CONCLUSIONS

Despite their controversial nature, it is undeniable that the presence of machine translation and post-editing, not only in the translation market but also in society, is here to stay. Being aware of the significant opposition of a large part of the professional community to the use of these tools, their increasingly widespread adoption, not only by large multinational companies but also by professionals seeking to increase their own productivity, requires us to consider their progressive future imposition in the medium to long term.

Once again, academic institutions and, in particular, those responsible for the most traditional official training programmes—mainly bachelor’s and master’s degrees—are always working to adapt their curricula to the changing nature of the T&I market. They are aware of the need to update and adapt translators-to-be to the needs and demands of their future reality, but all too often initiatives still lag behind and have not yet succeeded in filling this gap. Technology is no longer perceived as an additional competence but

as a core transversal must in the training process. Tools for speech recognition, editing and proofreading, terminology management, Desktop Publishing, diverse CAT Tools and, of course, MTPE have come to be essential instruments for professional practice. In this sense, the increasingly frequent presence of these tools in the curricula and training of translators-to-be, graduates and professionals in search of reskilling in the use of all sorts of technologies simply reflects the reality of the professional market. The multiplicity of tools and their continuous technological improvement means that companies not only impose their use in order to increase quality and productivity—which do not always go together—but professionals themselves are progressively adopting them on a voluntary basis in order to increase their workload, while at the same time acknowledging that they are somehow sacrificing the quality of their production.

This is evidenced by the data collected in the study presented here, where we aimed at comparing translators’ productivity when translating and/or post-editing a set of sports and news texts. We found a considerable reduction in working time in post-editing tasks: from 15% for sports broadcasts—with unscripted dialogues and not as frequently processed from the point of view of personal experience and professional practice—to 36% for news programmes, with a high language density—and speech rate—due to their generally scripted nature and their more extended presence in personal and professional spheres.

This time reduction, as might be expected, shows significant differences in the case of professional translators: compared to the 1.68 seconds / character (2.09 sports vs. 1.26 news) invested on average in the translation task by all the participants—professional translators,

210 trainers and mixed profiles—professional translators drew an average result of 1.29. In the case of post-editing tasks, with results reflecting a higher speed for all the participants, 1.31 on average, professional translators spent half the time to complete that same task.

As regards the volume of the texts produced, due, among other reasons, to the different average word length in English and Spanish, traditional practice holds that texts translated into Spanish tend to be 20% longer than their original versions. The results of this study, however, show that, although the majority of translated and post-edited texts are longer than their original versions—10 out of 18 in the case of translations and 11 out of 18 in the case of post-edited texts—the final word increase is far from the percentages that have traditionally been mentioned, with SP texts between 1% and 4% shorter in the case of sports videos and between 5% and 9.5% longer in SP news. In this regard, and due to the significant differences between the genres selected for the study, it is necessary to conduct further research to analyse in detail the origin of these differences. In the case of sports texts, the non-perfect nature of the source texts, with constant interruptions, rephrasing, truncated sentences and excessive orality could partially explain this fact, as both MT tools and the transcription tools used draw more errors in their processing. In any case, it would be necessary to analyse separately whether the mode of production—translation or MTPE—beyond the quality of the text produced also conditions the style and characteristics of the product.

Similarly, and in the light of the answers provided by the participants in the study, despite the fact that a large majority consider that MTPE is a tool that boosts translators' productivity, up to 44% of them also believe that it impoverishes the translator's work and

neglects the quality of end products. However, although the controversy is considerable and many professionals reject the use of these processes in AVT, it is also worth noting the gradual and voluntary incorporation of this tool by many professionals into their daily practice.

In any event, it is evident that the wide variety of materials—texts—addressed by AVT is so vast that it seems difficult to consider a homogeneous and standardised scenario for a generalised use for the application, or not, of MTPE or other tools, such as speech recognition, also used in this study. Similarly, the changing situation posed by the technological evolution on which they depend makes the situation ephemeral, placing professionals in a constant search for reskilling possibilities—self-training for most—in order to increase their productivity. Quality, often questioned, may not always be the first reason.

REFERENCES

- Agulló García, B. (2020). El paradigma actual de la subtitulación: cambios en la distribución de contenido, nuevos hábitos de consumo y avances tecnológicos. *La linterna del traductor*, 20, 53-62.
- Alves, F., Sarto Szpak, K., Gonçalves, J. L., Sekino, K., Aquino, M., Araújo e Castro, R., Koglin, A., de Lima Fonseca, N. B. & Mesa-Lao, B. (2016). Investigating cognitive effort in postediting: A relevance-theoretical approach. In S. Hansen-Schirra & S. Gruzca (Eds.), *Eyetracking and Applied Linguistics* (pp. 109-142). Language Science Press. <http://10.17169/langsci.b108.296>
- AppTek (2021, September 20). This Is How Automatic Speech Recognition & Machine Translation Are Revolutionizing Subtitling. *Slator*. <https://slator.com/automatic-speech-recognition-machine-translation-are-revolutionizing-subtitling/>
- Aranberri, N. (2014). Posedición, productividad y calidad. *Revista Tradumàtica. Tecnologies de la Traducció*, 12, 471-477. <https://doi.org/10.5565/rev/tradumatica.62>

- Asociación Española de Traductores, Correctores e Intérpretes (ASETRAD) (2021). *Posedición. Guía para profesionales*. Asetrad.
- ATAA (2022, October 3). Machine-translated subtitling – human translators are up in arms. *Association des traducteurs adaptateurs de l'audiovisuel*. <https://beta.ataa.fr/blog/article/machine-translated-subtitling-human-translators-are-up-in-arms>
- Athanasiadi, R. (2021). Technology as a driving force in subtitling. In L. Mejías-Climent & J. F. Carroero Martín (Eds.), *New Perspectives in Audiovisual Translation. Towards Future Research Trends* (pp. 139-166). Universitat de València.
- ATRAE (2021, October 13). Comunicado sobre la posesición. *Asociación de Traducción y Adaptación Audiovisual de España*. <https://atrae.org/comunicado-sobre-la-posesicion/>
- ATRAE (2022, September 27). Comunicado sobre el uso de lenguas puente. *Asociación de Traducción y Adaptación Audiovisual de España*. <https://atrae.org/comunicado-sobre-el-uso-de-lenguas-puente/>
- AVTE (2021). Machine translation manifesto. *Audiovisual Translators Europe. The European Federation of Audiovisual Translators*. <https://avteurope.eu/avte-machine-translation-manifesto/>
- Bellés-Calvera, L. & Caro Quintana, R. (2021). Audiovisual Translation through NMT and Subtitling in the Netflix Series *Cable Girls*. *Translation and Interpreting Technology Online*, 142-148. https://doi.org/10.26615/978-954-452-071-7_015
- Bolaños-García-Escribano, A., Díaz-Cintas, J. & Massidda, S. (2021). Subtitlers on the Cloud: The Use of Professional Web-based Systems in Subtitling Practice and Training. *Revista Tradumàtica. Tecnologies de la Traducció*, 19, 1-21. <https://doi.org/10.5565/rev/tradumatica.276>
- Bruno, L., Miloro, A., Sabaté-Carrové, P. E. & Sabaté-Carrové, M. (2021). Preconceptions from pre-professionals about MTPE. In M. Sabaté-Carrové & L. Baudo (Eds.), *The translation process series. Multiple perspective from teaching to professional practice*. Volume 1 (pp. 41-44). Edicions de la Universitat de Lleida.
- Bundgaard, K. (2017). Translator attitudes towards translator-computer interaction - Findings from a workplace study. *Hermes-Journal of Language and Communication in Business*, 56, 125-144. <https://doi.org/10.7146/hjlc.v0i56.97228>
- Burchardt, A., Lommel, A., Bywood, L., Harris, K. & Popović, M. (2016). Machine translation quality in an audiovisual context. *Target* 28(2), 206-221.
- Bywood, L. & Georgakopoulou, Y. (2017). Embracing the threat: machine translation as a solution for subtitling. *Perspectives Studies in Translatology* 25(3), 1-17. <http://10.1080/0907676X.2017.1291695>
- Cadwell, P., O'Brien, S. & Teixeira, C. S. C. (2017). Resistance and Accommodation: Factors for the (Non)adoption of machine translation among professional translators. *Perspectives*, 26(3), 301-321. <https://doi.org/10.1080/0907676X.2017.1337210>
- Cantos, P. & Sánchez, A. (2011). El inglés y el español desde una perspectiva cuantitativa y distributiva: equivalencias y contrastes. *Estudios Ingleses de la Universidad Complutense*, 19, 15-44.
- Carl, M., Dragsted, B., Elming, J., Hardt, D. & Jakobsen, A. L. (2011). The process of post-editing: a pilot study. In B. Sharp, M. Zock, M. Carl & A. L. Jakobsen (Eds.), *Proceedings of the 8th international NLPSC workshop. Special theme: Human-machine interaction in translation* (pp. 131-142). Copenhagen Business School.
- Díaz-Cintas, J. & S. Massidda (2020). Technological advances in audiovisual translation. In M. O'Hagan (Ed.), *The Routledge Handbook of Translation and Technology* (pp. 225-270). Routledge.
- European Language Industry Survey (ELIS) (2023). Trends, expectations and concerns of the European language industry. <https://elis-survey.org/wp-content/uploads/2023/03/ELIS-2023-report.pdf>
- Eriksen (2019, March 6). Text Expansion and contraction in Translation. *Eriksen Translation*. <https://eriksen.com/language/text-expansion/>
- Estopace, E. (2018, January 22). From Indonesia to Nigeria: Inside iflix's Localization Rollout. *Slator*. <https://slator.com/indonesia-nigeria-inside-iflixs-localization-rollout/>
- Fernández Moriano, P. (2019). Herramientas informáticas para traductores audiovisuales. In J. F. Carrero Martín, B. Cerezo Merchán, J. J. Martínez Sierra & G. Zaragoza Ninet (Eds.), *La traducción*

- audiovisual: aproximaciones desde la academia y la industria* (pp. 73-92). Comares.
- Fernández-Torné, A. & Matamala, A. (2015). Text-to-Speech vs. Human Voiced Audio Descriptions: A Reception Study in Films Dubbed into Catalan. *The Journal of Specialised Translation*, 24, 61-88.
- Firat, G. (2021). Uberization of translation. Impacts on working conditions. *The Journal of Internationalisation and Localization*, 8(1), 48-75. <https://doi.org/10.1075/jial.20006.fir>
- Gaspari, F., Almaghout, F. & Doherty, S. (2015). A survey of machine translation competences: Insights for translation technology educators and practitioners. *Perspectives*, 23(3), 333-358.
- Georgakopoulou, P. (2006). Subtitling and globalisation. *The Journal of Specialised Translation*, 6, 115-120.
- Georgakopoulou, P. (2019). Template files: The Holy Grail of subtitling. *Journal of Audiovisual Translation*, 2(2), 137-160.
- González-Pastor, D. & Rico, C. (2021). POSEditTrad: la traducción automática y la posesición para la formación de traductores e intérpretes. *Revista Digital de Investigación en Docencia Universitaria (RIDU)*, 15(1). <http://dx.doi.org/10.19083/ridu.2021.1213>
- Guerberof-Arenas, A. (2009). Productivity and quality in MT post-editing. In *Proceedings of the MT Summit XII Workshop Beyond Translation Memories: New Tools for Translators MT*.
- Guerberof-Arenas, A. (2013). What do professional translators think about post-editing. *The Journal of Specialised Translation*, 19(19), 75-95.
- Karakanta, A. (2022). Experimental research in automatic subtitling: At the crossroads between Machine Translation and Audiovisual Translation. In G. M. Greco, A. Jankowska & A. Szarkowska (Eds.), *Methodological Issues in Experimental Research in Audiovisual Translation and Media Accessibility*, *Translation Spaces 11*(1), 89-112. <https://doi.org/10.1075/ts.21021.kar>
- Karakanta, A., Bentivogli, L., Cettolo, M., Negri, M. & Turchi, M. (2022). Post-editing in Automatic Subtitling: A Subtitlers' perspective. In *Proceedings of the 23rd Annual Conference of the European Association for Machine Translation* (pp. 261-270). European Association for Machine Translation. <https://aclanthology.org/2022.eamt-1.29.pdf>
- Kenny, D. & Doherty, S. (2014). Statistical machine translation in the translation curriculum: overcoming obstacles and empowering translators. *The Interpreter and Translator Trainer*, 8(2), 276-294.
- Kenny, D. (2022). Machine translation for everyone: Empowering users in the age of artificial intelligence. *Language Science Press*. <http://10.5281/zenodo.6653406>
- Koponen, M., Sulubacak, U., Vitikainen, K. & Tiedemann, J. (2020). MT for subtitling: Investigating professional translator's user experience and feedback. In *Proceedings of 1st Workshop on Post-Editing in Modern-Day Translation*. Virtual. Association for Machine Translation in the Americas (pp. 79-92).
- Krings, H. P. (2011). *Repairing texts: empirical investigations of machine translation post-editing processes*. The Kent University Press.
- Läubli, S., Fishel, M., Massey, G., Ehrensberger-Dow, M. & Volk, M. (2013). Assessing Post-Editing Efficiency in a Realistic Translation Environment. In S. O'Brien, M. Simard & L. Specia (Eds.), *Proceedings of MT Summit XIV Workshop on Post-editing Technology and Practice* (pp. 83-91). European Association for Machine Translation. <https://doi.org/10.5167/uzh-80891>
- Lionbridge (2022, August 23). Lionbridge Language Cloud: Is Machine Translation getting better? *The Lionbridge Team*. <https://www.lionbridge.com/blog/translation-localization/language-cloud-is-machine-translation-getting-better/>
- Lommel, A., Uszkoreit, H. & Burchardt, A. (2014). Multidimensional Quality Metrics (MQM): A Framework for Declaring and Describing Translation Quality Metrics. *Revista Tradumàtica. Tecnologies de la Traducció*, 12, 455-463.
- Martín-Mor, A. & Sánchez-Gijón, P. (2016). Machine translation and audiovisual products: a case study. *The Journal of Specialised Translation*, 26, 172-186.
- Martínez Martínez, J. M. & Vela, M. (2016). SubCo: A Learner Translation Corpus of Human and Machine Subtitles. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)* (pp. 2246-2254). European Language Resources Association (ELRA). <https://aclanthology.org/L16-1357.pdf>

- Matusov, E., Wilken, P., & Georgakopoulou, Y. (2019). Customizing neural machine translation for subtitling. In *Proceedings of the Fourth Conference on Machine Translation (WMT)*, Volume 1 Research Papers (pp. 82-93). Association for Computational Linguistics.
- Mejías-Climent, L. & de los Reyes Lozano, J. (2021). Traducción automática y posesición en el aula de doblaje: resultados de una experiencia docente. *Hikma*, 20(2), 203-227. <https://doi.org/10.21071/hikma.v20i2>
- Moorkens, J., Toral, A., Castilho, S., & Way, A. (2018). Translators' perceptions of literary post-editing using statistical and neural machine translation. *Translation Spaces*, 7(2), 240-262. <https://doi.org/10.1075/ts.18014.moo>
- Nunes Vieira, L. & Alonso, E. (2018). *The use of machine translation in human translation workflows: Practices, perceptions and knowledge exchange*. Technical report. University of Bristol and Universidad Pablo de Olavide.
- O'Curran, E. (2014). Translation quality in post-edited versus human-translated segments: a case study. In S. O'Brien, M. Simard & L. Specia (Eds.), *Proceedings of AMTA Third Workshop on Post-editing technology and Practice* (pp. 113-118). Association for Machine Translation in the Americas.
- Ortiz-Boix, C. (2016). Machine translation and post-editing in wildlife documentaries: Challenges and possible solutions. *Hermeneus: Revista de la Facultad de Traducción e Interpretación de Soria*, 18, 269-313.
- Parra Escartín, C. (2018). ¿Cómo ha evolucionado la traducción automática en los últimos años? *La linterna del traductor*, 16, 20-28.
- Pedersen, J. (2017). The FAR model: assessing quality in interlingual subtitling. *The Journal of Specialised Translation*, 28, 210-229.
- Petukhova, V., Agerri, R., Fishel, M., Georgakopoulou, Y., Penkale, S., del Pozo, A., Sepesy M., Volk, M. & Way, A. (2012). SUMAT: Data Collection and Parallel Corpus Compilation for Machine Translation of Subtitles. In *Proceedings of the 8th Language Resources and Evaluation Conference* (pp. 21-28).
- Quesada-Granja, C. (2009, January 19). Distribución por longitud de las palabras de diferentes idiomas. *Slideshare*. <http://www.slideshare.net/quesadagranja/distribucion-por-longitud-de-las-palabras-de-diferentes-idiomas-presentation>
- Rico, C. & González Pastor, D. (2022). The role of machine translation in translation education: A thematic analysis of translator educators' beliefs. *The international Journal of Translation & Interpreting*, 14(1), 177-197.
- Rinsche, A. & Portera-Zanotti, N. (2009). Study report to the Directorate General for Translation of the European Commission. Final Version. Study on the Size of the Language Industry in the EU. The Language Technology Centre Ltd.
- Romero-Fresco, P. & Martínez, J. (2015). Accuracy Rate in Live Subtitling: The NER model. In R. B. Piñero & J. Díaz-Cintas (Eds.), *Audiovisual Translation in a Global Context. Palgrave Studies in Translating and Interpreting* (pp. 28-50). Palgrave Macmillan. https://doi.org/10.1057/9781137552891_3
- Romero-Fresco, P. & Pöchhacker, F. (2017). Quality assessment in interlingual live subtitling: the NTR model. *Linguistica Antverpiensia, New Series: Themes in Translation Studies*, 16, 149-167. <https://doi.org/10.52034/lanstts.v16i0>
- Sagar, R. (2020, June 5). Netflix Is Using AI For Its Subtitles. *Aim*. <https://analyticsindiamag.com/netflix-ai-subtitles-translation/>
- Sakamoto, A. (2021). The value of Translation in the Era of Automation: An examination of Threats. In R. Desjardins, C. Larsonneur & P. Lacour (Eds.), *When Translation Goes Digital* (pp. 231-256). Palgrave Studies in Translating and Interpreting MacMillan.
- Sánchez-Gijón, P. (2016). La posesición: hacia una definición competencial del perfil y una descripción multidimensional del fenómeno. *Sendebarr*, 27, 151-162. <https://doi.org/10.30827/sendebarr.v27i0.4016>
- Szarkowska, A., Díaz-Cintas, J. & Gerber-Morón, O. (2021). Quality is in the eye of the stakeholders: what do professional subtitlers and viewers think about subtitling? *Universal Access in the Information Society*, 20, 661-675. <https://doi.org/10.1007/s10209-020-00739-2>
- Translation Automation User Society (TAUS) (2010). *Machine Translation post-editing guidelines*. De Rijp.

- 214 Torres-Hostench, O., Presas, M. & Cid-Leal, P. (co-ords.) (2016). El uso de traducción automática y posesición en las empresas de servicios lingüísticos españolas: Informe de investigación ProjectA 2015. <https://ddd.uab.cat/record/148361>
- Varga, C. (2021). Online Automatic Subtitling Platforms and Machine Translation. An Analysis of Quality in AVT. *Scientific Bulletin of the Politechnica University of Timișoara*, 20(1), 37-49.
- Vercauteren, G., Revers, N. & Steyaert, K. (2021). Evaluating the effectiveness of machine translation of audiodescription: the results of two pilot studies in the English-Dutch language pair. *Revista Tradumàtica. Tecnología de la Traducción*, 19, 226-252.