Universitat Autònoma de Barcelona

Departament de Traducció i d'Interpretació

i d'Estudis de l'Àsia Oriental

Doctorat en Traducció i Estudis Interculturals

Audio description and technologies

Study on the semi-automatisation of the translation

and voicing of audio descriptions

PhD thesis submitted by:

Anna Fernández-Torné

Supervised by:

Dr. Anna Matamala

2016

Annexes

Annex 1: Articles as published

Annex 1.1. 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.

The Journal of Specialised Translation Issue 24 – July 2015

Text-to-speech vs. human voiced audio descriptions: a reception study in films dubbed into Catalan

Anna Fernández-Torné, Universitat Autònoma de Barcelona
Anna Matamala, Universitat Autònoma de Barcelona

ABSTRACT

This article presents an experiment that aims to determine whether blind and visually impaired people would accept the implementation of text-to-speech in the audio description of dubbed feature films in the Catalan context. A user study was conducted with 67 blind and partially sighted people who assessed two synthetic voices when applied to audio description, as compared to two natural voices. All of the voices had been previously selected in a preliminary test. The analysis of the data (both quantitative and qualitative) concludes that most participants accept Catalan text-to-speech audio description as an alternative solution to the standard human-voiced audio description. However, natural voices obtain statistically higher scores than synthetic voices and are still the preferred solution.

KEYWORDS

Accessibility, audio description, audiovisual translation, text-to-speech, speech synthesis, Catalan language, blind, visually impaired.

1. Introduction

Accessibility has become a major concern in society in recent decades, and laws are being enforced to guarantee disabled people's rights. Present legislation states that sensorial accessibility to audiovisual content should be provided: theatres, museums, TV broadcasters and web designers, among others, are endeavouring to make their content accessible to persons with disabilities and to comply with regulations.

For users who are blind or visually impaired, audio description (AD) allows access to visual information (images) appearing on screen, which they would otherwise miss. Audio description can be defined as an inter-semiotic translation in which the visuals are transferred into words that are received aurally by end users (Orero 2007; Orero and Matamala 2007). In films these oral descriptions are inserted in the silent gaps in the dialogue, i.e. when characters are not talking, and create a coherent whole with the film dialogues and soundtrack (Braun 2011). However, because creating and voicing an audio description is a time-consuming and costly process, this access service is not as widely available as one might expect. This is especially striking in social media environments, but also in other traditional broadcasting contexts.

In view of the need for wider availability of audio described audiovisual products, research on technological processes which fully or partially automate the audio description workflow is considered relevant, from a

scientific, social and economic point of view. Within this general framework, this article aims to present the results of research in which Catalan audio description using text-to-speech (TTS) software was assessed, and compared to standard human-voiced audio descriptions. Our final aim was to find out whether TTS AD in Catalan would be accepted by blind and visually impaired patrons as an alternative solution and to compare the scores attributed to both natural and artificial voices on key aspects. The project's novelty lies in the language under analysis (Catalan). In addition, the methodological approach is also new in comparison with existing text-to-speech audio description tests: on the one hand, it provides a detailed analysis of many features instead of asking about general opinions or perceptions; secondly, it assesses text-to-speech audio description against human-voiced audio description instead of evaluating it in isolation, as further explained in section 3.2.

The article presents, first of all, a review of related work, focussing on text-to-speech audio description but also widening the scope to present other text-to-speech applications in audiovisual translation and media accessibility (section 2). Methodological aspects are detailed in section 3, and results are discussed in section 4. Conclusions and possibilities for further research close the article.

2. Text-to-speech audio description: an overview

Blind and visually impaired people use text-to-speech in many contexts, and its usefulness has already been proved in different domains. Cryer and Home (2008) analyse the use of synthetic speech technology by blind and partially sighted people. Inspired by Freitas and Kouroupetroglou (2008), they list the many areas in which speech technologies can be used: mobility aids (for instance, GPS navigation devices), educational tools (talking dictionaries, audio textbooks), entertainment (audio subtitles (AST), speaking electronic programming guides) and communication (screen reading software on computers). Speech synthesis seems to offer quicker access to information (Llisterri *et al.* 1993) and guarantees independence of the user (González García 2004), among other aspects. Cryer and Home (2008) point out two relevant research results of their overview of text-to-speech usage by blind and partially sighted people: firstly, the direct impact of each user experience on the acceptance of synthetic speech, as people gradually get used to synthetic voices, and, secondly, the impact of the naturalness and the context where the artificial voice is being used.

As for text-to-speech audio description (TTS AD), it has been researched within a project developed at the University of Warsaw, Poland, aiming to assess its feasibility and its reception among visually impaired people. Szarkowska (2011: 144) states that "instead of recording a human voice reading out the AD script, TTS AD can be read by speech synthesis

software". This guarantees the cost-effectiveness of the AD production in comparison with traditional methods of AD production.
The project analysed the application of TTS AD in several types of audiovisual products:

- in a monolingual feature film in Polish (Szarkowska 2011), where the artificial voice tested was Ewa (female voice), by Ivo Software;
- in a dubbed educational TV series for children (Walczak and Szarkowska 2010), where the artificial voice tested was Ewa (female voice), by Ivo Software;
- in a foreign fiction film, with voice-over (Szarkowska and Jankowska 2012), where the artificial voice tested was Krzysztof (male), by Loquendo;
- in a non-fiction film, with audio subtitling (Mączyńska 2011), where the artificial voices tested were Zosia (female voice) for the AD, and Krzysztof (male voice) for the AST, both by Loquendo;
- in a dubbed feature film (Drożdż-Kubik 2011), where the artificial voice tested was Ewa (female voice), by Ivo Software.

The number of participants ranged from 17 in Drożdż-Kubik (2011) to 76 in Walczak and Szarkowska (2010). The conclusions for each study were as follows: Szarkowska (2011) and Szarkowska and Jankowska (2012) stated that most respondents accepted TTS AD both as an interim solution and as a permanent option; Walczak and Szarkowska (2010) emphasised that most participants enjoyed the voice used in the test, and Mączyńska (2011) and Drożdż-Kubik (2011) explained that a majority of respondents found TTS AD acceptable, although it was not the preferred solution. Hence all five studies showed that most viewers accept TTS in AD.

On a similar note, and inspired by Chapdelaine and Gagnon's work (2009) on an accessible website platform for rendering different levels of audio description (as far as quantity and quality of AD is concerned) on demand, Kobayashi *et al.* (2009: 249) describe a "technique to use synthesized speech to add AD to online videos on any websites." The three steps of their project include determining whether or not synthesised voice can compete with real voices, designing a text-based format to describe the AD scripts, and developing authoring software. Step one is thoroughly explained in Kobayashi *et al.* (2010): 115 visually-impaired adult participants took part in an informal survey in Japan where three kinds of voice were tested (human, standard TTS, and prototype TTS). This first experiment was followed by an in-depth interview session with three participants. The study continued in the US, where 236 participants completed a survey, followed by an additional in-depth interview session with eight participants. A follow-up study with 24 participants closed the research. It included additional variables such as long vs short stimuli, expressive TTS technology vs standard TTS, expert vs novice descriptions, and standard vs extended descriptions. All in all, this broad study showed

that synthesised descriptions are generally accepted, especially for relatively short videos and informational content.

With their more experimental approach, Encelle *et al.* (2011) present an exploratory work on video accessibility for the blind and visually impaired with "audio enrichments composed of speech synthesis and earcons (i.e. nonverbal audio messages)" (123). Their study with 21 blind volunteers show that earcons associated with speech synthesis are useful for understanding set-related information, i.e. enriching videos with the use of earcons to complement speech synthesis helps convey visual information.

Moving from academia to industry, the firm Swiss TXT is already planning to offer audio description in which text-to-speech technologies are implemented (Caruso 2012). A web-based editor for transforming text into speech which can be used for audio description has also been developed by Mieskes and Martínez (2011). The editor contains features which allow the speaking rate and pitch to be set, as well as phonetic tuning functionalities. The described scenario would allow a user to upload an existing description or create a new one, upload the corresponding movie and synthesise the descriptions. Similarly, Oncins *et al.* (2013) have developed the Universal Accessibility System, a multi-language and multi-system mobile application to make live performing arts accessible. The system is designed to offer automatic AD through speech synthesis as well as other features (subtitling, spoken subtitles, an emergency pack, etc.).

Research on text-to-speech in audiovisual translation (AVT) goes beyond audio description and is especially relevant in a strongly related transfer mode: audio subtitling or spoken subtitles, where a synthetic voice is used to automatically read aloud the subtitles and make them accessible not only to blind and visually-impaired people, but also to people with reading difficulties. This service has been implemented in television broadcasts in countries such as the Netherlands (Verboom *et al.* 2002) and Sweden (De Jong 2006), where two digital boxes are needed to make it work. To expand the availability of spoken subtitles and avoid the need for a special decoder, a user-based device for reading aloud subtitles (Subpal) has been proposed by Nielsen and Bothe (2007), and a free and open-source tool has been developed by Derbring, Ljunglöf and Olsson (2012) within the SubTTS project (Derbring, Ljunglöf and Olsson 2010).

Finally, it is worth mentioning that, focusing exclusively on the language under analysis in this research, Alías, Iriondo and Socoró (2011) present the state of speech synthesis implementation in Catalonia which includes the most relevant companies, research centres and products relating to Catalan synthetic voice generation, and carry out field work to map the actual usage of text-to-speech in Catalan audiovisual media. In a specific

section of their article devoted to blind and visually impaired users, they point out that most of them think text-to-speech could be used in AD as long as more natural and expressive voices can be developed, although no specific quantitative data are given.

3. Methodological aspects: materials and method

This section describes the participants involved in the current experiment, the voices used, the film and clip selection process, the evaluation questionnaires drafted, the actual development of the test, and the statistical methods used.

3.1. Participants

Since it was "impossible to map 'the population' from which a random sample" was to be taken (Bryman 2012: 416), an a priori generic purpose sampling strategy was adopted. Such a strategy implied the establishment of certain criteria for selecting participants at the outset of the research. A total of 67 persons participated in the test (55% female, 45% male). The mean age was 52, with ages ranging from 21 years old to 85 years old. Thirty-three participants (49%) were 50 or younger, the others being older (51%). A more detailed distribution of the participants is shown in Table 1:

	Women	Men	Total
From 20 to 34	5	4	9 (13%)
From 35 to 44	5	7	12 (18%)
From 45 to 54	7	8	15 (22%)
From 55 to 64	9	6	15 (22%)
From 65 to 74	9	4	13 (19%)
More than 80	2	1	3 (4%)
TOTAL	37 (55%)	30 (45%)	67

Table 1. Participants' distribution based on sex and age.

The age range was not limited to account for the whole spectrum of the adult population to which AD is offered. Additionally, acceptance of synthetic voices is often linked to their usage, and limiting the age range to younger or older participants would have probably had an effect on the results.

Using the World Health Organisation's classification of visual impairments (2013), 51% of the participants described their disability as blindness, whereas 49% declared it to be low vision, with visual impairment being from birth in 30 cases (45%).

As far as the participants' educational background is concerned, 51% reached at least first degree university level (Bachelor's degree or equivalent), whilst 24% did not reach the first stage of secondary school. 46% reported being unemployed, whilst 13% declared to be employed in clerical posts.

3.2. Voice selection

It was decided that a total of four voices (a male and a female artificial and a male and a female natural voice) would be included in the experiment to avoid any gender bias. In order to select them, a pre-test with 20 participants was carried out, as described in Matamala, Fernández-Torné and Ortiz-Boix (2013). Ten synthetic voices (see Table 2) and ten natural voices (both professional and non-professional voice artists selected by the Catalan School of Dubbing ECAD) were assessed by the participants (see Fernández-Torné and Matamala 2013 for further details on the methodological aspects of the pre-test).

Female	Male
Laia by Acapela	Jan by FestCat
Anna by iSpeech	Teo by FestCat
Meritxell by Verbio	Oriol by Verbio
Montserrat by Loquendo	Jordi by Loquendo
Ona by FestCAt	Pep by FestCat

Table 2. Artificial voices.

This pre-test allowed us to select the voices to be used in the experiment, namely a professional voice talent (a female natural voice), a non-professional but trained voice talent (a male natural voice), Laia by Acapela (a female synthetic voice), and Oriol by Verbio (a male synthetic voice).

3.3. Film and clip selection

The voices were tested in an audio described film excerpt. Various factors influenced the film selection process: first of all, this experiment is part of a wider project in which other technologies such as machine translation are to be tested in the English-Catalan language pair (Fernández-Torné, Matamala and Ortiz-Boix 2012). Therefore, a film which had already been audio described in Catalan (for the TTS AD experiments) and that had also been audio described in English (for the machine translation tests) was required. A dubbed fiction feature film or a children's animation film were the only options, as these were the only dubbed audiovisual products that were audio described in Catalan at the time the experiment took place.

Children's animation films were disregarded as our intended target audience were adults; hence a dubbed fiction film had to be selected.

Secondly, defining the specific genre was also considered relevant, since in TTS evaluation studies in other fields such as audiobooks, the text type has been shown to have a significant influence on the results. For instance, Hinterleitner *et al.* (2011) have proven that seven out of the 11 rating scales used in their study were influenced by the type of text when assessing the quality of the same synthetic voice. Our final decision was not to favour any particular film genre, and a film belonging to a "miscellaneous" category according to Salway *et al.*'s (2004) classification was chosen.

Finally, from a more practical point of view, it was considered that the availability of the English original script, the English AD script, the Catalan dubbed script, and the Catalan AD script would speed up the research process, and *Closer* (2004, directed by Mike Nichols) was selected. However, to limit the duration of the experiment, it was decided to carry out the experiment using short clips rather than the whole film, unlike the five studies within the TTS project developed at the University of Warsaw and the Jagiellonian University of Krakow (Szarkowska and Jankowska 2012).

As far as the clip selection was concerned, it was decided that two different clips, one clip for female voices and another one for male voices, would be chosen to minimis e fatigue and the impact of a learning effect on the subjects. Additionally, an in-depth analysis of the film, of the AD script and of the individual AD units was performed, in order to obtain two comparable clips in terms of content (neutral in both cases, with no potentially distracting and/or offensive content), length (3 minutes in clip 1 vs 3 minutes and 6 seconds in clip 2), intervening characters (Anna and Dan in both clips), background music (the same opera for both), and AD density (571 characters vs 537 characters respectively). Clips were randomly assigned a voice gender, either masculine or feminine, for the audio description.

3.4. Evaluation questionnaires

For the human assessment of synthetic voices, the Telecommunication Standardization Sector of the International Telecommunication Union (ITU-T) recommends using a Mean Opinion Score (MOS) test, by which listeners are asked to rate several systems taking into account various items (ITU Recommendation P.85 1994), hence this was our chosen approach. The items to be included in our questionnaire were selected after a thorough comparison of various tests in text-to-speech evaluation. These are:

- ITU Recommendation P.85 (1994), which includes seven 5-point scales and one 2-point (yes-no) scale;
- Viswanathan and Viswanathan (2005), who propose 11 items to be assessed on a 5-point scale;
- Cryer, Home and Morley Wilkins (2010), who suggest twelve 5-point scales; and
- Hinterleitner *et al.* (2011), who put forward an evaluation protocol for the assessment of TTS in audiobook reading tasks, concluding that eight scales out of the eleven they tested should be kept, with a continuous 7-point rating scale.

It was finally decided to limit the number of items and to focus on issues directly linked to end-user reception rather than on the intelligibility dimension, since intelligibility was taken for granted in the selected voices and was deemed more relevant for system performance testing. The final list of items included in our questionnaire is listed next, in the same order as they were presented to participants when given the instructions. Participants assessed each item on a 5-point scale.

Overall impression: a global score, the general opinion participants have of the voice of the audio description.

Accentuation: this score assesses whether the stress is put on the right syllable.

Pronunciation: measures to what extent words are correctly uttered according to Catalan phonetics.

Speech pauses: evaluates whether the voice stops when needed between sentence components and between sentences.

Intonation: assesses whether the pitch curve accurately represents the sentence type (whether it is a question, an exclamation or a declarative sentence).

Naturalness: in synthetic voices, this item assesses to which extent the voice resembles a human voice; in natural voices, it is related to the degree the human voice is forced and dramatised.

Pleasantness: conveys to what extent the listener finds the voice pleasant.

Listening effort: involves subjectively assessing whether listening to the voice for a long period of time would be tiring or tedious.

Acceptance: is used to indicate whether the voice is deemed adequate to voice audio descriptions.

It must be stressed that a careful translation into Catalan of each of the previous items, validated by a professional translator and tested in a pilot test, was carried out. It was also decided that heading descriptors were not to be used in the real test since the choice of an oral delivery mode for the test instead of a written one made the use of headings before a question quite awkward and it actually did not enhance comprehension. Therefore, participants were directly asked the questions and read aloud the 5 possible answers to each question preceded by their corresponding score: from least positive (1) to most positive (5) (see Annex 1 for the back translation into English of the actual Catalan questionnaire).

Regarding the order of the items, the overall impression and acceptance items were kept in the first and last positions respectively following the other tests. A logical order was proposed for the remaining scales, from more specific questions to broader ones: word-centered questions (accentuation and pronunciation), phrase-centered questions (speech pauses and intonation), voice-centered questions (naturalness and pleasantness) and a global question (listening effort).

As well as the questionnaire, a post-questionnaire was included, inspired by the works of Walczak (2010), Mączyńska (2011), Chmiel and Mazur (2012) and Pazos (2012). Its aim was to gather information on the participant demographics and to get more subjective information on personal preferences and usage of audio described audiovisual products and TTS applications in devices and/or computers. As in previous studies in the field (Walczak 2010, Mączyńska 2011), such questions were included in a post-questionnaire rather than in a pre-questionnaire. This decision was motivated by our wish to be as tactful as possible, trying not to ask potentially sensitive questions at the beginning of the test. The post-questionnaire, translated from Catalan into English, is included in Annex 2.

3.5. Procedure

Participants did the experiment on a one to one basis in a sound proof booth, following approved ethical procedures. Listening conditions were controlled: the stimuli were played with VLC Media Player and presented through professional headphones, Beats mixr by Dr. Dre. All participants were volunteers and listened to all stimuli, following a within-subjects design.

The experimental session was initially tested in a pilot test which was developed as follows: participants were given an overview of the project and the actual experiment, and were required to sign a Participant Information Sheet and Consent Form previously approved by the University Ethical Committee. They were instructed to assess each AD voice independently, and a thorough explanation of the nine items for

which they were to give ratings was provided (see previous section). A warm-up task using a voice that was not included in the actual experiment was also carried out.

The main experiment then started, and participants were asked to listen to the four voices, replicating always the same pattern: audio stimulus reproduction, 5-second pause, questions 1 to 9 read aloud by the researcher, oral reply by the participant that was written down by the researcher, and a final 3-second pause. The listening order of the voices was randomised across participants, always presenting the synthetic voices first to avoid a negative impact on the TTS system evaluation, as suggested by van Santen (1993), and Viswanathan and Viswanathan (2005: 62). This part of the experiment lasted 22 minutes and 36 seconds, and the test finished with the post-questionnaire, which was read aloud by the researcher, who would again write down the answers in the corresponding form.

3.6. Statistical methods

For the eight items to be considered (accentuation, pronunciation, speech pauses, intonation, naturalness, pleasantness, listening effort, acceptance) descriptive statistics (mean, median, standard deviation, minimum, maximum and percentiles) were calculated. Figures 2 and 3 display the means and the medians for all the items. A multinomial model was established for each item under analysis as the dependent variable and the type of voice as the independent variable. However, some of the items had very low frequencies in some of the categories, so they were recategorised as a binary outcome (scores 1, 2, and 3 were grouped under the category "low score", whereas scores 4 and 5 were grouped under the category "high score"). Then logistic regression models were used to assess the probability of obtaining a high score.

Overall impression was also analysed using a multinomial model, taking into account the voice, the gender, the age (categorised in under and above 50 to balance groups) and the disability type as independent variables.
All results were obtained using SAS, v 9.2 (SAS Institute Inc, USA). For the decisions, significance level was fixed at 0.05.

4. Results and discussion

From the mean scores of the items (Figure 2) we notice that the natural male voice obtains higher scores in:
- accentuation (4.761, stdev=0.495),
- acceptance (4.687, stdev=0.583),
- intonation (4.478, stdev=0.682),
- listening effort (4.597, stdev=0.605), and

- speech pauses (4.627, stdev=0.624).

The natural feminine voice, by contrast, obtains higher scores in
- pleasantness (4.373, stdev=0.671),
- naturalness (4.522, stdev=0.725),
- overall impression (4.478, stdev=0.725), and
- pronunciation (4.731, stdev=0.479).

The lowest scores for natural voices are related to the female voice acceptance (3.970, stdev=0.244) and intonation (4.343, stdev=0.708). However, for the purposes of our study, what is especially interesting is not which voice gets higher scores on what items, but to observe that the results for the synthetic voices is close to that of natural voices, and that all the scores of the synthetic voices are above 3.1, reaching 4.313 in the accentuation of the synthetic male voice and 4.284 for the pronunciation of the feminine synthetic voice.

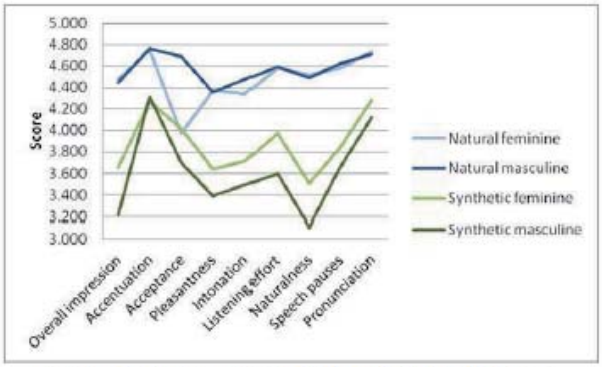

Figure 1. Mean scores of all scales for all voices.

However, since all items were collected as scores between 1 and 5, the medians (see Figure 3) may be more robust than the means. It must be stressed that all median scores are between 3.0 and 5.0. Both male and female natural voices obtain 5.0 in accentuation, listening effort, naturalness, pronunciation, speech pauses, and overall impression, and 4.0 in pleasantness. However, in acceptance and intonation the male natural voice gets higher scores (5.0 vs 4.0). This shows how the more subjective aspects which relate to end users' preferences (for instance, acceptance) present greater variation, whilst standard features that a professional describer masters (e.g. accentuation and pronunciation) are more stable. It also shows how even a natural voice may not get the highest mark in terms of pleasantness or intonation.

In as far as artificial voices are concerned, the female voice obtains 4.0 in all items under analysis, whilst the male artificial voice ranges from 3.0 (pleasantness, naturalness, overall impression) to 5.0 (accentuation), with

most items rated 4.0 on a 5-point scale (acceptance, intonation, listening effort, speech pauses, pronunciation). Again, what is especially relevant is the fact that all items are assessed above 3.0 and that in some items the median scores are the same for some natural and artificial voices. This is the case of accentuation (same scores for both natural voices and the male artificial voice), acceptance (same scores for the female natural voice and both artificial voices), pleasantness (same scores for natural voices and the female artificial voice), and intonation (same scores for the natural female voice and both artificial voices).

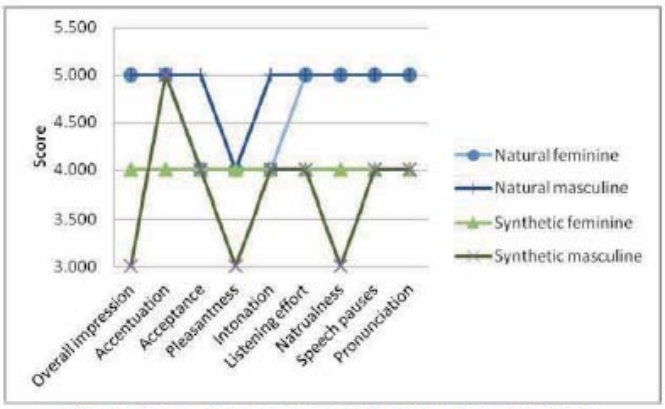

Figure 2. Median scores of all scales for all voices.

An analysis taking into account the ordinal characteristic of the items is based on the multinomial or logistic models. Statistically significant differences between the synthetic voices and their natural counterparts were found in all items under analysis. In all cases the natural voices were considered to be better than the artificial ones (see Annex 3 for further details).

When comparing the two artificial voices, the analysis shows that the synthetic feminine voice was more accepted, required less effort, was considered to be more natural and obtained a better score in the overall impression than the synthetic masculine one. As for the rest of the items (accentuation, pleasantness, intonation, speech pauses and pronunciation), no statistically significant differences were found.

Focusing on the overall impression, the multinomial model allows us to conclude that women (OR=1.67, IC=(0.96,2.90)) and the group below 50 (OR=1.89, IC=(1.09,3.30)) gave statistically significant higher scores than men and people older than 50, respectively. No statistically significant differences were found related to the disability type.

To complement previous statistical analyses, the post-questionnaire provides qualitative data that will be discussed next. When asked about their preferences in terms of a male or a female AD voice, 72% declared they did not have any preferences, with only 16% stating that it depends on the audiovisual product. The reasons for preferring either a female or a male voice in such cases were the topic (in 7 out of the 11 cases, that is 64%) and the characters (in 4 instances, that is 36%).

When asked about their preferences regarding a human or a synthetic voice, 81% of the informants stated that they preferred a human voice to read the AD, 1% declared that they preferred a synthetic voice, 3% said that it depended on the audiovisual product, and 15% declared they did not have any specific preferences as long as the artificial voice sounded natural enough and was not tiring. It must be noted, for example, that in the case of the synthetic voices tested, the naturalness mean scores were 3.507 for the female voice and 3.104 for the male one, and the listening effort mean scores were 3.836 and 3.657 respectively, which are quite strong results in a 5-point scale. It must be also stressed that 51 informants (76%) said they normally use electronic devices with synthetic voice applications on a daily basis.

When explicitly asked about the TTS AD as an alternative solution to human voiced audio description, 94% of participants responded positively. Twenty-two participants, i.e. 33%, stated that the main reason for accepting TTS AD as an alternative solution was that it would definitely increase the amount of audio described audiovisual products. Eight out of these 22 participants explained that it would reduce both the costs and time for creating such products. Nine participants (13%) stated that it could be an alternative solution because the quality of synthetic voices was already good enough. On the other hand, 10 informants (15%) stated that synthetically voiced AD was better than no AD at all, while 9 respondents (13%) argued it should only be an alternative, not the usual situation.

When questioned about specific kinds of audiovisual products, the preferences varied slightly, as shown in Figure 4: most of the participants agreed on applying TTS AD in documentaries (48 respondents), series (48 respondents) and films (49 respondents); not so many people agreed on applying it to cartoons (36 respondents) and even less informants were willing to implement it in live plays (24 respondents), with 4 participants being against implementing it at all.

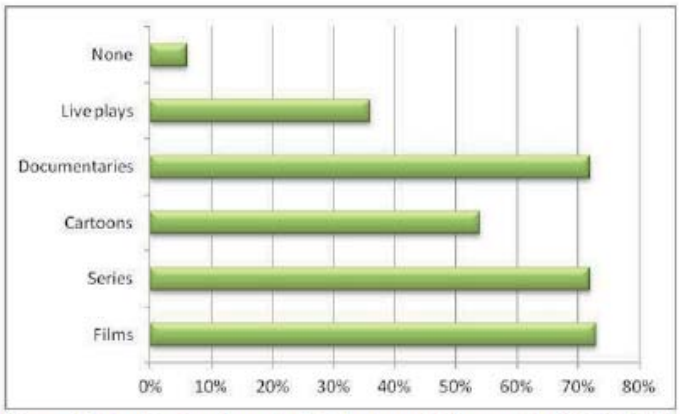

Figure 3. Audiovisual products that could be used with TTS AD.

Finally, a question about their opinion after listening to the four voices included in the experiment showed a preference for the masculine natural voice (42%) and the feminine natural voice (38%), although 14% said they preferred the feminine synthetic voice and 6% selected the male synthetic voice. These qualitative data match with the results obtained both in the descriptive and inferential statistics, which actually graded voices in the same order: the natural masculine voice was the one which obtained better mean scores, closely followed by the natural feminine, then the synthetic feminine and finally the synthetic masculine.

5. Conclusions

This article has presented a first analysis of text-to-speech audio description in Catalan, as compared to human-voiced audio descriptions, using both male and female voices. Participants have assessed the voices taking into account various items (overall impression, accentuation, pronunciation, speech pauses, intonation, naturalness, pleasantness, listening effort, and acceptance), providing data of both a quantitative and a qualitative nature.

Results show that natural voices in our experiment have statistically higher scores than synthetic voices. They also show that the synthetic feminine voice has higher mean scores than the synthetic masculine voice in all items but accentuation. This proves that the preferential choice of blind and partially sighted persons is the audio description voiced by a human, rather than by a speech synthesis system. This does not mean, though, that TTS AD is not accepted by end users, as shown by the fact that 94% of the participants consider TTS an alternative acceptable solution, and 20% of the respondents actually state that their preferred voice from the four under analysis is a synthetic one. Moreover, it is

particularly relevant that no mean score of any of the items goes under 3.1 on a 5-point scale. As an example, the acceptance item's lowest score is a 3.7 (for the synthetic masculine voice) and the overall impression item's lowest score is a 3.2 (also for the synthetic masculine voice).

This experiment follows previous research on TTS AD carried out in Poland and Japan but it is the first of its kind in Catalan. However, it also has its own limitations. First of all, since the study used a non-probability sampling approach (Bryman 2012: 418), the results cannot be generalised to the whole Catalan blind and partially sighted population. Another of its setbacks is the length of the clips: it remains to be seen whether the results would remain the same in longer productions and in various genres. It would also be highly interesting to see whether reception varies in productions originally shot in Catalan and in dubbed productions, since the language and the sound conditions are different. Another topic worth researching would not only be the perceived quality based on a list of previously selected items, but also the engagement of the audience, in line with Fryer and Freeman's research (2013). Finally, it would also be worth researching end users behaviour if given the possibility of tuning their own AD preferences, at least as far as voice, voice gender and volume are concerned, in line with Walczak and Szarkowska's approach (2012).

All in all, it is our hope that this type of research will allow us to find new ways of increasing access to culture and entertainment for the blind and visually impaired, both on traditional and new media. We are convinced that speech technologies but also other language and visual processing technologies will play a key role and will open a myriad of research possibilities.

Acknowledgements
This work has been carried out within the scope of the doctoral program in Translation and Intercultural Studies offered in the Department of Translation and Interpreting at the Universitat Autònoma de Barcelona, with the financial support of the ALST project ("Linguistic and sensorial accessibility", project ref. code FFI2012-31024 of the Spanish Ministerio de Economía y Competitividad). Anna Matamala is a TransMedia Catalonia member, funded by Generalitat de Catalunya (2014SGR27). We would like to thank Iola Ledesma and Jose Navarro, of the Escola Catalana de Doblatge (ECAD), for helping in the selection of the natural voices and in the recording of the clips used in the experiments, as well as Ana Vázquez and Anna Espinal, of the Applied Statistics Service of the Autonomous University of Barcelona, for their collaboration in defining the sample and in the statistical analysis of the data. Special thanks to the Associació Discapacitat Visual de Catalunya (ADVC) and Associació Catalana per a la Integració del Cec (ACIC) for their kindness and support and for providing us with so many volunteers for the experiment.

Bibliography

- **Alías, Francesc, Ignasi Iriondo, and Joan Claudi Socoró** (2011). "Aplicació de tècniques de generació automàtica de la parla en producció audiovisual." *Quaderns del CAC*, 37(1), 105-114. http://www.cac.cat/pfw_files/cma/recerca/quaderns_cac/Q37_Alias_etal.pdf (consulted 30.07.2014)

- **Braun, Sabine** (2011). "Creating coherence in Audio Description." *Meta: Journal des Traducteurs / Meta: Translator's Journal*, 56(3), 645-662. http://www.erudit.org/revue/meta/2011/v56/n3/1008338ar.pdf (consulted 30.07.2014)

- **Bryman, Alan** (2012). *Social Research Methods*. Oxford: Oxford University Press.

- **Caruso, Beatrice** (2012) "Audio Description Using Speech Synthesis." Languages and the Media. 9th International Conference on Language Transfer in Audiovisual Media. Conference Catalogue. Berlin, Germany: ICWE, 59-60.

- **Chapdelaine, Claude and Langis Gagnon** (2009). "Accessible Videodescription On-Demand." *ASSETS '09. Proceedings of the 11th international ACM SIGACCESS conference on Computers and accessibility*. New York: ACM, 221-222.

- **Chmiel, Agnieska and Iwona Mazur** (2012). "AD reception research: Some methodological considerations." Elisa Perego (ed.) (2012). *Emerging Topics in Translation: Audio Description*. Trieste: EUT Edizioni Università di Trieste, 57-80.

- **Cryer, Heather and Sarah Home** (2008). *Exploring the use of synthetic speech by blind and partially sighted people*. Literature review #2. Birmingham: RNIB Centre for Accessible Information (CAI).

- **Cryer, Heather, Sarah Home and Sarah Morley Wilkins** (2010). *Synthetic speech evaluation protocol*. Technical report #7. Birmingham: RNIB Centre for Accessible Information (CAI).

- **De Jong, Frans** (2006). "Access Services for Digital Television." Ricardo Perez-Amat and Álvaro Pérez-Ugena (eds) (2006). *Sociedad, integración y televisión en España*. Madrid, Spain: Laberinto Comunicación, 331-344.

- **Derbring, Sandra, Peter Ljunglöf and Maria Olsson** (2009). "SubTTS: Light-weight automatic reading of subtitles." Kristiina Jokinen and Eckhard Bick (eds) (2009). *Nodalida'09: Proceedings of the 17th Nordic Conference of Computational Linguistics*. NEALT Proceedings Series, vol. 4. Odense, Denmark: Northern European Association for Language Technology.

- **Drożdż-Kubik, Justyna** (2011). "Harry Potter i Kamień Filozoficzny słowem malowany – czyli badanie odbioru filmu z audiodeskrypcją z syntezą mowy." MA Thesis. Jagiellonian University.

- **Encelle, Benoît, Magali Ollagnier-Beldame, Stéphanie Pouchot and Yannick Prié** (2011). "Annotation-based video enrichment for blind people: A pilot study on the use of earcons and speech synthesis." *ASSETS '11: Proceedings of the 13th International ACM SIGACCESS Conference on Computers and Accessibility*. New York, USA: ACM, 123-130.

- **Fernández-Torné, Anna, Anna Matamala and Carla Ortiz-Boix** (2012). "Technology for accessibility in multilingual settings: the way forward in AD?" Paper presented at *The translation and reception of multilingual films* (University of Montpellier 3, 15-16 June 2012). http://ddd.uab.cat/record/117160 (consulted 30.07.2014)

- **Fernández-Torné, Anna and Anna Matamala** (2013). "Methodological considerations for the evaluation of TTS AD's acceptance in the Catalan context." Paper presented at *ARSAD (Advanced Research Seminar)*. (Autonomous University of Barcelona, 13-14 March 2013). http://ddd.uab.cat/record/117078 (consulted 30.07.2014)

- **Freitas, Diamantino and Georgios Kouroupetroglou** (2008). "Speech technologies for blind and low vision persons." *Technology and Disability* 20, 135-156.

- **Fryer, Louise and Jonathan Freeman** (2013). "Visual impairment and presence: measuring the effect of audio description." *Proceedings of the 2013 Inputs-Outputs Conference: An Interdisciplinary Conference on Engagement in HCI and Performance*. New York, USA: ACM, article n° 4.

- **González García, Luis** (2004). "Assessment of text reading comprehension by Spanish-speaking blind persons." *British Journal of Visual Impairment* 22 (1), 4-12.

- **Hinterleitner, Florian *et al.*** (2011). "An Evaluation Protocol for the Subjective Assessment of Text-to-Speech in Audiobook Reading Tasks." *Proceedings of the Blizzard Challenge Workshop*. International Speech Communication Association (ISCA).

- **ITU-T Recommendation** P.85 (1994). *Telephone transmission quality subjective opinion tests. A method for subjective performance assessment of the quality of speech voice output devices*. Geneva, Switzerland: ITU. http: //www.itu.int/rec/T-REC-P.85-199406-I/en (consulted 23.07.2014).

- **Kobayashi, Masatomo, Kentarou Fukuda, Hironobu Takagi, and Chieko Asakawa** (2009). "Providing synthesized audio description for online videos." *ASSETS '09: Proceedings of the 11th International ACM SIGACCESS Conference on Computers and Accessibility*. New York, USA: ACM, 249-250.

- **Kobayashi, Masatomo, Trisha O'Connell, Bryan Gould, Hironobu Takagi, and Chieko Asakawa** (2010). "Are Synthesized Video Descriptions Acceptable?" *ASSETS '10: Proceedings of the 12th International ACM SIGACCESS Conference on Computers and Accessibility*. New York, USA: ACM, 163-170.

- **Ljunglöf, Peter, Sandra Derbring and Maria Olsson** (2012). "A free and open-source tool that reads movie subtitles aloud." *Proceedings of the Third Workshop on Speech and Language Processing for Assistive Technologies*. Montreal, Canada: Association for Computational Linguistics (ACL), 1-4.

- **Llisterri, Joaquim, Natividad Fernàndez, Francesc Gudayol, Juan José Poyatos and Josep Martí** (1993). "Testing users' acceptance of Ciber232, a test to speech system used by blind people." Granström, B., Hunnicutt, S., and Spense, K.E. (eds) (1993) *Speech and Language Technology for Disabled Persons. Proceeding of an ESCA Workshop*. Stockholm, Sweden. 203-206.

- **Mączyńska, Magdalena** (2011). *TTS AD with audio subtitling to a non-fiction film. A case study based on* La Soufriere *by Werner Herzog*. MA Thesis. University of Warsaw.

- **Matamala, Anna, Anna Fernández-Torné and Carla Ortiz-Boix** (2013). "Enhancing sensorial and linguistic accessibility: further developments in the TECNACC and ALST projects." Paper presented at the *5th International Conference Media for All. Audiovisual Translation: Expanding Borders*. (Dubrovnik, Croatia, 25-27 September 2013). http://ddd.uab.cat/record/116868 (consulted 30.07.2014)

- **Mieskes, Margot and Juan Martínez Pérez** (2011). "A Web-based Editor for Audio-titling using Synthetic Speech." *3rd International Symposium on Live Subtitling with Speech Recognition*. Antwerp http://www.respeaking.net/Antwerp%202011/Webbased_editor.pdf (consulted 30.07.2014)

- **Nielsen, Simon and Hans-Heinrich Bothe** (2007). "SUBPAL: A Device for Reading Aloud Subtitles from Television and Cinema." Marion A. Hersh and James Ohene-Djan (eds) (2008). *Proceedings of the Conference & Workshop on Assistive Technologies for People with Vision & Hearing Impairments Assistive Technology for All Ages CVHI 2007*. CEUR Workshop Proceedings, vol. 415. http://ceur-ws.org/Vol-415/paper17.pdf (consulted 30.07.2014)

- **Oncins, Estella, Oscar Lopes, Pilar Orero, Javier Serrano and Jordi Carrabina** (2013). "All together now: a multi-language and multi-system mobile application to make living performing arts accessible." *The Journal of Specialised Translation* 20, 147-164.

- **Orero, Pilar** (2007). "Sampling audio description in Europe." Jorge Díaz Cintas, Pilar Orero and Aline Remael (eds) (2007). *Media for All. Subtitling for the Deaf, Audio Description, and Sign Language*. Amsterdam/New York: Rodopi, 111-125.

- **Orero, Pilar and Anna Matamala** (2007). "Accessible opera: overcoming linguistic and sensorial barriers." *Perspectives. Studies in Translatology* 15(4), 262-277. https://ddd.uab.cat/record/117149 (consulted 30.07.2014)

- **Pazos, Patricia** (2012). *Audiosubtitulación: una posible solución para la accesibilidad a los medios audiovisuals*. MA thesis. Autonomous University of Barcelona.

- **Salway, Andrew, Elia Tomadaki and Andrew Vassiliou** (2004). *Building and analysing a corpus of AD scripts. TIWO Television in Words. Report on Workpackage 2*. Surrey, UK: University of Surrey.

- **Szarkowska, Agnieska** (2011). "Text-to-speech audio description: towards wider availability of AD." *The Journal of Specialised Translation* 15, 142-162.

- **Szarkowska, Agnieska and Anna Jankowska** (2012). "Text-to-speech audio description of voice-over films. A case study of audio described *Volver* in Polish." Elisa Perego (ed.) (2012). *Emerging topics in translation: Audio description*. Trieste, Italy: EUT Edizioni Università di Trieste, 81-98.

- **van Santen, Jan P.H.** (1993). "Perceptual experiments for diagnostic testing of text-to-speech systems." Computer Speech & Language, 7(1): 49–100.

- **Verboom, Maarten, David Crombie, Evelien Dijk and Mildred Theunisz** (2002). "Spoken Subtitles: Making Subtitled TV Programmes Accessible." Klaus Miesenberger, Joachim Klaus and Wolfgang L. Zagler (eds) (2002). *Proceedings of Computers Helping People with Special Needs, 8th International Conference, ICCHP 2002*. Berlin-Heidelberg, Germany: Springer-Verlag, 295-302.

- **Viswanathan, Mahesh and Madhubalan Viswanathan** (2005). "Measuring speech quality for text-to-speech systems development and assessment of a modified mean opinion score (MOS) scale." *Computer Speech and Language* 19, 55-83.
- **Walczak, Agnieska** (2010). *Audio description for children. A case study of text-to-speech audio description of educational animation series* Once Upon a Time... Life. MA Thesis. University of Warsaw.
- **Walczak, Agnieska and Agnieska Szarkowska** (2012). "Text-to-speech audio description of educational materials for visually impaired children." Silvia Bruti and Elena Di Giovanni (eds) (2012). *Audio Visual Translation across Europe: An Ever-Changing Landscape*. Berna/Berlin: Peter Lang, 209-234.

Website

World Health Organisation (2013) Fact Sheet no. 282. http://www.who.int/mediacentre/factsheets/fs282/en/ (consulted 17.07.2014).

Biography

Anna Fernández-Torné holds an MA in Audiovisual Translation (UAB), MA in Language Consultancy in the Media (UAB) and a European MA in Audiovisual Translation (Parma University). Her research centres on audio description and technologies, both text-to-speech and machine translation. She has been a freelance translator since 2004, specialising in audiovisual translation, and lectures at the MA in Audiovisual Translation at Universitat Autònoma de Barcelona. anna.torne@gmail.com

Anna Matamala, holds a PhD in Applied Linguistics (UPF) and is a tenured lecturer at the UAB. She has published in international refereed journals such as *Meta, Perspectives, Babel, VIAL* and *Linguistica Antverpiensia*, among others. She is the author of a book on interjections and lexicography (IEC, 2005), co-author of a book on voice-over (Peter Lang, 2010), and co-editor of three volumes on AVT and media accessibility. anna.matamala@uab.cat

The Journal of Specialised Translation Issue 24 – July 2015

Annex 1. Questionnaire

How would you describe the quality of the voice you have just heard?
1. Bad
2. Regular
3. Neutral
4. Good
5. Excellent

Did you detect anomalies in terms of the accentuation of words?
1. Yes, al lot of them
2. Yes, many
3. Yes, some
4. Yes, but only a few
5. No, none

Did you notice anomalies in terms of pronunciation?
1. Yes, al lot of them
2. Yes, many
3. Yes, some
4. Yes, but only a few
5. No, none

Do you think the voice makes pauses when it is needed?
1. No, never
2. No, almost never
3. Yes, normally
4. Yes, almost always
5. Yes, always

How would you rate the intonation of sentences?
1. Very bad
2. Bad
3. Good
4. Quite good
5. Very good

How would you define the degree of naturalness of the voice?
1. Very unnatural
2. Unnatural
3. Natural
4. Quite natural
5. Very natural

To what extent do you deem this voice to be pleasant?
1. Very unpleasant
2. Unpleasant

3. Neutral
4. Pleasant
5. Very pleasant

Do you think listening to this voice for a long time would be tiring?
1. Yes, a lot
2. Yes, quite a lot
3. Yes, a little bit
4. No, not much
5. No, not at all

Do you think this voice could be used for voicing audio descriptions?
1. No, never
2. No, almost never
3. Yes, in some cases
4. Yes, in many cases
5. Yes, always

Annex 2. Post-questionnaire

*Mandatory field

- Identifier *

Enter your initials (first name initial, first surname initial and second surname initial) followed by your age. Do not leave any blank space in between.

- Age*
- Sex*

Male / Female

- Level of studies reached*

Lower than first stage of secondary school
Secondary education, first stage
Secondary education, second stage
Advanced vocational education
First cycle university education (diploma, degree, engineering or graduate studies)
Second cycle university education (master, postgraduate or doctoral studies)

- In case you have reached university education, please specify.
- Occupation*

Public administration management and management of companies with 10 or more wage earners.
Management of companies with less than 10 wage earners
Management of companies without wage earners
Professions associated with 2nd and 3rd cycle university degrees and the like
Professions associated with a 1st cycle university degree and the like
Support technicians and professionals
Administrative type employees
Catering services workers and personal services workers
Protection and security service workers
Retail workers and the like
Workers skilled in agriculture and fishing
Skilled construction workers, except machinery operators
Skilled workers in the extractive industry, metallurgy, construction of machinery and related trades.
Skilled workers from the graphic arts, textile and tailoring, elaboration of food, cabinetmakers, craftspersons and other similar industries
Fixed machinery and industrial installation operators; fitters and assemblers.
Mobile machinery drivers and operators
Unskilled workers in the service sector (except transports)
Agriculture, fishing, construction, manufacturing industries and transport labourers.
Armed forces
Unemployed for longer than one year
Unemployed, seeking a first job

- Profession in your own words
- Kind of visual impairment according to WHO*
Blindness / Low vision
- How long have you been visually impaired for? *
From birth / For less than 1 year / For between 1 and 10 years / For between 11 and 20 years / For more than 20 years any
- Have you ever seen an audio described product (films, series, theatre plays, etc.)?*
Yes / No
- In case you have, which kind of products? (You can tick more than one answer)
Films / Series / Cartoons / Theatre plays / Opera plays
- How often do you use audio described products?*
At least once a day / At least once a week / At least once a month / Never / Other
- Do you prefer the AD to be read by*
A man / A woman / It depends on the audiovisual product / I don't care
- If it depends on the audiovisual product, what does it depend on exactly?
- You prefer the AD to be read by*
A human voice / An artificial voice / It depends on the audiovisual product / I don't care
- If it depends on the audiovisual product, what does it depend on exactly?
- Do you use electronic devices with synthetic voice applications, such as mobile phones or computers?*
Yes / No
- How often do you use them?*
At least once a day / At least once a week / At least once a month / Never
- Have you ever used audio described products with synthetic voice?*
Yes / No
- Do you think it is an alternative solution to human voiced audio description?*
Yes / No
- Why do you think so?*
- What kind of products would you use with synthetic voiced AD? (You can tick more than one answer)*
Films / Series / Cartoons / Documentaries / Live plays / None
- Which voice, from the 4 voices you have just heard, did you like the most?*
The masculine synthetic voice / The masculine natural voice / The feminine synthetic voice / The feminine natural voice
- Would you be able to rank them in order, from the one you like the most to the one you liked the least?
- Other comments

Annex 3. Odds ratio (OR) tables

Voice	OR Estimate	OR IC Lower	OR IC Upper
Sinthetic masculine vs synthetic feminine	0.3878	0.2041	0.7367
Synthetic masculine vs natural masculine	0.06400	0.03085	0.1328
Synthetic masculine vs natural feminine	0.05216	0.02484	0.1095
Synthetic feminine vs natural masculine	0.1650	0.08252	0.3300
Synthetic feminine vs natural feminine	0.1345	0.06652	0.2719
Natural masculine vs natural feminine	0.8150	0.4144	1.6029

Overall impression

Voice	OR Estimate	OR IC Lower	OR IC Upper
Sinthetic masculine vs synthetic feminine	1.1856	0.5856	2.4005
Synthetic masculine vs natural masculine	0.1604	0.06801	0.3783
Synthetic masculine vs natural feminine	0.1774	0.07617	0.4129
Synthetic feminine vs natural masculine	0.1353	0.05743	0.3187
Synthetic feminine vs natural feminine	0.1496	0.06433	0.3478
Natural masculine vs natural feminine	1.1058	0.4334	2.8212

Accentuation

Voice	OR Estimate	OR IC Lower	OR IC Upper
Sinthetic masculine vs synthetic feminine	0.6593	0.3323	1.3081
Synthetic masculine vs natural masculine	0.1128	0.05074	0.2506
Synthetic masculine vs natural feminine	0.1000	0.04445	0.2251
Synthetic feminine vs natural masculine	0.1710	0.07730	0.3784
Synthetic feminine vs natural feminine	0.1517	0.06774	0.3398
Natural masculine vs natural feminine	0.8871	0.3765	2.0898

Pronunciation

Voice	OR Estimate	OR IC Lower	OR IC Upper
Sinthetic masculine vs synthetic feminine	0.6419	0.3409	1.2088
Synthetic masculine vs natural masculine	0.08431	0.04036	0.1761
Synthetic masculine vs natural feminine	0.1045	0.05101	0.2142
Synthetic feminine vs natural masculine	0.1313	0.06367	0.2710
Synthetic feminine vs natural feminine	0.1628	0.08043	0.3297
Natural masculine vs natural feminine	1.2397	0.5845	2.6297

Speech pauses

Voice	OR Estimate	OR IC Lower	OR IC Upper
Sinthetic masculine vs synthetic feminine	0.6922	0.3463	1.3835
Synthetic masculine vs natural masculine	0.1176	0.04655	0.2973
Synthetic masculine vs natural feminine	0.1568	0.06638	0.3701
Synthetic feminine vs natural masculine	0.1700	0.06693	0.4316
Synthetic feminine vs natural feminine	0.2265	0.09541	0.5375
Natural masculine vs natural feminine	1.3324	0.4610	3.8508

Intonation

Voice	OR Estimate	OR IC Lower	OR IC Upper
Sinthetic masculine vs synthetic feminine	0.4920	0.2652	0.9126
Synthetic masculine vs natural masculine	0.07831	0.03867	0.1586
Synthetic masculine vs natural feminine	0.07904	0.03905	0.1600
Synthetic feminine vs natural masculine	0.1592	0.08044	0.3150
Synthetic feminine vs natural feminine	0.1607	0.08124	0.3177
Natural masculine vs natural feminine	1.0092	0.4981	2.0447

Naturalness

Voice	OR Estimate	OR IC Lower	OR IC Upper
Sinthetic masculine vs synthetic feminine	0.5685	0.2815	1.1482
Synthetic masculine vs natural masculine	0.07245	0.02695	0.1948
Synthetic masculine vs natural feminine	0.08628	0.03367	0.2211
Synthetic feminine vs natural masculine	0.1274	0.04734	0.3431
Synthetic feminine vs natural feminine	0.1518	0.05913	0.3895
Natural masculine vs natural feminine	1.1909	0.3705	3.8281

Pleasantness

Voice	OR Estimate	OR IC Lower	OR IC Upper
Sinthetic masculine vs synthetic feminine	0.4473	0.2363	0.8468
Synthetic masculine vs natural masculine	0.1148	0.05619	0.2347
Synthetic masculine vs natural feminine	0.1205	0.05921	0.2453
Synthetic feminine vs natural masculine	0.2568	0.1272	0.5181
Synthetic feminine vs natural feminine	0.2695	0.1340	0.5417
Natural masculine vs natural feminine	1.0494	0.5012	2.1973

Listening effort

Voice	OR Estimate	OR IC Lower	OR IC Upper
Sinthetic masculine vs synthetic feminine	0.3832	0.1814	0.8096
Synthetic masculine vs natural masculine	0.07689	0.02486	0.2378
Synthetic masculine vs natural feminine	0.01831	0.002365	0.1417
Synthetic feminine vs natural masculine	0.2007	0.06254	0.6438
Synthetic feminine vs natural feminine	0.04778	0.006047	0.3775
Natural masculine vs natural feminine	0.2381	0.02552	2.2217

Acceptance

Annex 1.2. Fernández-Torné, A. (Forthcoming). Machine Translation Evaluation through Post-Editing Measures in Audio Description. *inTRAlinea*. 2016, 18.

Annex 1.3. Fernández-Torné, A., Matamala, A. (2016). Machine Translation in Audio Description? Comparing Creation, Translation and Post-editing Efforts. *Skase*, 9(1), 64-85.

Machine translation in audio description?

Comparing creation, translation and post-editing efforts

Anna Fernández-Torné

Anna Matamala

Abstract: Machine translation has been proved worthwhile, in terms of time saving and productivity gains, in technical and administrative translation domains. In order to examine whether this also applies to audio description, an experiment comparing the efforts of creating an audio description from scratch, of translating it manually from English into Catalan and of post-editing its machine translated version has shown that the objective post-editing effort is lower than creating it ex novo. *However, the subjective effort is perceived to be higher.*

Keywords: accessibility; audio description; audiovisual translation; machine translation; Catalan language; post-editing effort

Introduction

The presence of audiovisual content in our society is increasing at a dramatic pace. New ways of making this growing volume of audiovisual content accessible to all audiences faster – and at lower costs, if possible – need to be researched and developed, and the implementation of technologies in audiovisual translation (AVT) seems to be the way forward, as it has already been proved efficient in other translation domains.

Machine translation (MT) is one of the technologies that is becoming common practice in the professional translation arena (Koponen 2015, Daems et al. 2015), and translators' productivity gains using MT have been broadly demonstrated (Guerberof 2009; Plitt and Masselot 2010). MT with post-editing (PE) – that is, with a revision by a professional – is already part of the workflow of many translation service providers dealing with technical texts and also of public administrations aiming "to quickly check the general meaning of incoming information" (European Commission n.d.). However, "[t]he adoption rate of MT and PE processes naturally varies in different countries and language pairs" (Koponen 2015: 3), and in translation domains, too. This is where audiovisual translation in general, and audio description in particular, lags behind. Audio descriptions, the translation of

images into words addressed to an audience who cannot access the visual content (Maszerowska et al. 2014), are nowadays generally created independently in each language and are only seldom translated, being the application of MT being non-existent to the best of our knowledge.

This article presents the results of an experiment in which MT was implemented in audio description (AD) for the English-Catalan language pair. The experiment compared the effort, both objective and subjective, in three different scenarios: when creating an audio description in Catalan (AD creation), when translating an English audio description into Catalan (AD translation), and when post-editing a machine-translated audio description from English into Catalan (AD PE). Our ultimate aim is to explore whether MT could be satisfactorily deployed in audio description, hence the focus of the analysis is the comparison of AD PE in relation to AD creation, which is currently the standard process. However, another possibility has also been taken into account, i.e. human translation, a process already discussed in the literature in relation to audio description (Matamala 2006, Jankowska 2013). Results in this regard are also provided, although they are discussed to a lesser extent.

The article begins with an overview of related work. Next, the experimental set-up is presented, with a thorough description of the participants, test data, effort assessment methods, test development, and statistical methods used. In the following section, a comprehensive exposition of the results is presented and discussed, and finally, conclusions are drawn while proposing directions for further research.

Related work

The application of MT to audiovisual content is still in its early stages. In recent decades the EU has funded several projects dealing with the automatic generation of subtitles and their translation into multiple languages both in media – MUSA (2002-2004), eTITLE (2003-2005), SUMAT (2011-2014) and EU-BRIDGE (2012-2014) –, and educational content – transLectures (2011-2014) and EMMA (2014-2016). Research has also been carried out to assess the quality of machine-translated or post-edited audiovisual translations such as subtitles (Armstrong et al. 2006; Volk 2009; Del Pozo et al. 2014) and, more recently, voice-overs (Ortiz-Boix and Matamala 2015). However, the implementation of MT in audio description has not yet attracted the attention of many researchers, and only the ALST project (Matamala 2015) has ventured into the topic, proving so far the feasibility of machine translating filmic AD in the Catalan-Spanish language pair (Ortiz-Boix 2012). This article is part of that project, and focuses on comparing the effort involved in generating an AD when using different methods. That is why this section will succinctly describe previous research in post-editing effort, placing special emphasis on its measurement.

The general framework used in many studies to assess post-editing effort is Krings' (2001) proposal. Krings differentiates between temporal, technical and cognitive effort. Temporal effort is the total time spent on post-editing a text, technical effort refers to the operations carried out to post-edit the text, and cognitive effort applies to the mental processes involved in identifying errors in raw machine-translated texts and in deciding on the necessary steps to correct them.

Measuring temporal effort is straightforward. Technical effort can also be directly observed by using methods such as key-logging technologies (Guerra 2003, Tatsumi and Roturier 2010). However, cognitive effort is not directly observable. Krings (2001) used think-aloud protocols to determine cognitive effort, but he noticed that this method affected the total process time. Other technologies, such as key-logging (O'Brien 2004) and eye-tracking (O'Brien 2011, Carl et al. 2011), have successfully been used, since they allow subjects' behaviour to be recorded unobtrusively in real time. Pauses have been considered a key indicator of cognitive effort. Indeed, in writing research pauses are "assumed to provide us with a window to the cognitive processes underlying language production" (Wengelin 2006: 108, cited in Chukharev-Hudilainen 2014: 64), and they are usually computed, particularly their frequency, duration and position. Lacruz, Denkowski and Lavie (2014) state that both average pause ratio (APR), i.e. the average time per pause divided into the average time per word, and pause to word ratio (PWR), i.e. the number of pauses divided into the number of words, correlate well with cognitive effort: the lower the APR and the higher the PWR, the higher the levels of cognitive effort.

Most of the research carried out so far in this area has focused on technical documents, since this is where machine translation is more extensively used. In the field of audiovisual translation, studies on post-editing effort are more limited: De Sousa, Aziz and Specia (2011) compare the temporal effort involved in translating subtitles from English into Brazilian Portuguese compared to post-editing draft versions produced using translation tools, both MT and TM. Results show that "translating from scratch consistently takes 70% longer than post-editing the same sentence" (*ibid.*: 5). On the other hand, Ortiz-Boix and Matamala (forthcoming) compare the effort involved in translating wildlife documentary excerpts compared to post-editing them. Their results seem to indicate that post-editing may imply less effort than translating, although statistically significant results are not achieved in all parameters under analysis.

Methodology

This section describes methodological aspects such as the selection of participants, the test data, the measurement tools, the test development, and the statistical methods. The whole procedure was approved by the Ethics Committee of the Universitat Autònoma de Barcelona.

Participants

The participants' profile was controlled to avoid high variability which could distort the results of the test. Volunteers were recruited from among native Catalan-speaking students of an MA in AVT.

Fourteen participants took part in the experiment, but for technical reasons only the results of twelve could be used. It should also be noted that one task of one participant was not adequately recorded (translation of clip A), but since the other data were available they were included in the analysis.

Two participants were male (17%), and ten were female (83%), with a mean age of 25.8 years. All but one had a BA in Translation and Interpreting and all of them finished their MA in Audiovisual Translation in June 2014, when the test took place. They had the same experience as far as AVT and AD creation was concerned: only as students had they translated audiovisual products and created ADs.

In relation to their attitude towards translating ADs and of post-editing machine-translated ADs, participants showed a general negative prejudice towards post-editing machine-translated ADs. Prior to the test, when presented with the statement "Machine translating ADs created in other languages and post-editing them conveniently is useful" and asked to express their level of agreement on a 5-point Likert scale (1 being "strongly disagree" and 5, "strongly agree"), two participants (16.6%) chose 1, six participants (50%) chose 2, and four participants (33.3%) chose 3. When the statement presented was "Translating ADs created in other languages is useful", only one participant selected 2 (8.3%), seven selected 3 (58.3%) and four participants (33.3%) selected 4, indicating a more positive attitude towards human translation. When asked to comment on their choices, they argued that MT plus PE would lack naturalness, would convey more calques, and the task itself would often be as time-consuming as creating an AD from scratch.

Test data selection

Three clips from the film *Closer* (2004, directed by Mike Nichols) were chosen as test data. This film was selected for various reasons. First, since this experiment was part of a wider project in which other technologies such as speech recognition (SR) (Delgado, Matamala and Serrano 2015) and text-to-speech (Fernández-Torné and Matamala 2015) were tested in AD, a film both in English and in Catalan (dubbed version), with AD in both languages, was required. Secondly, a film with a non-specific genre addressed to adults was favoured, so that children films were considered out of scope, and a film within a 'miscellaneous' category according to the classification by Salway et al. (2004) was searched for.

The clips' duration was established at approximately three minutes to minimise participants' fatigue, as they would have to create, translate and post-edit three different AD excerpts in just one session. The number of words included in the AD to be translated was also controlled to balance the test duration, so that in no case would the translation and post-editing take longer than one hour. Neutral clips in terms of content were chosen in order to avoid any potential distraction or offense to the participants. Finally, clip excerpts from the development of the plot, rather than the beginnings, were chosen as specific constraints are generally to be found in terms of AD creation at the beginning of a film (Remael and Vercauteren 2007).

For each clip three versions were available: (1) for the AD creation, the audiovisual Catalan dubbed version of the excerpts; (2) for the human translation task, the audiovisual Catalan dubbed version with the audio description in English provided as written text with time codes. This AD corresponds to the one included in the commercial DVD released in 2005; (3) for the AD PE task, the audiovisual Catalan dubbed version with the audio description in English, provided as written text with time codes, plus the machine translation generated by Google Translate of the English AD, also provided as written text with time codes. Google Translate was chosen as the best free online engine available in the chosen language pair and domain in a pre-test (Fernández-Torné forthcoming).

Assessment measures

Following Krings (2001), effort was split up into three categories: temporal effort, technical effort, and cognitive effort. Even though this classification was designed for the assessment of effort in post-editing tasks, it was also deemed adequate for evaluating creation and translation efforts, since they can all be considered comparable indicators of text production, as explained by Dam-Jensen and Heine (2013). According to the authors, there are three types of text production, i.e. writing, translation and adaptation, which relate differently to pre-existing texts. In this sense, adaptation – post-editing in our case – "can be seen as an 'intermediate type' as it depends on a source text (or more than one), as does translation, but involves a shift in text type by means of paraphrasing, revising or summarizing" (Dam-Jensen and Heine 2013: 92).

Although "post-editing time, a simple and objective annotation, can reliably indicate translation post-editing effort in a practical, task-based scenario" (Specia 2011: 73) and can also be seen "as a way to assess some of the cognitive effort involved in post-editing" (Koponen et al. 2012: 1), effort was assessed by measuring several parameters from each category, namely:

- Temporal effort: total process time, time spent in Subtitle Workshop.
- Technical effort: keyboard actions (including total character types and other keystrokes), mouse actions (including clicks, movements and scrolls), switches keyboard to mouse, and total window transitions.

- Cognitive effort: total pause time, mean pause time, number of pauses and PWR.

All these elements were automatically recorded by the key-logging tool InputLog 5.2.01 (Leijten and Van Waes 2013), and are referred to in the analysis as objective effort.

The total process time was measured to determine the temporal effort. An additional indicator was the time spent in the software where the actual creation, translation or post-editing took place: our belief is that the less the time spent in Subtitle Workshop, the more temporal effort involved in searching for information or solving doubts on the Internet.

Regarding the technical effort, both keyboard actions (itemising total characters typed and other keystrokes) and mouse actions (differentiating between clicks, movements and scrolls) were calculated. Although deletion and insertion operations are considered to be direct indicators of technical effort (Krings 2001: 179), they could not be recorded in the selected software. Instead, switches from keyboard to mouse and total number of window transitions were computed, as these are also operations made during the process.

Concerning cognitive effort, PWR was calculated as the main indicator of cognitive effort in post-editing (Lacruz, Denkowski and Lavie 2014). Other aspects related to pauses – total pause time, number of pauses, mean pause time – were also assessed, as pauses have been found to be good indicators of cognitive demand, not only in writing research but also in translation (Lacruz, Shreve and Angelone 2012). It must be highlighted at this point that O'Brien (2006) did not find significant evidence to prove that pauses are actually related to cognitive effort in post-editing, but since they have largely been proved to correlate well with cognitive load in both written and spoken language production and translation research, they were considered in the present study, where they are defined as any scriptural inactivity of more than 300 ms (Lacruz, Denkowski and Lavie 2014).

Apart from these objective measures, it was considered interesting to assess the participants' subjective effort, similar to what De Sousa, Aziz and Specia (2011) did. Data on participants' perceived effort and opinions were gathered via a questionnaire administered after each task, and was compared to the participants' expected effort and opinions, gathered also via a questionnaire.

Questionnaire design

A profile questionnaire (PQ) was designed to gather personal information on participants, such as age, sex, and level of education.

A general questionnaire (GQ) was developed to gather the participants' attitudes to post-editing and translating audio descriptions, and their opinions on various aspects both before performing the test (expectations) and after having performed it (perceptions). The GQ included four statements for each of the tasks under analysis (AD creation, AD translation,

AD PE) to which participants had to indicate their level of agreement on a 10-point numerical scale:

- Rate the tasks according to the effort you think they will involve for you
- Rate the tasks according to how much you think they will impair creativity
- Rate the task according to how much you think they will be boring
- Rate the task according to the quality you think they will achieve

A slight variation was included in the GQ to be administered after the experiments: verb tenses were changed from "will involve" to "have involved", and an additional open field to justify their choices was added.

As can be seen from the previous statements, the issues under analysis relate to effort, creativity impairment, boredom, calque conveyance, and output quality, as these are aspects often mentioned in relation to post-editing. Subjective ratings were deemed important not only to complement objective data, but also to check whether their expectations on the tasks were met and to examine whether their attitudes towards any of the tasks changed once they had performed them.

Three post-task questionnaires (PTQ) were also designed to obtain data on the participants' views immediately after performing each of the tasks. A first set of questions asked participants to rate their level of agreement with a series of statements on a 5-point Likert scale, including an open field for comments. The statements read:

a) In the AD creation PTQ:
   - The clip was easy to audio describe.
b) In the AD translation PTQ:
   - The source text was easy to translate.
   - The clip was easy to audio describe departing from the original AD.
c) In the AD PE PTQ:
   - The clip was easy to audio describe departing from the MT AD.
   - The machine-translated text was easy to post-edit.
   - The machine-translated text required no post-editing.
   - The machine-translated text was fluent Catalan.
   - All the information in the source text was present in the machine-translated text.

Additionally, in the AD translation and in the AD PE PTQ, a question specifically asked whether there were any elements participants had had to adapt from the departure text (be it the English AD or the MT output) and, if so, which. Possible answers included "amount of information", "length of descriptions, "frequency of descriptions", "number of incomplete sentences (with no verb)", "register (too formal or too colloquial"), and also an open field.

As can be seen from the previous statements, some of them allowed for an easy comparison between tasks, for instance in terms of ease.

Test development

The experiment was carried out in a controlled environment (laboratory conditions), following a within-subjects design. A pilot test allowed improvements to the experimental design.

The experiment was divided into two parts. In the first part participants were asked to fill in the PQ and the GQ. They were then requested to watch the Catalan dubbed version of the film *Closer* from beginning to end uninterruptedly, so that they all had the same contextual information. Then there was a 30-minute break.

In the second part of the experiment, they were asked to create the Catalan AD, to translate the English AD into Catalan and to fully post-edit the English to Catalan machine-translated AD of the three three-minute-long excerpts.

The instructions for the AD creation stated that they should deliver a Catalan audio description according to the Catalan AD style. As for the AD translation, they were told that an English AD with spotting (time coding of the AD units) would be given to them and their task was to create a Catalan AD, modifying time-codes and AD units if needed. They were told that they should adapt the original AD to the Catalan AD style, which should fit with the Catalan dubbed version provided. The same instructions were used for the AD PE task. Moreover, the following specific guidelines inspired by the works of O'Brien (2010), TAUS and CNGL (2010), Specia (2011), De Sousa, Aziz, and Specia (2011) and Housley (2012) were included for PE:

- Perform the minimum amount of editing necessary to make the AD translation ready for voicing retaining as much raw translation as possible
- Aim for grammatically, syntactically and semantically correct translation.
- Ensure that no information has been accidentally added or omitted.
- Ensure that the message transferred is accurate.
- Ensure that key terminology is correctly translated.
- Basic rules regarding spelling, punctuation and hyphenation apply.

The order of the tasks and clips was balanced across participants. Participants were asked to perform all three tasks using Subtitle Workshop 2.51 (http://subworkshop.sourceforge.net/index), a software they were all familiar with. Although it is a subtitling software, Subtitle Workshop was chosen because it includes an integrated video player and allows inserting or editing time codes where appropriate for the synchronisation of the audio description.

After performing each task, a PTQ was administered to all participants. Once all tasks were finished, they were asked to complete the GQ, as described in the previous sub-section.

Statistical methods

Descriptive statistics (mean, median, standard deviation, minimum and maximum) were computed for all quantitative variables. A bivariate analysis was performed to determine the relationship between each variable and the task being performed. For the comparison of the tasks, a repeated measures model was used, taking into account that each participant had performed all three tasks. All results were obtained using SAS, v9.3 (SAS Institute Inc., Cary, NC, USA). For the decisions, significance level was fixed at 0.05.

Results and discussion

This section presents and discusses the results in the three tasks under analysis: AD creation, AD translation, and AD PE. Objective effort results are presented first, followed by the analysis of subjective effort and participants' views. When differences between tasks are statistically significant, it is explicitly mentioned in the discussion. Non-statistically significant data are also provided because they may illustrate relevant differences in the processes.

Objective effort

Temporal effort

Mean total process times for the AD creation and AD PE tasks were quite close to each other: 2,696.880 seconds (44.95 minutes) was the mean total process time for AD creation, whereas 2,666.695 seconds (44.44 minutes) was the total for the AD PE task. Although the figure for AD translation was higher (2,919.641 seconds, i.e. 48.66 minutes), there were no statistically significant differences among the three tasks.

The amount of time spent in Subtitle Workshop, where the actual task was to be performed, was also calculated. AD PE and AD translation presented a closer mean time (2,238.552 seconds, i.e. 37.31 minutes, and 2,245.303 seconds, i.e. 37.42 minutes, respectively) spent on the software, and for AD creation the time spent was only slightly higher (2,415.218 seconds, 40.25). Again, the difference was not statistically significant.

When calculating relative values (see *Figure 1*), it was observed that in AD creation participants spent 90% of the time in Subtitle Workshop, which means it was the task requiring less research on the Internet, whereas AD translation was the task requiring most

time outside Subtitle Workshop (33%). Post-editing was somewhere in between, dedicating 84% to the Subtitle Workshop and 16% to searching the Internet. These results can be seen as a logical consequence of the processes associated with each task: while AD can be considered a creative and introspective task, translation is usually associated with dictionary searches and on-line consultations. On the other hand, PE mainly implies rewording, word reordering and error correction, which do not necessarily involve as many Internet searches.

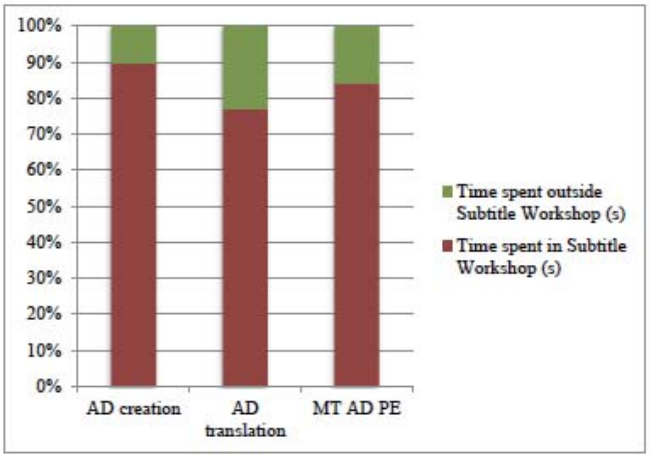

Figure 1 Time spent inside and outside Subtitle Workshop

Globally, although differences were not statistically significant, AD post-editing was the task that presented the lowest total process time and therefore, the least temporal effort.

Technical effort

Taking all keyboard actions as a whole, AD creation rendered the highest number of keyboard actions, with an average of 2,948.417. AD translation had an average of 2,656.545 actions, while AD PE presented only 1,973 actions on average. However, only the difference between AD creation and AD PE was statistically significant.

When only the total number of characters typed (including spaces) was taken into consideration, AD PE showed a significantly lower number of characters than the other two tasks: a mean total number of 885.083 characters typed against 1,520 for AD creation and 1,763.727 for AD translation. As for the rest of keystrokes, AD translation showed the lowest

number of keystrokes on average, with only 892.818, followed by AD PE (1,087.917) and AD creation (1,428.417). Even though the results for AD creation were higher, there were no statistically significant differences between any of the tasks.

Mouse actions presented quite similar mean figures: 1,473.667 for AD creation, 1,556.583 for AD PE and 1,666.545 for AD translation. In this respect, clicks and movements did not show significant differences either, but scrolls did (see *Table 1*). AD creation (23.417 scrolls on average) was statistically lower than both AD translation (65 scrolls) and post-editing (58.667 scrolls).

Concerning the number of switches from keyboard to mouse, all means ranged from 209 to 232, showing no statistically significant difference. It was in the total number of window transitions that significant differences were to be found again: AD translation presented a statistically higher number of transitions (209.727) than AD creation (99.167), but not AD PE (141). This result is in line with the distribution of time spent inside and outside Subtitle Workshop: AD creation was the task which spent proportionally more time in Subtitle Workshop and it was also the task showing the lowest amount of transitions, with the post-editing task falling between AD creation and AD translation.

Globally, in relation to technical effort, post-editing was statistically the least keyboard intensive task, with significantly the lowest number of characters typed, in accordance with O'Brien's (2010) findings. It was also the task entailing fewer mouse clicks and fewer switches from keyboard to mouse, while the rest of the values were not the highest for the three tasks in any case. All this seems to indicate that post-editing is the task involving less technical effort.

Cognitive effort

Concerning the mean total pause time, post-editing showed the lowest mean total pause time (1,394.345 seconds, i.e. 23.24 minutes), followed by AD translation (1,504.525 seconds, i.e. 25.08 minutes) and AD creation (1,625.437 seconds, i.e. 27.09 minutes). AD PE also presented the lowest mean number of pauses (961.083), although both AD creation and AD translation were not far away from that figure, presenting a very similar mean number of pauses (1,031.583 and 1,035.091, respectively). The mean time of such pauses did not differ much either: while AD PE presented the lowest mean pause time (1.505 seconds), AD translation had a mean pause time of 1.514 seconds and AD creation, of 1.724 seconds. No statistically significant differences were found in any of these items.

In connection with the pause to word ratio (PWR), AD PE showed a statistically lower mean ratio (4.081) than AD creation (6.009), but not AD translation (4.591).

It was deemed interesting to see whether the distribution between the time spent pausing and the time devoted to active writing diverged from task to task. AD creation seemed to be assigning more time to pauses (60.27%), while AD translation and post-editing devoted just a little more than half of the time to pausing (51.53% and 52.29% respectively)

(see *Figure 2*). Even though the difference was not significant, it is important to highlight that the task of creation involves more pausing than writing, which might be an indicator of a higher cognitive effort.

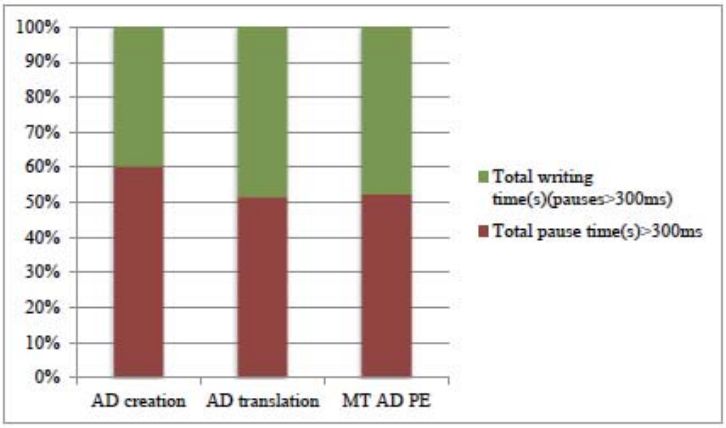

Figure 2 Distribution of pausing and writing during each task

All these data seem to indicate that post-editing was the least effort-involving task, especially if we focus on a key indicator such as PWR: AD PE presented the lowest number of pauses and the highest number of words, resulting in the lowest PWR, which is associated with low levels of cognitive effort. Conversely, AD creation seems to be the most demanding cognitively. *Table 1* presents an overview of objective results.

		AD creation	AD translation	AD post-editing
Temporal effort	total process time (seconds)	2,696.880	2,919.641	2,666.695
		(44.95 minutes)	(48.66 minutes)	(44.44 minutes)
	time spent in Subtitle Workshop (seconds)	2,415.218	2,245.303	2,238.552

		(40.25 minutes)	(37.42 minutes)	(37.31 minutes)
Technical effort	keyboard actions	2,948.417	2,656.545	1,973.000
	total number of characters typed (including spaces)	1,520.000	1,763.727	885.083
	other keystrokes	1,428.417	892.818	1,087.917
	mouse actions	1,473.667	1,666.545	1,556.583
	left clicks	615.333	616.364	567.000
	right and middle clicks	3.833	4.727	2.500
	movements	831.083	980.455	928.417
	scrolls	23.417	65.000	58.667
	switches from keyboard to mouse	231.333	223.182	209.583
	window transitions	99.167	209.727	141.000
Cognitive effort	total pause time (seconds)	1,625.437	1,504.525	1,394.345
		(27.09 minutes)	(25.08 minutes)	(23.24 minutes
	number of pauses	1,031.583	1,035.091	961.083
	mean pause time (seconds)	1.724	1.514	1.505
	pause to word ratio	6.009	4.591	4.081

Table 1 Overview of objective effort assessment results

Subjective effort and participants' opinions

Beyond objective effort, this research aimed to go a step further and gather data on participants' subjective views on effort and other relevant aspects. First of all, a comparison of the replies to the GQ, before and after the experiment, is presented, focusing first on effort and then on other items such as the degree of creativity impairment each task involves, boredom, calque conveyance, and final quality. Secondly, the participants' opinions after each task are analysed, adopting a contrastive approach where possible.

General questionnaire responses

Quantitative data on the participants' expected effort (prior to the tasks) and perceived effort (after the task) was gathered through a questionnaire, which included also an open field to justify their choices in its post-task version. Opinions on other aspects were also gathered. *Table 2* shows the means and medians obtained for each item under analysis before and after performing the tasks, on a 10-point scale where 1 is the lowest value. In the case of final quality, however, it must be clarified that "best quality" was number 1 whilst "worst quality" was number 10.

		AD creation		AD translation		AD post-editing	
		Pre	Post	Pre	Post	Pre	Post
Effort involved	Mean	8.25	7.17	6.17	5.58	6.50	7.50
	Median	8	7	6	6	6	8
Creativity impairment	Mean	3.09	3.82	7.45	7.27	8.45	9.36
	Median	3	4	8	7	9	10
Boredom	Mean	2.09	1.82	4.18	4.18	6.73	7.27
	Median	2	2	4	4	6	8
Calque conveyance	Mean	1.25	2.00	5.25	5.42	6.93	8.33
	Median	1	1.5	5	5	7	9
Final quality	Mean	1.67	2.58	2.75	3.25	4.83	5.08
	Median	2	2	2.5	3	4.5	5

Table 2 Comparison of opinions before and after the experiment

Results indicate that participants expected AD PE to be the task that would impair their creativity the most and would convey more calques. They also expected it to be the most boring task, and the one delivering the worst output quality, and involving more effort. However, in terms of effort, once the experiment was finished, both AD creation and translation were perceived as involving less effort than expected (mean=8.25 and 6.17 prior to the test to 7.17 and 5.58 after the test), while PE AD showed the opposite trend (6.50

changed into 7.50), becoming the task involving more effort according to our sample of participants. Regarding the other indicators, they all showed a clear evolution towards worse PE ratings after performing the task. This was also the case for most indicators in AD creation and AD translation, except for creativity impairment in AD translation (7.45 prior to the test, 7.27 after), and boredom in both AD translation (4.18 both prior and after the test) and AD creation (2.09 into 1.82). One possible explanation for this trend is the lack of experience of our participants.

As regards open questions that provide qualitative data, the fact that time codes were already given to participants both in the translation and post-editing tasks was often stressed as an advantage as far as effort was concerned, but the poor quality of the machine-translated text was seen as a drawback since "while a few sentences were translated correctly, most of them had mistakes or the structure needed changes" (Participant 3). Qualitative answers also reinforced the idea of post-editing being the most creativity-impairing task as it imposes "a constraint to the final text" (Participant 7). However, some participants pointed to the instruction indicating them to keep as much raw MT text as possible as the reason behind this creativity impairment rather than the actual usage of MT, which comes to show the importance and impact of instructions not only in the research arena but also in the professional world.

In connection with the degree of boredom of the tasks, responses reasserted that "[t]he AD creation task is the least boring task" (Participant 5) since "the more creative you can be, the less boring the activity will be" (Participant 7), which seems to indicate they enjoyed it more, although enjoyment was not directly assessed in the questionnaire. They also agreed in terms of conveying calques that "[i]n both the translation and the MT AD post-editing you risk to use [*sic*] calques because you do not create a new text, but depart from a source text in a foreign language" (Participant 9), and that "MT AD lacks quality because the audiodescriber [*sic*] departs from a text which is not perfectly translated" (Participant 9).

On the basis of the above, it seems that post-editing was the task involving the most subjective effort of all and presenting more drawbacks, which contrasts with objective data analysed previously.

Post-task questionnaires analysis

One of the questions included in all three post-task questionnaires assessed how easy participants felt a particular task was, immediately after performing it. Although not explicitly mentioning effort in the statement, this measure can somehow be linked to the effort participants perceived in the task. *Figure 3* shows how many participants selected a specific value on a 5-point Likert scale for each task, 1 being "strongly disagree" and 5 being "strongly agree".

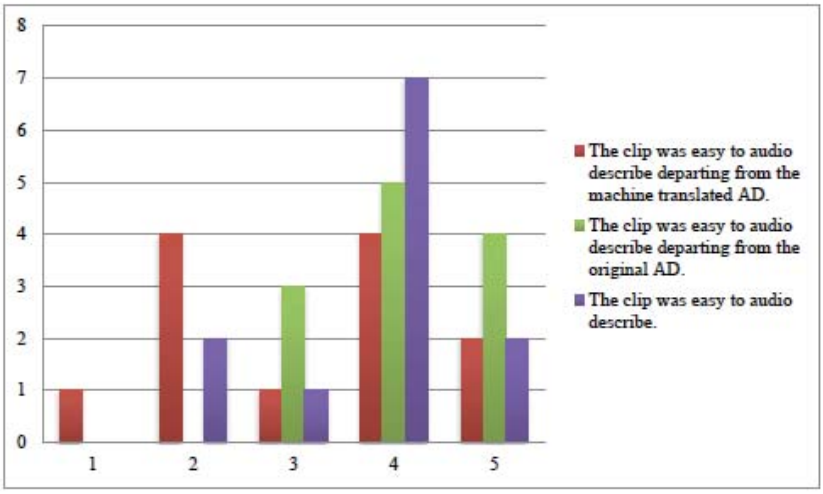

Figure 3 Self-reported ease of audio description in each scenario

The frequency chart indicates a higher variability in the answers for the post-editing task, ranging from 1 to 5 with the same number of participants selecting 2 and 4 (four participants), for instance. Regarding translation, the chosen values range from 3 to 5, showing that participants find it an easy task in a more unified way. Despite participants not having been trained in translating audio descriptions, they have a strong background and translation training, and this possibly affects the results. Finally, regarding AD creation, the vast majority selected 4 (7 out of 12), proving again that this is viewed as an easier task compared with post-editing, despite having only taken one course on audio description at MA level. When mean and median values are considered, the results are the following: AD creation (mean=3.75, median=4), AD translation (mean=4.08, median=4), AD PE (mean=3.17, median=3.5). Two additional statements looked further into the ease of the task: on the one hand, participants were asked their level of agreement with the statement "The source text was easy to translate", obtaining a mean value of 4.08 and a median of 4. When the same formulation was used for post-editing ("The machine-translated text was easy to post-edit"), the values were 2.67 and 2.5, respectively. This proves again how translation is perceived as an easier task than post-editing, at least when referring to the texts provided in the experiment. Needless to say that many factors can impact on these results: on the one hand, previous training of the participants; on the other, the quality of the machine translation output. In this regard, when participants were asked to report their level of agreement with the statement "The machine-translated text required no post-editing" on a 5-point Likert scale, values were almost the lowest possible (mean= 1.08, median=1). In response to the

sentence "The machine-translated text was fluent Catalan", the mean was 1.75 and the median was 2, and figures were slightly higher when assessing the statement "All the information in the source text was present in the machine-translated text" (mean=3.25, median=3.5).

Conclusions

This article has presented an experiment in which the efforts of creating an AD, translating an AD and post-editing a machine-translated AD were compared, with the ultimate aim of exploring whether machine translation could be satisfactorily deployed in audio description. After presenting an overview of the current state of the art, and describing the experimental design, results were discussed.

Post-editing is generally considered to be faster than human translation (Daems et al. 2015: 31), and many existing experiments prove this (de Sousa, Aziz and Specia 2011, Koglin 2015). In our test, despite being the fastest option, the differences are extremely low: on average post-editing takes only four minutes less than translating, and the difference between post-editing and creating an AD is just a few seconds. However, other indicators tend to present wider differences, and both technical and cognitive effort seem to be less demanding in post-editing. Moreover, even though no statistically significant differences were found in most cases – probably due to sample size limitations –, post-editing is usually the task displaying the most homogeneous results and, therefore, less variability, which makes the mean values obtained more reliable. It would therefore seem that implementing machine translation for audio description may be a feasible solution, or at least one which merits further investigation.

Nonetheless, if subjective effort assessments are to be considered, post-editing is generally expected to be the most demanding task in terms of effort, an idea that is reinforced once the task is performed, when the effort perception has the lowest value. This is in sharp contrast with objective data, and makes us think of the need to carry out studies which not only provide numerical data on already established indicators that can be objectively measured but also gather feed-back from users. New technological solutions cannot only be measured in terms of time or productivity, and this also applies to a possible implementation of machine translation in the audio description field.

Due to its exploratory nature, this experiment has several limitations, opening the door to further research. First of all, as already stated above – and despite it being common practice in this kind of research (Temizöz 2012) –, the small sample size does not allow statistically robust conclusions to be drawn. A larger sample would allow for sounder extrapolations to be made.

Secondly, the participants' profile has undoubtedly had an impact on the results. It was decided to use postgraduate students from the same MA programme in order to ensure a uniformly comparable sample. It remains to be seen what would happen if more experienced translators, post-editors or audio describers were selected for the test rather than AV students. One could hypothesise that time spent on the tasks by professionals compared with novices would be lower, as demonstrated by Moorkens and O'Brien (2015), but, as the same authors point out, professional attitudes towards technology may be more negative. Additionally, it would be interesting to find out whether there would be any differences between professionals with different profiles (audio describers, translators, post-editors), as it would be considerably more difficult to find professionals with completely comparable experience in these three fields.

Thirdly, evaluating the output quality, not just the process, as in this paper, would be a necessary next step. Assessing the output for the three scenarios under analysis, both by experts and by end users – mainly blind and visually impaired audiences –, would undoubtedly offer more information on this topic.

Finally, it would be worthwhile replicating the same experiment with other data sets and language pairs, to get a wider overview of the possibilities of machine translation in this new field. Many research possibilities emerge, but this paper can be considered a first step in a rather under-researched topic in the field of audiovisual translation.

Acknowledgements

We would like to thank Mãrielle Leijten and Eric Van Horenbeeck for their help and support with the keystroke-logging tool Inputlog. Special thanks are also due to the students of the MA in Audiovisual Translation who volunteered to take part in the experiments. This work was supported by the Spanish Ministerio de Ciencia e Innovación project ALST (reference code FFI2012-31024). Anna Matamala is a TransMedia Catalonia member, funded by Generalitat de Catalunya (2014SGR27).

References:

ARMSTRONG, Stephen, WAY, Andy, CAFFREY, Colm, FLANAGAN, Marian, KENNY, Dorothy, O'HAGAN, Minako. 2006. Improving the quality of automated DVD subtitles via example-based machine translation [online]. In *Proceedings of Translating and the Computer*, 2006, vol. 28 [cit. 2015-12-19], no page numbers. Available at: <http://www.mt-archive.info/Aslib-2006-Armstrong.pdf>.

CARL, Michael, DRAGSTED, Barbara, ELMING, Jakob, HARDT, Daniel, JAKOBSEN, Arnt Lykke. 2011. The Process of Post-Editing: a Pilot Study [online]. In *Proceedings of the 8th International NLPSC Workshop. Copenhagen Studies in Language*. 2011, vol. 41 [cit. 2015-12-19], pp. 131-142. Available at: <http://www.mt-archive.info/NLPCS-2011-Carl-1.pdf>.

CHUKHAREV-HUDILAINEN, Evgeny. 2014. Pauses in spontaneous written communication: A keystroke logging study [online]. In *Journal of Writing Research*. 2014, vol. 6, no. 1 [cit. 2015-12-19], pp. 61-84. Available at: <http://dx.doi.org/10.17239/jowr-2014.06.01.3>.

DAEMS, Joke, VANDEPITTE, Sonia, HARTSUIKER, Robert, MACKEN, Lieve. 2015. The impact of machine translation error types on post-editing effort indicators [online]. In *Proceedings of 4th Workshop on Post-Editing Technology and Practice (WPTP4)*. 2015 [cit. 2015-12-19], pp. 31-45. Available at: <http://amtaweb.org/wp-content/uploads/2015/10/MTSummitXV_WPTP4Proceedings.pdf>.

DAM-JENSEN, Helle, HEINE, Carmen. 2013. Writing and translation process research: Bridging the gap [online]. In *Journal of Writing Research*. 2013, vol. 5, no. 1 [cit. 2015-12-19], pp. 89-101. Available at: <http://dx.doi.org/10.17239/jowr-2013.05.01.4>.

DELGADO, Héctor, MATAMALA, Anna, SERRANO, Javier. 2015. Speaker Diarization and Speech Recognition in the Semi-Automatization of Audio Description: An Exploratory Study on Future Possibilities. In *Cadernos de Traduçao*, vol. 35, no. 2, pp. 308-324.

DEL POZO, Arantza. 2014. *SUMAT final report* [online]. 2014 [cit. 2015-06-01]. Available at: <http://www.sumat-project.eu/uploads/2014/07/D1-5_Final-Report-June-2014.pdf>

DE SOUSA, Sheila C. M., AZIZ, Wilker, SPECIA, Lucia. 2011. Assessing the post-editing effort for automatic and semi-automatic translations of DVD subtitles [online]. In *Proceedings of Recent Advances in Natural Language Processing*. 2011 [cit. 2015-12-19], pp. 97-103. Available at: <http://aclweb.org/anthology/R11-1014>.

EUROPEAN COMMISSION. n.d. MT@EC A machine translation service, covering all of the EU's official languages [online] [cit. 2015-12-19]. Available at: <http://ec.europa.eu/isa/ready-to-use-solutions/mt-ec_en.htm>.

FERNÁNDEZ-TORNÉ, Anna. Forthcoming. Machine Translation Evaluation through Post-Editing Measures in Audio Description.

FERNÁNDEZ-TORNÉ, Anna, MATAMALA, Anna. 2015. Text-to-Speech vs Human Voiced Audio Descriptions: A Reception Study in Films Dubbed into Catalan. In *The Journal of Specialised Translation* [online]. 2015, vol. 24 [cit. 2015-12-19], pp. 61-88. Available at: <http://www.jostrans.org/issue24/art_fernandez.pdf>.

GUERBEROF, Ana. 2009. Productivity and quality in MT post-editing [online]. In *MT Summit XII-Workshop: Beyond Translation Memories: New Tools for Translators MT*. 2009 [cit. 2015-12-19], no page numbers. Available at: < http://www.mt-archive.info/MTS-2009-Guerberof.pdf>.

GUERRA, Lorena. 2003. *Human translation versus machine translation and full post-editing of raw machine translation output*. MA diss. Dublin : Dublin City University, 2003.

HOUSLEY, Jason K. 2012. Ruqual: A system for assessing post-editing [online]. In *All Theses and Dissertations*. 2012 [cit. 2015-12-19], no. 3106. Available at: <http://scholarsarchive.byu.edu/cgi/viewcontent.cgi?article=4105andcontext=etd>.

JANKOWSKA, Anna. 2013. Tłumaczenie skryptów audiodeskrypcji z języka angielskiego jako alternatywna metoda tworzenia skryptów audiodeskrypcji. [Translation of audio description scripts from English as an alternative method of audio description scripts creation]. PhD diss. Cracow : Jagiellonian University, 2013.

KOGLIN, Arlene. 2015. An empirical investigation of cognitive effort required to post-edit machine translated metaphors compared to the translation of metaphors [online]. In *Translation and Interpreting*. 2015, vol. 7, no. 1 [cit. 2015-12-19], pp. 126-141. Available at: <http://www.trans-int.org/index.php/transint/issue/view/29>.

KOPONEN, Maarit. 2015. How to teach machine translation post-editing? Experiences from a post-editing course [online]. In *Proceedings of 4th Workshop on Post-Editing Technology*

and Practice (WPTP4). 2015 [cit. 2015-12-19], pp. 2-15. Available at: <http://amtaweb.org/wp-content/uploads/2015/10/MTSummitXV_WPTP4Proceedings.pdf>.

KOPONEN, Maarit, RAMOS, Luciana, AZIS, Wilker, SPECIA, Lucia. 2012. Post-Editing Time as a Measure of Cognitive Effort [online]. In *Proceedings of 1st Workshop on Post-Editing Technology and Practice (WPTP1)*. 2012 [cit. 2015-12-19], pp. 11-20. Available at: <http://amta2012.amtaweb.org/AMTA2012Files/start.htm>.

KRINGS, Hans P. 2001. *Repairing Texts: Empirical Investigations of Machine Translation Post-Editing Processes*. Kent : Kent State University Press, 2001.

LACRUZ, Isabel, DENKOWSKI, Michael, LAVIE, Alon. 2014. Cognitive Demand and Cognitive Effort in Post-Editing [online]. In *Proceedings of the 3rd Workshop on Post-Editing Technology and Practice (WPTP3)*, 2014 [cit. 2015-12-19], pp. 73-84. Available at: <http://amtaweb.org/AMTA2014Proceedings/AMTA2014Proceedings_PEWorkshop_final.pdf>.

LACRUZ, Isabel, SHREVE, Gregory M., ANGELONE, Erik. 2012. Average Pause Ratio as an Indicator of Cognitive Effort in Post-Editing: A Case Study. In *Proceedings of 1st Workshop on Post-Editing Technology and Practice (WPTP1)*. 2012 [cit. 2015-12-19], pp. 29-38. Available at: <http://amta2012.amtaweb.org/AMTA2012Files/start.htm>.

LEIJTEN, Mariëlle, VAN WAES, Luuk. 2013. Keystroke Logging in Writing Research: Using Inputlog to Analyze and Visualize Writing Processes. In *Written Communication*, vol. 30, no. 3, pp. 358-392.

MASZEROWSKA, Anna, MATAMALA, Anna, ORERO, Pilar. 2014. *Audio Description. New perspectives illustrated.* Amsterdam : John Benjamins, 2014.

MATAMALA, Anna. 2006. La accesibilidad en los medios: aspectos lingüísticos y retos de formación. In PÉREZ-AMAT, Ricardo, PÉREZ-UGENA, Álvaro. *Sociedad, integración y televisión en España*. Madrid : Laberinto, 2006, pp. 293-306.

MATAMALA, Anna. 2015. The ALST Project: Technologies for Audiovisual Translation. In *Translating and the Computer*, November 2015, vol. 37, pp. 79-89.

MOORKENS, Joss, O'BRIEN, Sharon. 2015. Post-editing evaluations: trade-offs between novice and profesional participant [online]. In *Proceedings of the 18th annual conference of the European Association for Machine Translation (EAMT 2015)*. 2015 [cit. 2015-12-19], pp. 75-81. Available at: <https://aclweb.org/anthology/W/W15/W15-4910.pdf>.

NICHOLS, Mike. 2004. [Motion picture]. *Closer*. USA : Columbia Pictures, 2004. Ident. No. CDR 37281.

O'BRIEN, Sharon. 2004. Machine translatability and post-editing effort: How do they relate? In *Translating and the Computer,* November 2004, vol. 26, no page numbers.

O'BRIEN, Sharon. 2006. Pauses as indicators of cognitive effort in post-editing machine translating output. In *Across Languages and Cultures*, 2006, vol. 7, no. 1, pp. 1-21.

O'BRIEN, Sharon. 2009. Eye tracking in translation process research: methodological challenges and solutions. In MEES, Inger M., ALVES, Fabio, GOPFERICH, Susanne, (eds.) *Methodology, technology and innovation in translation process research: a tribute to Arnt Lykke Jakobsen. Copenhagen studies in language*. Copenhagen : Samfundslitteratur, 2009, no. 38, pp. 251-266.

O'BRIEN, Sharon. 2010. Introduction to post-editing: Who, what, how and where to next [online]. In *Association for Machine Translation in the Americas AMTA 2010*. 2010 [cit. 2015-12-19], no page numbers. Available at: <http://amta2010.amtaweb.org/AMTA/papers/6-01-ObrienPostEdit.pdf>.

O'BRIEN, Sharon. 2011. Towards predicting post-editing productivity [online]. In *Machine Translation*. 2011, vol. 25 [cit. 2015-12-19], pp. 197-215. Available at: <http://doras.dcu.ie/17154/1/Towards_Predicting_Postediting_Productivity_Final_2.pdf>.

ORTIZ-BOIX, Carla. 2012. *Technologies for audio description: study on the application of machine translation and text-to-speech to the audio description in Spanish*, MA diss. Barcelona : Universitat Autònoma de Barcelona, 2012.

ORTIZ-BOIX, Carla, MATAMALA, Anna. 2015. Quality Assessment of Post-Edited versus Translated Wildlife Documentary Films: A Three-Level Approach. [online]. In *Proceedings*

of 4th Workshop on Post-Editing Technology and Practice (WPTP4). 2015 [cit. 2015-12-19], pp. 2-15. Available at: <http://amtaweb.org/wp-content/uploads/2015/10/MTSummitXV_WPTP4Proceedings.pdf>.

ORTIZ-BOIX, Carla, MATAMALA, Anna. Forthcoming. Post-editing Wildlife Documentary Films: A New Possible Scenario? In *Jostrans*, vol. 26.

PLITT, Mirko, MASSELOT, François. 2010. A Productivity Test of Statistical Machine Translation Post-Editing in a Typical Localisation Context. In *The Prague Bulletin of Mathematical Linguistics,* vol. 93, pp. 7-16.

REMAEL, Aline, VERCAUTEREN, Gert. 2007. Audio describing the exposition phase of films. Teaching students what to choose. In *TRANS*, 2007, vol. 11, pp. 73-93.

SALWAY, Andrew. 2004. AuDesc system specification and prototypes, TIWO: Television in Words [online]. 2004, vol. 3 [cit. 2015-12-19]. Available at: <http://www.bbrel.co.uk/pdfs/TIWO_Television_in_Words_Deliverable_3.pdf>.

SPECIA, Lucia. 2011. Exploiting objective annotations for measuring translation post-editing effort [online]. In *Conference of the European Association for Machine Translation*. 2011, vol. 15 [cit. 2015-12-19], pp. 73-80. Available at: <http://www.mt-archive.info/EAMT-2011-Specia.pdf>.

TATSUMI, Midori, ROTURIER, Johann. 2010. Source Text Characteristics and Technical and Temporal Post-Editing Effort: What is Their Relationship? [online]. In *Proceedings of the Second Joint EM+/CNGL Workshop "Bringing MT to the User: Research on Integrating MT in the Translation Industry"*. 2010 [cit. 2015-12-19], pp. 43-51. Available at: <http://www.mt-archive.info/JEC-2010-Tatsumi.pdf>.

TAUS, CNGL. 2010. *Machine translation postediting guidelines* [online]. 2010 [cit. 2015-12-19]. Available at: <http://taus-website-media.s3.amazonaws.com/images/stories/guidelines/taus-cngl-machine-translation-postediting-guidelines.pdf>.

TEMIZÖZ, Özlem. 2012. Machine translation and postediting. In *The European Society for Translation Studies Research Committee State-of-the-Art Research Report*.

VOLK, Martin. 2009. The automatic translation of film subtitles. A machine translation success story? [online[. In *JLCL*, 2009, vol. 24, no. 3 [cit. 2015-12-19], pp. 113-125. Available at: <http://www.uzh.ch/news/articles/2008/3000/Volk__MT_of_Subtitles.pdf>.

Anna Fernández-Torné

Department of Translation and Interpreting and East Asian Studies, Universitat Autònoma de Barcelona

La Vinya, 11B

08329 Teià (Barcelona)

Spain

e-mail: anna.torne@gmail.com

Anna Matamala

Department of Translation and Interpreting and East Asian Studies, Universitat Autònoma de Barcelona

Despatx K-1002, Campus de la UAB

08193 Bellaterra (Barcelona)

Spain

e-mail: anna.matamala@uab.cat

In SKASE Journal of Translation and Interpretation [online]. 2015, vol. 9, no. 1 [cit. 2016-18-04]. Available online <http://www.skase.sk/Volumes/JTI10/pdf_doc/05.pdf>. ISSN 1336-7811

Annex 2: Documents related to the TTS pre-test for the selection of best-rated voices

Annex 2.1. Cold calling email for the recruitment of volunteer voice talents

Benvolguts:

Em dic Anna Fernández Torné i sóc doctoranda del Centre d'Accessibilitat i Intel·ligència Artificial de Catalunya, de la Universitat Autònoma de Barcelona. Necessito col·laboradors voluntaris per a la realització d'un experiment per a la meva tesi, centrada en l'estudi de l'audiodescripció i les tecnologies. Es tracta simplement de locutar 5 frases.

Els qui estigueu interessats a participar en l'experiment heu de complir els següents requisits:

- Tenir el català com a llengua materna (dialecte central)

- No tenir deficiències en la parla

La gravació es durà a terme a l'ECAD i la data es convindrà en funció de la disponibilitat de sales. Si hi voleu participar, em podeu enviar un missatge a la següent adreça de correu electrònic ana.fernandez.torne@uab.cat.

Si coneixeu algú altre que creieu que hi pugui estar interessat, no dubteu a remetre-li aquest missatge o a facilitar-li la meva adreça de correu electrònic perquè es posi en contacte amb mi.

Gràcies!

English translation

Dear all,

My name is Anna Fernández Torné and I am a PhD student of the Accessibility and Artificial Intelligence Centre of Catalonia, of the Universitat Autònoma de Barcelona. I am in need of recruiting volunteers to carry out an experiment for my thesis, focused on the study of audio description and technologies. Only 5 sentences should be voiced.

Those who are interested in taking part in the experiment must meet the following requirements:

- Be Catalan native speakers (central dialect)

- Have no speech deficiencies

The recording will take place at the ECAD premises and the date will be agreed on depending on the room availability. If you would like to participate, you can send me a message to the following email address: ana.fernandez.torne@uab.cat.

If you know any other person that you think might be interested in taking part, do not hesitate to forward them this message or to give them my email address so that they can contact me.

Thank you!

Annex 2.2. AD unit selection

Number of characters	Number of words	Sample 1 AD units
13	3	Fosa a blanc.
224	42	Avança fins a l'escenari següent, on tres noies ballen al voltant d'una barra vertical. Mentre una penja cap per avall de la barra, una altra s'ajup amb el cul enfora i un client li posa un bitllet sota la cintura del tanga.
71	12	Ella el segueix a l'interior de l'edifici i l'escodrinya amb la mirada.
47	11	Se li acosta i li arregla el nus de la corbata.
116	24	El noi mig somriu i s'hi encamina mentre ella ho desa somrient al maletí i en treu una poma. Té una ferida al front.
471	**92**	
		Sample 2 AD units
13	3	Fosa a negre.
230	46	Asseguda en una cadira, la noia obre el maletí. En treu un full de diari doblegat, hi fa un cop d'ull i el torna al seu lloc. Mentre ell l'observa seriós des del llindar, en treu un embolcall de paper d'alumini i hi fa una ullada.
71	13	En veure-la despenjar el telèfon, salta per sobre del llit i l'hi pren.
47	11	Li petoneja les cames i se li estira a la vora.
116	23	Ella es tomba i fixa la vista a l'altra banda de la finestra. L'autobús, vermell i de dos pisos, passa sota un pont.
477	**96**	

Annex 2.3. AD units order

A (Acapela Laia / Marta Serratosa)	1	Fosa a blanc.
	2	Avança fins a l'escenari següent, on tres noies ballen al voltant d'una barra vertical. Mentre una penja cap per avall de la barra, una altra s'ajup amb el cul enfora i un client li posa un bitllet sota la cintura del tanga.
	3	Ella el segueix a l'interior de l'edifici i l'escodrinya amb la mirada.
	4	Se li acosta i li arregla el nus de la corbata.
	5	El noi mig somriu i s'hi encamina mentre ella ho desa somrient al maletí i en treu una poma. Té una ferida al front.
B (iSpeech Anna / Sílvia Majó)	2	Avança fins a l'escenari següent, on tres noies ballen al voltant d'una barra vertical. Mentre una penja cap per avall de la barra, una altra s'ajup amb el cul enfora i un client li posa un bitllet sota la cintura del tanga.
	3	Ella el segueix a l'interior de l'edifici i l'escodrinya amb la mirada.
	4	Se li acosta i li arregla el nus de la corbata.
	5	El noi mig somriu i s'hi encamina mentre ella ho desa somrient al maletí i en treu una poma. Té una ferida al front.
	1	Fosa a blanc.
C (Verbio Meritxell / Anna Romagosa)	5	El noi mig somriu i s'hi encamina mentre ella ho desa somrient al maletí i en treu una poma. Té una ferida al front.
	4	Se li acosta i li arregla el nus de la corbata.
	1	Fosa a blanc.
	2	Avança fins a l'escenari següent, on tres noies ballen al voltant d'una barra vertical. Mentre una penja cap per avall de la barra, una altra s'ajup amb el cul enfora i un client li posa un bitllet sota la cintura del tanga.
	3	Ella el segueix a l'interior de l'edifici i l'escodrinya amb la mirada.
D (Loquendo Montserrat / Belén Roca)	4	Se li acosta i li arregla el nus de la corbata.
	1	Fosa a blanc.
	5	El noi mig somriu i s'hi encamina mentre ella ho desa somrient al maletí i en treu una poma. Té una ferida al front.
	3	Ella el segueix a l'interior de l'edifici i l'escodrinya amb la mirada.
	2	Avança fins a l'escenari següent, on tres noies ballen al voltant d'una barra vertical. Mentre una penja cap per avall de la barra, una altra s'ajup amb el cul enfora i un client li posa un bitllet sota la cintura del tanga.
E (Talp Ona / Clàudia Carreras)	3	Ella el segueix a l'interior de l'edifici i l'escodrinya amb la mirada.
	5	El noi mig somriu i s'hi encamina mentre ella ho desa somrient al maletí i en treu una poma. Té una ferida al front.
	2	Avança fins a l'escenari següent, on tres noies ballen al voltant d'una barra vertical. Mentre una penja cap per avall de la barra, una altra s'ajup amb el cul enfora i un client li posa un bitllet sota la cintura del tanga.
	1	Fosa a blanc.
	4	Se li acosta i li arregla el nus de la corbata.

F (Talp Jan / Albert Martínez)	5	Ella es tomba i fixa la vista a l'altra banda de la finestra. L'autobús, vermell i de dos pisos, passa sota un pont.
	4	Li petoneja les cames i se li estira a la vora.
	3	En veure-la despenjar el telèfon, salta per sobre del llit i l'hi pren.
	2	Asseguda en una cadira, la noia obre el maletí. En treu un full de diari doblegat, hi fa un cop d'ull i el torna al seu lloc. Mentre ell l'observa seriós des del llindar, en treu un embolcall de paper d'alumini i hi fa una ullada.
	1	Fosa a negre.
G (Talp Teo / David Garcia)	4	Li petoneja les cames i se li estira a la vora.
	1	Fosa a negre.
	2	Asseguda en una cadira, la noia obre el maletí. En treu un full de diari doblegat, hi fa un cop d'ull i el torna al seu lloc. Mentre ell l'observa seriós des del llindar, en treu un embolcall de paper d'alumini i hi fa una ullada.
	5	Ella es tomba i fixa la vista a l'altra banda de la finestra. L'autobús, vermell i de dos pisos, passa sota un pont.
	3	En veure-la despenjar el telèfon, salta per sobre del llit i l'hi pren.
H (Verbio Oriol / Arià Paco)	3	En veure-la despenjar el telèfon, salta per sobre del llit i l'hi pren.
	5	Ella es tomba i fixa la vista a l'altra banda de la finestra. L'autobús, vermell i de dos pisos, passa sota un pont.
	4	Li petoneja les cames i se li estira a la vora.
	1	Fosa a negre.
	2	Asseguda en una cadira, la noia obre el maletí. En treu un full de diari doblegat, hi fa un cop d'ull i el torna al seu lloc. Mentre ell l'observa seriós des del llindar, en treu un embolcall de paper d'alumini i hi fa una ullada.
I (Loquendo Jordi / Dani Albiac)	1	Fosa a negre.
	2	Asseguda en una cadira, la noia obre el maletí. En treu un full de diari doblegat, hi fa un cop d'ull i el torna al seu lloc. Mentre ell l'observa seriós des del llindar, en treu un embolcall de paper d'alumini i hi fa una ullada.
	5	Ella es tomba i fixa la vista a l'altra banda de la finestra. L'autobús, vermell i de dos pisos, passa sota un pont.
	3	En veure-la despenjar el telèfon, salta per sobre del llit i l'hi pren.
	4	Li petoneja les cames i se li estira a la vora.
J (Talp Pep / Albert Llort)	2	Asseguda en una cadira, la noia obre el maletí. En treu un full de diari doblegat, hi fa un cop d'ull i el torna al seu lloc. Mentre ell l'observa seriós des del llindar, en treu un embolcall de paper d'alumini i hi fa una ullada.
	3	En veure-la despenjar el telèfon, salta per sobre del llit i l'hi pren.
	1	Fosa a negre.
	4	Li petoneja les cames i se li estira a la vora.
	5	Ella es tomba i fixa la vista a l'altra banda de la finestra. L'autobús, vermell i de dos pisos, passa sota un pont.

Annex 2.4. Listening order of the speech samples

Participant								
1	A	F	B	G	C	H	D	I
2	B	G	C	H	D	I	E	J
3	C	H	D	I	E	J	A	F
4	D	I	E	J	A	F	B	G
5	E	J	A	F	B	G	C	H
6	E	J	D	I	C	H	B	G
7	D	I	C	H	B	G	A	F
8	C	H	B	G	A	F	E	J
9	B	G	A	F	E	J	D	I
10	A	F	E	J	D	I	C	H
11	F	A	G	B	H	C	I	D
12	G	B	H	C	I	D	J	E
13	H	C	I	D	J	E	F	A
14	I	D	J	E	F	A	G	B
15	J	E	F	A	G	B	H	C
16	J	E	I	D	H	C	G	B
17	I	D	H	C	G	B	F	A
18	H	C	G	B	F	A	J	E
19	G	B	F	A	J	E	I	D
20	F	A	J	E	I	D	H	C

Annex 2.5. Voice samples' questionnaire

Impressió general
Com descriuries la qualitat de la veu que acabes de sentir?
1. Dolenta
2. Regular
3. Normal
4. Bona
5. Excel·lent

Accentuació
Fins a quin punt consideres que és agradable aquesta veu?
1. Sí, moltes
2. Sí, bastantes
3. Sí, algunes
4. Sí, però poques
5. No, cap

Pronúncia
Has percebut anomalies pel que fa a la pronúncia?
1. Sí, moltes
2. Sí, bastantes
3. Sí, algunes
4. Sí, però poques
5. No, cap

Pauses discursives
Creus que la veu fa pauses quan cal?
1. No, mai
2. No, gairebé mai
3. Sí, normalment
4. Sí, gairebé sempre
5. Sí, sempre

Entonació
Com qualificaries l'entonació de les frases?
1. Molt dolenta
2. Dolenta
3. Bona
4. Bastant bona
5. Molt bona

Naturalitat
Com definiries el grau de naturalitat de la veu?
1. No gens natural
2. Poc natural
3. Natural
4. Bastant natural
5. Molt natural

Agradabilitat
Fins a quin punt consideres que és agradable aquesta veu?
1. Molt desagradable
2. Desagradable
3. Normal
4. Agradable
5. Molt agradable

Esforç d'escolta
Creus que et cansaria escoltar aquesta veu durant molt de temps seguit?
1. Sí, molt
2. Sí, bastant
3. Sí, una mica
4. No, no gaire
5. No, no gens

Acceptació
Creus que aquesta veu es podria fer servir per locutar audiodescripcions?
1. No, mai
2. No, gairebé mai
3. Sí, en alguns casos
4. Sí, en bastants casos
5. Sí, sempre

English translation

How would you describe the quality of the voice you have just heard?
1. Bad
2. Regular
3. Neutral
4. Good
5. Excellent

Did you detect anomalies in terms of the accentuation of words?
1. Yes, al lot of them
2. Yes, many
3. Yes, some
4. Yes, but only a few
5. No, none

Did you notice anomalies in terms of pronunciation?
1. Yes, al lot of them
2. Yes, many
3. Yes, some
4. Yes, but only a few
5. No, none

Do you think the voice makes pauses when it is needed?
1. No, never
2. No, almost never
3. Yes, normally
4. Yes, almost always
5. Yes, always

How would you rate the intonation of sentences?
1. Very bad
2. Bad
3. Good
4. Quite good
5. Very good

How would you define the degree of naturalness of the voice?
1. Very unnatural
2. Unnatural
3. Natural
4. Quite natural
5. Very natural

To what extent do you deem this voice to be pleasant?
1. Very unpleasant
2. Unpleasant
3. Neutral
4. Pleasant
5. Very pleasant

Do you think listening to this voice for a long time would be tiring?
1. Yes, a lot
2. Yes, quite a lot
3. Yes, a little bit
4. No, not much
5. No, not at all

Do you think this voice could be used for voicing audio descriptions?
1. No, never
2. No, almost never
3. Yes, in some cases
4. Yes, in many cases
5. Yes, always

Annex 2.6. Cold-calling email for the recruitment of volunteer participants

Benvolguts:

Em dic Anna Fernández Torné i necessito col·laboradors voluntaris per a la realització d'un experiment per a la meva tesi, centrada en l'estudi de l'audiodescripció i les tecnologies. Es tracta simplement d'escoltar 20 mostres de veu, 10 de femenines i 10 de masculines, i de respondre un qüestionari per avaluar cadascuna de les veus. L'experiment es farà en dues sessions diferents, d'uns 45 minuts cadascuna. Hi haurà un espai de temps de 2 dies com a mínim entre cada sessió.

Els qui estigueu interessats a participar en l'experiment heu de complir els següents requisits:

- Ser vidents

- Ser catalanoparlants

- No tenir deficiències auditives

- No tenir experiència en síntesi de parla (veus artificials)

- No tenir experiència en audiodescripció (no haver vist mai cap producte audiovisual audiodescrit)

Si hi voleu participar, em podeu enviar un missatge a la següent adreça de correu electrònic ana.fernandez.torne@uab.cat i mirarem de concretar l'hora que us vagi millor.

Si coneixeu algú que creieu que hi pugui estar interessat, no dubteu a remetre-li aquest missatge o a facilitar-li la meva adreça de correu electrònic perquè es posi en contacte amb mi.

Gràcies!

English translation

Dear all,

My name is Anna Fernández Torné and I am in need of recruiting volunteers to carry out an experiment for my thesis, focused on the study of audio description and technologies. You should just listen to 20 voice samples, 10 feminine and 10 masculine, and answer a questionnaire to assess each voice. The experiment will be carried out in two different sessions, around 45 minutes each. There will be at least a lapse of time of two days between both sessions.

Those who are interested in taking part in the experiment must meet the following requirements:

- Be sighted

- Be Catalan speakers

- Have no hearing impairments

- Have no experience in speech synthesis (artificial voices)

- Have no experience in audio description (not having seen any audio described audiovisual product)

If you would like to participate, you can send me a message to the following email address ana.fernandez.torne@uab.cat and we will try to schedule a time that suits you better.

If you know any other person that you think might be interested in taking part, do not hesitate to forward them this message or to give them my email address so that they can contact me.

Thank you!

Annex 2.7. Instructions for the TTS pre-test performance

CONTEXTUALITZACIÓ I INSTRUCCIONS PER ALS PARTICIPANTS

Com ja sabeu, em dic Anna Fernández Torné i sóc doctoranda del Centre d'Accessibilitat i Intel·ligència Artificial de Catalunya, de la Universitat Autònoma de Barcelona. Sou voluntaris per a un dels experiments que formen part de la meva tesi, centrada en l'estudi de l'audiodescripció i les tecnologies.

L'audiodescripció és la tècnica que s'utilitza per convertir la informació visual (les imatges) en informació oral perquè les persones cegues i amb deficiències visuals puguin tenir accés al que passa en pantalla. Es tracta de descriure el que passa en els espais de silenci, quan els personatges no parlen. La tecnologia que s'hi pretén aplicar és, en aquest cas, la síntesi de parla. Per síntesi de parla entenem la producció artificial de parla humana mitjançant veus artificials.

Aquest experiment en concret pretén avaluar la qualitat d'una sèrie de veus sintètiques i naturals que es volen aplicar a l'audiodescripció de pel·lícules doblades al català. És per això que necessitem la vostra ajuda. L'únic que s'espera de vosaltres és que escolteu atentament aquestes veus, que les avalueu (només les veus, deixant de banda el contingut, el que s'hi diu) i que respongueu els qüestionaris corresponents per donar-nos-en la vostra opinió. És important que tingueu en compte que no heu de comparar les veus entre elles, sinó que les heu d'avaluar una per una, independentment de les altres. Convé, a més, que no tingueu por de puntuar als extrems. Val més que quedi clara la vostra opinió que no pas que tot siguin puntuacions neutres.

La prova consta de dues parts. En la primera part sentireu dues mostres de veu d'exemple, una de masculina i una de femenina, i us familiaritzareu amb les preguntes i les respostes del qüestionari. Quan hagueu sentit i avaluat les dues veus d'exemple, hi haurà una pausa perquè pugueu plantejar-nos els dubtes que us hagin sorgit. Després, un cop resolts els dubtes, començarà la segona part, el qüestionari real.

Les mostres de veu estan formades per frases inconnexes. Haureu de parar-hi atenció i escoltar-les atentament. Sentireu cada mostra dos cops. Un cop hagueu sentit cada mostra, caldrà que respongueu un qüestionari de 9 preguntes. Les respostes a aquestes preguntes es presenten amb escales de l'1 al 5, on 1 és la valoració més negativa i 5, la més positiva.

Haureu de marcar la resposta que voleu donar amb el ratolí i pitjar tot seguit el botó Continuar per passar a la següent pregunta. Un cop contestada la pregunta, no podreu anar enrere per canviar la resposta, encara que al formulari veureu que hi apareix el botó Enrere.

Les preguntes seran sempre les mateixes per a totes les mostres de veu i es corresponen amb els següents aspectes:

- Impressió global: doneu una puntuació de conjunt de la veu, l'opinió general.
- Accentuació: tingueu en compte si es pronuncien les paraules amb l'accent allà on toca (si fa "rèptil" i no "reptil"), és a dir, si la síl·laba tònica de cada paraula es fa correctament.
- Pronúncia: avalueu si la veu distingeix bé els sons que formen cada paraula, si es fan correctament d'acord amb la fonètica catalana (si diu "cos" i no "cos", "passa" i no "pasa", "ètic" i no "étic", entre d'altres) .
- Pauses discursives: considereu si la veu s'atura quan fa falta entre els components de la frase i entre frases, si no sobta perquè para en llocs no habituals en català, com ara "La... casa vermella està en... runes".
- Entonació: sospeseu fins a quin punt la veu dibuixa la corba melòdica que pertoca a la frase que s'està dient, si queda clar quan és una pregunta o no, per exemple.
- Naturalitat: en el cas de les veus sintètiques, indiqueu fins a quin punt la veu s'assembla a una veu humana; en el cas de les veus naturals, indiqueu fins a quin punt la veu és forçada, artificial.
- Agradabilitat: indiqueu si la veu és agradable.

- Esforç d'escolta: responeu fins a quin punt us resultaria pesat, avorrit, sentir la veu en la situació hipotètica d'haver de veure una pel·lícula sencera audiodescrita amb aquesta veu.
- Acceptació: indiqueu si la veu és adequada per fer audiodescripcions.

Com veieu, primer se us presenta una pregunta de caràcter global (la impressió global), després dues preguntes centrades en les paraules (l'accentuació i la pronúncia), després dues més de centrades més aviat en les frases (les pauses discursives i l'entonació), a continuació tres de relacionades directament amb la veu (la naturalitat, l'agradabilitat i l'esforç d'escolta) i finalment una de relacionada amb l'aplicació concreta de l'audiodescripció (l'acceptació).

Haureu d'avaluar 10 mostres de veu en total, 5 de femenines i 5 de masculines, alternades, i haureu d'emplenar un qüestionari per a cada veu.

La participació en aquest experiment és del tot voluntària i, tot i que us demanarem algunes dades personals, com ara l'edat, els estudis previs i la professió, la informació es tractarà amb la màxima confidencialitat i les dades es processaran de forma anònima. Al davant hi teniu un full informatiu i de consentiment. Sisplau, llegiu-lo, empleneu-lo, signeu-lo si hi esteu d'acord i me'l doneu.

Ara, veureu que a la pantalla hi teniu oberts el navegador i el programa de reproducció de so VLC. Primer haureu d'emplenar el primer full que us apareix al navegador, el de DADES PERSONALS. Ompliu-lo, sisplau.

Si ja l'heu omplert, podeu tancar la pestanya AIXÍ.

Ara, haureu d'anar obrint les altres pestanyes en l'ordre exacte en què apareixen en pantalla (d'esquerra a dreta) i emplenant només els vostres identificadors. Quan hagueu emplenat l'identificador, feu clic a Continuar i no toqueu res més, i continueu amb el següent, però sense tancar la pestanya.

Ara, comproveu que teniu oberta la pestanya del navegador Exemple de veu masculina sintètica. Bé.

A mesura que sentiu les preguntes, marqueu la resposta i feu clic amb el ratolí al botó Continuar. Quan hagueu acabat de contestar totes les preguntes, feu clic al botó Enviar i tanqueu la pestanya del navegador.

Provem-ho amb els dos exemples i m'aviseu si teniu cap dubte.

Comencem.

English translation

CONTEXTUALITSATION AND INSTRUCTIONS FOR THE PARTICIPANTS

As you already know, my name is Anna Fernández Torné and I am a PhD student of the Accessibility and Artificial Intelligence Centre of Catalonia, of the Universitat Autònoma de Barcelona. You are volunteers for one of the experiments that constitute my thesis, focused on the study of audio description and technologies.

Audio description is the technique used to turn visual information (images) into aural information so that blind and partially sighted persons may have access to what is happening on screen. It involves describing what is happening in the silent gaps, when characters are not talking. The technology to be applied is, in this case, speech synthesis. We consider speech synthesis as the artificial production of human speech by means of artificial voices.

This experiment in particular aims at assessing the quality of a selection of synthetic and natural voices that are to be applied to the audio description of dubbed films into Catalan. That is why we need your help. The only thing you are expected to do is to listen attentively to these voices, to assess them (only the voices, thus neglecting the

contents, what is actually being said) and to answer the corresponding questionnaires to give us your opinion. It is important to bear in mind that you do not have to compare the voices among them, but to assess them individually, separately from the others. Additionally, it is desirable that you are not afraid of giving extreme answers. It is preferable that your opinion is stated clearly than having only neutral ratings.

The test is divided into two parts. In the first part, you will listen to two voice samples as an example, one masculine and one feminine, and you will familiarise with the questions and answers of the questionnaire. Once you have listened to and assessed the two example voices, there will be a pause in order to raise any concerns you may have. Afterwards, once concerns have been cleared up, the second part of the test will begin, the real questionnaire.

Voice samples are made up of disjointed sentences. You will have to pay attention and listen to them attentively. You will listen to each sample twice. Once you have listened to each sample, you will have to answer a 9-question questionnaire. Answers to these questions are presented in scales from 1 to 5, where 1 is the most negative rating and 5 is the most positive one.

You will have to click on the desired answer with the mouse and press the Continue button to go to the following question. Once the question has been answered, you will not be allowed to go back to change it, even though you will see in the form that there is a Back button.

Questions will always be the same for all voice samples and they refer to the following aspects:

- Overall impression: give a global score for the voice, your general opinion.
- Accentuation: consider whether words are pronounced placing the stress on the right syllable (whether “rèptil” is pronounced instead of “reptil”), i.e. if the accented syllable is the one it should be.

- Pronunciation: assess whether the voice clearly distinguishes the sounds that form each word, if they are properly pronounced according to Catalan phonetics (if “cos” is said instead of “cós”, “passa” instead of “pasa”, “ètic” instead of “étic”, among others).
- Speech pauses: consider whether the voice stops when needed between sentence components and between sentences, if it is not shocking because it stops in uncommon places in Catalan, such as “La... casa vermella està en... runes”.
- Intonation: assess whether the pitch curve accurately represents the sentence it is being said, if it is clear to you whether it is a question or not, for example.
- Naturalness: in synthetic voices, indicate to which extent the voice resembles a human voice; in natural voices, indicate to which extent the human voice is forced, artificial.
- Pleasantness: indicate whether the voice is pleasant.
- Listening effort: assess to which extent listening to the voice would be tiring or tedious in case of having to watch an entire film audio described with this voice.
- Acceptance: indicate if the voice is adequate to voice audio descriptions.

As you can see, you are first presented with a global question (overall impression), followed by two questions focused on words (accentuation and pronunciation), followed by two others focused on sentences (speech pauses and intonation), then three questions directly related to the voice (naturalness, pleasantness and listening effort) and finally one question regarding the particular application of audio description (acceptance).

You will have to assess a total of 10 voice samples, 5 feminine and 5 masculine, alternating feminine and masculine, and you will have to fill in one questionnaire per voice.

Your participation on this study is on a voluntary basis and, even though we will ask for some personal details, such as your age, education and profession, the information will be treated with utmost confidentiality and data will be processed anonymously. The printed information sheet and consent form is placed in front of you. Please, read it, fill it in, sign it if you agree and give it back to me.

You will now see on screen that the web browser and the VLC player are open. You will first need to fill in the first tab appearing on the browser, the PERSONAL DETAILS form. Please, fill it in.

If you already have, you can close the tab LIKE THIS.

You will now need to open the other tabs in the exact order they appear on screen (from left to right) filling in just your identifiers. Once you have filled in the identifier, click on Continue and do not fill in anything else, and continue with the following identifier without closing the tab.

Now, check that the open tab on the browser is Example of masculine synthetic voice. Good.

As you listen to the questions, click on the desired answer with the mouse and then click on Continue. Once you have answered all questions, click on the Send button and close the browser tab.

Let’s try with the two examples and ask any questions to me if needed.

Let’s start.

Annex 2.8. Information sheet

FULL D'INFORMACIÓ SOBRE L'EXPERIMENT

L'objectiu de l'estudi és avaluar l'acceptació de diverses veus en l'àmbit de l'audiodescripció en català.

La vostra participació en l'experiment serà la següent: haureu d'avaluar 20 mostres de veu mitjançant qüestionaris seguint les instruccions que us donarem.

La vostra participació és totalment voluntària i us podeu retirar de l'estudi en qualsevol moment sense haver-ho de justificar de cap manera ni sense que això us afecti negativament de cap manera.

Les vostres dades seran totalment anònimes ja que al full escrit només us demanarem que hi escriviu el vostre sexe (home/dona), l'edat, els estudis previs i la professió.

No rebreu cap compensació econòmica per aquesta participació, la qual en cap cas les farà servir amb ànim de lucre.

Les investigadores que duran a terme l'experiment són l'Anna Fernández Torné (que participarà en les proves) i la Dra. Anna Matamala (que dirigeix la investigació). Aquesta darrera és la responsable i la podeu trobar a anna.matamala@uab.cat, al despatx K-2012 de la UAB o bé al telèfon 935811646.

Si voleu continuar informats de les evolucions de la investigació, només cal que ho demaneu a la investigadora principal al correu electrònic anterior.

MOLTES GRÀCIES per la vostra participació!

(Nota: a l'altra cara d'aquest full hi teniu el document de consentiment informat)

English translation

INFORMATION SHEET

The aim of this study is to assess the acceptance of several Catalan voices in the audio description field.

Your participation in the experiment will be as follows: you will have to assess 20 voice samples by means of questionnaires following the instructions we will give to you.

Your participation is absolutely voluntary and you may withdraw from the study at any time without further justification and without that negatively affecting you in any way.

Your personal details will be completely anonymous since in the written form you will only be asked to indicate your gender (man/woman), your age, your education and your profession.

You will receive no economic compensation for your participation, which will not be used for profit under any circumstance.

The researchers who will conduct the experiment are Anna Fernández Torné (who will take part in the tests) and Dr. Anna Matamala, who is the supervisor and responsible for this research. You can contact her at anna.matamala@uab.cat, or by phone at 935811646. She can also be found at the office K-2012 in the UAB.

If you would like to be informed of the research progress, please ask the principal investigator by contacting her at the previous email address.

Thank you very much for your participation!

(Note: you will find the informed consent form on the other side of this page)

Annex 2.9. Informed consent form

CONSENTIMENT INFORMAT

Títol de l'experiment: Avaluació subjectiva de veus per a l'audiodescripció

Jo, (NOM I COGNOMS) _____

- He llegit el full d'informació que m'han donat com també el de consentiment informat.
- He rebut prou informació sobre l'estudi i l'he entesa.
- He pogut fer preguntes sobre l'estudi.

He parlat amb: **Anna Fernández Torné**

- Entenc que la meva participació és voluntària
- Entenc que la meva participació és no remunerada
- Entenc que la meva informació serà confidencial
- Entenc que em puc retirar de l'estudi quan vulgui i sense haver de donar explicacions ni sense que hi hagi repercussions negatives.

Dono lliurament la meva conformitat per poder participar en l'estudi.

Data

Firma del participant

(Nota: a l'altra cara d'aquest full hi teniu el document d'informació)

English translation

INFORMED CONSENT

Title of the experiment: Subjective assessment of voices for the audio description

(NAME AND SURNAME) _____

- I have read the information sheet I have been given, as well as the informed consent form.
- I have received enough information about the study and I have understood it.
- I have been able to make questions about the study.

I have talked to: **Anna Fernández Torné**

- I understand that my participation is voluntary
- I understand that my participation is unpaid
- I understand that my personal details will be treated confidentially
- I understand that I can withdraw from the study whenever I want giving no particular reason for it and having no negative repercussions.

I freely consent to participate on this research study.

Date

Signature of participant

(Note: you will find the information sheet on the other side of this page)

Annex 2.10. Personal details form

Dades personals

* Necessari

Identificador *

Introdueix les teves inicials (la del nom, la del primer cognom i la del segon cognom) seguit de la teva edat. No hi deixis cap espai entremig.

Edat *

Sexe *

- o Home
- o Dona

Nivell d'estudis assolit *

- o Inferiors a la primera etapa d'educació secundària
- o Educació secundària, primer cicle
- o Segona etapa d'educació secundària
- o Formació professional de grau superior
- o Estudis universitaris de primer cicle (diplomatura, llicenciatura, enginyeria o grau)
- o Estudis universitaris de segon cicle (màster, postgrau o doctorat)

En cas d'haver cursat estudis superiors, digues quins.

Nivell professional *

- o Direcció de l'administració pública i de les empreses de 10 o més assalariats
- o Gerència d'empreses amb menys de 10 assalariats
- o Gerència d'empreses sense assalariats
- o Professions associades a titulació de segon i tercer cicle universitari i afins
- o Professions associades a titulació de primer cicle universitari i afins
- o Tècnics i professionals de suport
- o Empleats de tipus administratiu
- o Treballadors de serveis de restauració i personals

- Treballadors de serveis de protecció i seguretat
- Dependents de comerç i assimilats
- Treballadors qualificats en agricultura i pesca
- Treballadors qualificats de construcció, excepte operadors de màquines
- Treballadors qualificats d'indústries extractives, metal·lúrgia, construcció de màquines i assimilats
- Treballadors qualificats d'indústries d'arts gràfiques, tèxtil i confecció, elaboració d'aliments, ebenistes, artesans
- Operadors d'instal·lacions industrials de maquinària fixa; muntadors i assembladors
- Conductors i operadors de maquinària mòbil
- Treballadors no qualificats en serveis (excepte transports)
- Peons d'agricultura, pesca, alimentació, indústries manufactureres i transports
- Forces armades
- Persones que han deixat l'última feina fa més d'un any
- Parats que busquen una primera feina

English translation

* Mandatory field

Identifier *

Enter your initials (first name initial, first surname initial and second surname initial) followed by your age. Do not leave any blank space in between.

Age *

Sex *

Male

Female

Level of studies reached *

- Lower than first stage of secondary school
- Secondary education, first stage
- Secondary education, second stage
- Advanced vocational education
- First cycle university education (diploma, degree, engineering or graduate studies)
- Second cycle university education (master, postgraduate or doctoral studies)

In case you have reached university education, please specify.

Occupation *

- Public administration management and management of companies with 10 or more wage earners.
- Management of companies with less than 10 wage earners
- Management of companies without wage earners
- Professions associated with 2nd and 3rd cycle university degrees and the like
- Professions associated with a 1st cycle university degree and the like
- Support technicians and professionals
- Administrative type employees
- Catering services workers and personal services workers
- Protection and security service workers
- Retail workers and the like
- Workers skilled in agriculture and fishing
- Skilled construction workers, except machinery operators
- Skilled workers in the extractive industry, metallurgy, construction of machinery and related trades.
- Skilled workers from the graphic arts, textile and tailoring, elaboration of food, cabinetmakers, craftspersons and other similar industries
- Fixed machinery and industrial installation operators; fitters and assemblers.
- Mobile machinery drivers and operators
- Unskilled workers in the service sector (except transports)
- Agriculture, fishing, construction, manufacturing industries and transport labourers.
- Armed forces
- Unemployed for longer than one year
- Unemployed, seeking a first job

Annex 3: Documents related to the TTS main experiment on the reception of TTS AD

Annex 3.1. Clips' AD scripts

Clip 1

36 En un estudi fotogràfic, una dona rossa, esvelta i atractiva fa anar una càmera.
37 El noi del maletí posa assegut en un tamboret. No porta ulleres i va molt més elegant.
38 Ell torna el paquet de tabac a la butxaca de l'americana, que descansa en un t amboret. Torna a asseure's davant la càmera i mira la fotògrafa amb un mig somriure.
39 Ell travessa l'estança, contempla un mural de fotografies i torna al tamboret.
40 Se li acosta i li arregla el nus de la corbata.
41 Ell obeeix i la fotògrafa ajusta el trípode a l'altura adequada i l'enfoca amb la càmera.
42 Intercanvien un mig somriure.

Clip 2

175 En Dan s'espera en un vestíbul luxós, recolzat en una barana.
176 Passeja inquiet amunt i avall amb un llibre a les mans.
177 S'aboca per sobre la barana i fa un cop d'ull al pis de baix. L'Anna accedeix a l'edifici i passa de dret cap a unes escales.
178 Al primer pis, en Dan l'espera amb un posat impacient.
179 Entren al bar i s'apropen a la barra.
180 L'Anna segueix en Dan fins a una taula.
181 S'alça i fa per marxar, però ella no li deixa anar la mà.
182 Ell se li acosta a tocar i la besa.
183 Marxa mentre ella el segueix amb la mirada amb un mig somriure als llavis.

Annex 3.2. Post-questionnaire

* Necessari

Identificador *

Introdueix les teves inicials (la del nom, la del primer cognom i la del segon cognom) seguit de la teva edat. No hi deixis cap espai entremig.

Edat *

Sexe *

- Home
- Dona

Nivell d'estudis assolit *

- Inferiors a la primera etapa d'educació secundària
- Educació secundària, primer cicle
- Segona etapa d'educació secundària
- Formació professional de grau superior
- Estudis universitaris de primer cicle (diplomatura, llicenciatura, enginyeria o grau)
- Estudis universitaris de segon cicle (màster, postgrau o doctorat)

En cas d'haver cursat estudis superiors, digues quins.

Nivell professional *

- Direcció de l'administració pública i de les empreses de 10 o més assalariats
- Gerència d'empreses amb menys de 10 assalariats
- Gerència d'empreses sense assalariats
- Professions associades a titulació de segon i tercer cicle universitari i afins
- Professions associades a titulació de primer cicle universitari i afins
- Tècnics i professionals de suport
- Empleats de tipus administratiu
- Treballadors de serveis de restauració i personals
- Treballadors de serveis de protecció i seguretat
- Dependents de comerç i assimilats
- Treballadors qualificats en agricultura i pesca
- Treballadors qualificats de construcció, excepte operadors de màquines
- Treballadors qualificats d'indústries extractives, metal·lúrgia, construcció de màquines i assimilats
- Treballadors qualificats d'indústries d'arts gràfiques, tèxtil i confecció, elaboració d'aliments, ebenistes, artesans

- Operadors d'instal·lacions industrials de maquinària fixa; muntadors i assembladors
- Conductors i operadors de maquinària mòbil
- Treballadors no qualificats en serveis (excepte transports)
- Peons d'agricultura, pesca, alimentació, indústries manufactureres i transports
- Forces armades
- Persones que han deixat l'última feina fa més d'un any
- Parats que busquen una primera feina

Professió en les teves paraules

Tipus de discapacitat segons l'OMS *

- Ceguesa
- Baixa visió

Quant fa que tens deficiència visual? *

- De naixement
- Fa menys d'un any
- Fa entre un i deu anys
- Fa entre onze i vint anys
- Fa més de vint anys

Havies vist mai un producte audiodescrit (pel·lícules, sèries, obres de teatre, etc.)? *

- Sí
- No

Si n'havies vist, de quin tipus? (Pots marcar més d'una resposta)

- Pel·lícules
- Sèries
- Dibuixos animats
- Obres de teatre
- Òpera

Amb quina freqüència utilitzes productes audiodescrits? *

- Com a mínim 1 cop al dia
- Com a mínim 1 cop a la setmana
- Com a mínim 1 cop al mes
- Mai
- Altres:

Prefereixes que l'AD la llegeixi *

- un home
- una dona
- depèn del producte audiovisual
- tant me fa

Si depèn del producte audiovisual, de què depèn exactament?

Prefereixes que l'AD la llegeixi *

- una veu humana
- una veu artificial
- depèn del producte audiovisual
- tant me fa

Si depèn del producte audiovisual, de què depèn exactament?

Utilitzes dispositius electrònics amb aplicacions de veu sintetitzada, com ara mòbils o ordinadors? *

- Sí
- No

Amb quina freqüència? *

- Com a mínim 1 cop al dia
- Com a mínim 1 cop a la setmana
- Com a mínim 1 cop al mes
- Mai

Alguna vegada havies fet servir productes audiodescrits amb veu sintetitzada? *

- Sí
- No

Creus que és una solució alternativa a l'audiodescripció amb veu humana? *

- Sí
- No

Per què? *

Quina mena de productes utilitzaries amb audiodescripció amb veu sintetitzada? (Pots marcar més d'una resposta) *

- Pel·lícules
- Sèries
- Dibuixos animats
- Documentals
- Obres en directe
- Cap

De les 4 veus que has sentit, quina t'ha agradat més? *

- La masculina sintètica
- La masculina natural
- La femenina sintètica
- La femenina natural

Series capaç de classificar-les per ordre, començant per la que més t'ha agradat i acabant per la que menys t'ha agradat?

Altres comentaris

English translation

* Mandatory field

Identifier *

Enter your initials (first name initial, first surname initial and second surname initial) followed by your age. Do not leave any blank space in between.

Age *

Sex *

Male

Female

Level of studies reached *

- Lower than first stage of secondary school
- Secondary education, first stage
- Secondary education, second stage
- Advanced vocational education
- First cycle university education (diploma, degree, engineering or graduate studies)
- Second cycle university education (master, postgraduate or doctoral studies)

In case you have reached university education, please specify.

Occupation*

- Public administration management and management of companies with 10 or more wage earners.
- Management of companies with less than 10 wage earners
- Management of companies without wage earners
- Professions associated with 2nd and 3rd cycle university degrees and the like
- Professions associated with a 1st cycle university degree and the like
- Support technicians and professionals
- Administrative type employees
- Catering services workers and personal services workers
- Protection and security service workers
- Retail workers and the like
- Workers skilled in agriculture and fishing
- Skilled construction workers, except machinery operators
- Skilled workers in the extractive industry, metallurgy, construction of machinery and related trades.

- Skilled workers from the graphic arts, textile and tailoring, elaboration of food, cabinetmakers, craftspersons and other similar industries
- Fixed machinery and industrial installation operators; fitters and assemblers.
- Mobile machinery drivers and operators
- Unskilled workers in the service sector (except transports)
- Agriculture, fishing, construction, manufacturing industries and transport labourers.
- Armed forces
- Unemployed for longer than one year
- Unemployed, seeking a first job

Profession in your own words

Kind of visual impairment according to WHO *

- Blindness
- Low vision

How long have you been visually impaired for? *

- From birth
- For less than 1 year
- For between 1 and 10 years
- For between 11 and 20 years
- For more than 20 years any

Have you ever seen an audio described product (films, series, theatre plays, etc.)? *

- Yes
- No

In case you have, which kind of products? (You can tick more than one answer)

- Films
- Series
- Cartoons
- Theatre plays
- Opera plays

How often do you use audio described products? *

- At least once a day
- At least once a week
- At least once a month
- Never
- Other

Do you prefer the AD to be read by *

- A man
- A woman
- It depends on the audiovisual product
- I don't care

If it depends on the audiovisual product, what does it depend on exactly?

You prefer the AD to be read by *

- A human voice
- An artificial voice
- It depends on the audiovisual product
- I don't care

If it depends on the audiovisual product, what does it depend on exactly?

Do you use electronic devices with synthetic voice applications, such as mobile phones or computers? *

- Yes
- No

How often do you use them? *

At least once a day / At least once a week / At least once a month / Never

Have you ever used audio described products with synthetic voice? *

- Yes
- No

Do you think it is an alternative solution to human voiced audio description? *

- Yes
- No

Why do you think so? *

What kind of products would you use with synthetic voiced AD? (You can tick more than one answer) *

- Films
- Series
- Cartoons
- Documentaries
- Live plays
- None

Which voice, from the 4 voices you have just heard, did you like the most? *

- The masculine synthetic voice
- The masculine natural voice
- The feminine synthetic voice
- The feminine natural voice

Would you be able to rank them in order, from the one you like the most to the one you liked the least?

Other comments

Annex 3.3. Sampling plan

2011 DATA ACCORDING TO IDESCAT: TOTAL=31256

SEX	Frequency	%
Female	17485	55.94
Male	13771	44.06
Total	31256	100.00

AGE GROUP	Frequency	%
18 - 19	108	0.35
20 - 34	1434	4.59
35 - 44	2569	8.22
45 - 54	4069	13.02
55 - 64	5706	18.26
65 - 74	6530	20.89
From 75 onwards	10840	34.68
Total	31256	100.00

IMPAIRMENT	Frequency	%
MILD BINOCULAR VISUAL ACUITY LOSS	3517	11.25
MODERATE BINOCULAR VISUAL ACUITY LOSS	6321	20.22
SEVERE BINOCULAR VISUAL ACUITY LOSS	3152	10.08
PROFOUND BINOCULAR VISUAL ACUITY LOSS	11620	37.18
BLINDNESS	1514	4.84
MONOCULAR VISION LOSS	1566	5.01
VISUAL EFFICIENCY DECREASE	3566	11.41
Total	31256	100.00

IMPAIRMENT DEGREE	Frequency	%
Less than 65%	9731	31.13
65-74%	5464	17.48
More than or =75%	16061	51.39
Total	31256	100.00

AGE GROUP	Frequency	%	% women	Women	% men	Men
18-19	108	0.35	40.74	44	59.26	64
20-34	1434	4.59	43.79	628	56.21	806
35-44	2569	8.22	45.74	1175	54.26	1394
45-54	4069	13.01	49.84	2028	50.16	2041
55-64	5706	18.26	51.72	2951	48.28	2755
65-74	6530	20.89	56.25	3673	43.75	2857
75 and more	10840	34.68	64.45	6986	35.55	3854
Total	31256	100.00				

If the sample was	**67**	**participants**			
AGE GROUP	Total	% women	Women	% men	Men
18-19	0	40,74	0	59,26	0
20-34	3	43,79	1	56,21	2
35-44	6	45,74	3	54,26	3
45-54	9	49,84	4	50,16	4
55-64	12	51,72	6	48,28	6
65-74	14	56,25	8	43,75	6
75 and more	23	64,45	15	35,55	8

ACTUAL SAMPLE	**67**	**participants**	
AGE GROUP	Total	Women	Men
18-19	0	0	0
20-34	9	5	4
35-44	12	5	7
45-54	15	7	8
55-64	15	9	6
65-74	13	9	4
75 and more	3	2	1
TOTAL	67	37	30

Annex 3.4. Cold-calling email for the recruitment of volunteer participants

Hola:

Sóc l'Anna Fernández Torné, doctoranda del Centre d'Accessibilitat I Intel·ligència Artificial de Catalunya, de la Universitat Autònoma de Barcelona. Em poso en contacte amb vosaltres perquè necessitaria voluntaris catalanoparlants i amb qualsevol grau de discapacitat visual per a un dels experiments que formen part de la meva tesi, titulada "Audiodescripció I tecnologies: estudi sobre la semi-automatització de la traducció i locució d'audiodescripcions", inscrita en el projecte TECNACC, Tecnologia per a l'accessibilitat: estratègies d'automatització de la traducció i locució d'audiodescripcions, guardonat amb el premi Aposta 2011, de la UAB, i en el projecte ALST, Accessibilitat lingüística i sensorial: tecnologies per a les veus superposades i l'audiodescripció, finançat pel Ministeri d'Economia i Competitivitat (FFI2012-31023).

L'experiment que pretenc dur a terme dura aproximadament una hora. L'únic que cal fer és escoltar atentament les veus de l'audiodescripció, avaluar-les (només les veus, deixant de banda el contingut, el que s'hi diu) i respondre uns qüestionaris per donar-nos-en la vostra opinió.

La participació en aquest experiment és del tot voluntària i, tot i que demanarem algunes dades personals, com ara l'edat, els estudis previs i la professió, la informació es tractarà amb la màxima confidencialitat i les dades es processaran de forma anònima. Per això caldrà que els participants donin el seu consentiment, que hauré d'enregistrar perquè en quedi constància.

Si teniu cap pregunta, no dubteu a fer-me-la arribar per correu electrònic o telefònicament al 629126498.

Gràcies.

English translation

Dear all,

My name is Anna Fernández Torné and I am a PhD student of the Accessibility and Artificial Intelligence Centre of Catalonia, of the Universitat Autònoma de Barcelona. I am contacting you because I would need to recruit Catalan speaking volunteers having any degree of visual impairment for one of the experiments that constitute my thesis, entitled "Audio description and technologies: study on the semi-automatisation of the translation and voicing of audio descriptions", developed within the framework of the

TECNACC, Technology for accessibility: strategies for the automatisation of the translation and voicing of audio descriptions, awarded with the UAB Aposta 2011 prize, and within the framework of the ALST project, Linguistic and sensorial accessibility: technologies for the voice-over and audio description, funded by the Spanish Ministerio de Economía y Competitividad (under grant number FFI2012-31023).

The experiment I intend to conduct takes around one hour. The only thing you need to do is to listen attentively to the voices of the audio description, assess them (only the voices, thus neglecting the contents, what is actually being said) and answer some questionnaires to give us your opinion.

Your participation on this study is on a voluntary basis and, even though we will ask for some personal details, such as your age, education and profession, the information will be treated with utmost confidentiality and data will be processed anonymously. This is why you will be asked to give your consent, which I will need to record for purposes of record.

Should you have any question, please do not hesitate to let me know either by email or by phone at 629126498.

Thank you!

Annex 3.5. Information sheet

FULL D'INFORMACIÓ SOBRE L'EXPERIMENT

L'objectiu de l'estudi és avaluar l'acceptació de diverses veus en l'àmbit de l'audiodescripció en català.

La vostra participació en l'experiment serà la següent: haureu d'avaluar 4 mostres de veu mitjançant qüestionaris seguint les instruccions que us donarem.

La vostra participació és totalment voluntària i us podeu retirar de l'estudi en qualsevol moment sense haver-ho de justificar de cap manera ni sense que això us afecti negativament de cap manera.

Les vostres dades seran totalment anònimes ja que al full escrit només us demanarem que hi escriviu el vostre sexe (home/dona), l'edat, els estudis previs i la professió.

No rebreu cap compensació econòmica per aquesta participació, la qual en cap cas les farà servir amb ànim de lucre.

Les investigadores que duran a terme l'experiment són l'Anna Fernández Torné (que participarà en les proves) i la Dra. Anna Matamala (que dirigeix la investigació). Aquesta darrera és la responsable i la podeu trobar a anna.matamala@uab.cat, al despatx K-2012 de la UAB o bé al telèfon 935811646.

Si voleu continuar informats de les evolucions de la investigació, només cal que ho demaneu a la investigadora principal al correu electrònic anterior.

MOLTES GRÀCIES per la vostra participació!

(Nota: a l'altra cara d'aquest full hi teniu el document de consentiment informat)

English translation

INFORMATION SHEET

The aim of this study is to assess the acceptance of several Catalan voices in the audio description field.

Your participation in the experiment will be as follows: you will have to assess 4 voice samples by means of questionnaires following the instructions we will give to you.

Your participation is absolutely voluntary and you may withdraw from the study at any time without further justification and without that negatively affecting you in any way.

Your personal details will be completely anonymous since in the written form you will only be asked to indicate your gender (man/woman), your age, your education and your profession.

You will receive no economic compensation for your participation, which will not be used for profit under any circumstance.

The researchers who will conduct the experiment are Anna Fernández Torné (who will take part in the tests) and Dr. Anna Matamala, who is the supervisor and responsible for this research. You can contact her at anna.matamala@uab.cat, or by phone at 935811646. She can also be found at the office K-2012 in the UAB.

If you would like to be informed of the research progress, please ask the principal investigator by contacting her at the previous email address.

Thank you very much for your participation!

(Note: you will find the informed consent form on the other side of this page)

Annex 3.6. Informed consent form

CONSENTIMENT INFORMAT

Títol de l'experiment: Avaluació subjectiva de veus per a l'audiodescripció

Jo, (NOM I COGNOMS) _____

- He llegit el full d'informació que m'han donat com també el de consentiment informat.
- He rebut prou informació sobre l'estudi i l'he entesa.
- He pogut fer preguntes sobre l'estudi.

He parlat amb: Anna Fernández Torné

- Entenc que la meva participació és voluntària
- Entenc que la meva participació és no remunerada
- Entenc que la meva informació serà confidencial
- Entenc que em puc retirar de l'estudi quan vulgui i sense haver de donar explicacions ni sense que hi hagi repercussions negatives.

Dono lliurament la meva conformitat per poder participar en l'estudi.

Data

Firma del participant

(Nota: a l'altra cara d'aquest full hi teniu el document d'informació)

English translation

INFORMED CONSENT

Title of the experiment: Subjective assessment of voices for the audio description

(NAME AND SURNAME) _____

- I have read the information sheet I have been given, as well as the informed consent form.
- I have received enough information about the study and I have understood it.
- I have been able to make questions about the study.

I have talked to: **Anna Fernández Torné**

- I understand that my participation is voluntary
- I understand that my participation is unpaid
- I understand that my personal details will be treated confidentially
- I understand that I can withdraw from the study whenever I want giving no particular reason for it and having no negative repercussions.

I freely consent to participate on this research study.

Date

Signature of participant

(Note: you will find the information sheet on the other side of this page)

Annex 4: Documents related to the MT pre-test for the selection of a MT engine

Annex 4.1. AD script

AD unit		Duration (seconds)	Words	Characters
1	A professional camera rests on its tripod. A woman peering down through the viewfinder lifts her head.	4.600	17	102
2	Dan sits stiffly on a stool in front of a screen. The beautiful photographer turns away.	4.240	16	88
3	Dressed all in black, Dan puts back his cigarette packet back in his jacket pocket and eyeing the photographer, who is in her thirties, tall and slim, with a chiselled large-featured face. He sits back down. She studies him with a glint in her eye.	11.240	45	248
4	She smiles warmly.	1.040	3	18
5	She nods. Dan stares steadily at her unsmiling. As she turns away again he gets to his feet and crosses the studio.	6.240	22	115
6	Dan looks at some of her photos, which hang on the walls. They are mainly of people.	4.160	17	84

7	Dan wanders back towards the stool and sits.	2.320	8	44
8	She looks coolly at him.	1.480	5	24
9	He straightens his back as she continues to take pictures. She tilts her head to one side regarding him thoughtfully.	5.280	20	117
10	He raises them again flashing a smile. The photographer steps purposefully towards him and adjusts his tie. He looks up at her.	6.600	22	127
11	She goes back to her camera and looks through the viewfinder at Dan. Then lifts her head to look directly at him.	5.040	22	113
12	He stands. She raises the camera on the tripod.	2.200	9	47
13	Dan's piercing eyes dart to one side then fall on the photographer, who meets his gaze and smiles softly, her eyes glistening.	6.680	22	126
14	Her smile gone, she stands motionless, her eyes still fixed on him.	3.160	12	67

Annex 4.2. Ranking task

Rank the translation from best to worst, assigning numbers to each unit from 5 (best) to 1 (worst) in the left column.

	A professional camera rests on its tripod. A woman peering down through the viewfinder lifts her head.
	Unes restes de càmera professionals en el seu tripod. Una dona que mira atentament avall a través del viewfinder ascensors el seu cap.
	Una càmera professional es basa en el seu trípode. Una dona mirant cap avall a través del visor aixeca el cap.
	Un professional de la càmera es basa en el seu trípode . Una dona mirant cap avall a través del visor aixeca el cap .
	Una càmera professional es basa en el trípode. Una dona mirant cap avall a través del visor aixeca el seu cap.
	Una càmera professional es recolza al seu trípode. Una dona que fita avall el viewfinder alça el seu cap.
	Dan sits stiffly on a stool in front of a screen. The beautiful photographer turns away.
	Dan rígid asseu en un tamboret davant d'una pantalla. El fotògraf bella allunya.
	Dan està assegut rígidament en una cadira davant d'una pantalla . El fotògraf bella s'allunya .
	Dan seu stiffly en un stool davant d'una pantalla. Les voltes de fotògraf boniques fora.
	Dan seu rígidament en un tamboret davant d'una pantalla. El fotògraf bonic s'allunya.
	Dan es troba stiffly en un tamboret davant d'una pantalla. La bella fotògraf es converteix distància.
	Dressed all in black, Dan puts back his cigarette packet back in his jacket pocket and eyeing the photographer, who is in her thirties, tall and slim, with a chiselled large-featured face. He sits back down. She studies him with a glint in her eye.
	Vestida de negre , Dan torna a posar el seu paquet de cigarrets a la butxaca de la jaqueta i mirant al fotògraf , que és d'uns trenta anys , alt i prim , amb una cara gran equipat cisellat . Ell torna a seure. Ella ho estudia amb una brillantor en els seus ulls .
	va Vestir tot en negre, Dan posa endarrere el seu paquet de cigarret endarrere en la seva butxaca de jaqueta i eyeing el fotògraf, qui és en el seu thirties, alt i esvelt, amb un chiselled gran-cara presentada. Seu endarrere avall. Li estudia amb un glint en el seu ull.
	Vestida de negre, Dan posa de nou el seu paquet de tabac enrere a la butxaca de la

	jaqueta i eyeing el fotògraf, que en els seus trenta anys, alt i prim, amb una escarpada gran destacats cara. Ell és tornar a baixar. Estudia-lo amb un glint en la seva mirada.
	Vestida tota de negre, Dan posa altra vegada el seu paquet de cigarrets tornada a la seva butxaca de la jaqueta i mirant el fotògraf, que està en els seus trenta, allargades i estretes, amb una cara immensament gran presentat. S'asseu. Estudia-lo amb un glint al seu ull.
	Vestit completament amb negre,Dan desa|endarrereix el seu paquet de cigarrets altra vegada|de tornada a la seva butxaca de jaquetes i eyeing el fotògraf, que és|està en les seves trentenes, altes i esveltes, amb una cara de manera gran presentada cisellada. S'asseu avall. L'estudia amb un centellejar al seu ull.
	She smiles warmly.
	Ella somriu càlida.
	Somriu de manera calenta.
	somriu warmly.
	Ella somriu amb gust .
	Ella somriu calorosament.
	She nods. Dan stares steadily at her unsmiling. As she turns away again he gets to his feet and crosses the studio.
	Ella assenteix amb el cap. Dan escales de manera constant en la seva unsmiling. Com ella torna de distància de nou es posa als seus peus i creua l'estudi.
	Ella assenteix amb. Dan mira constantment al seu unsmiling. Com ella allunya nou ell arriba als seus peus i creua l'estudi.
	assenteix amb el cap. Dan Mira fixament fermament al seu unsmiling. Mentre gira fora una altra vegada aconsegueix als seus peus i travessa l'estudi.
	Fa que sí. Dan fitafermament|constantment que inen. Mentre|Com s'allunya una altra vegada arriba als seus peus i creua l'estudi.
	Ella assenteix amb el cap . Donen es queda mirant fixament a ella sense somriure . Com ella ho rebutja de nou es posa dret i creua l'estudi .
	Dan looks at some of her photos, which hang on the walls. They are mainly of people.
	Donen es veu en algunes de les seves fotos , que pengen de les parets . Es tracta principalment de persones .
	Donen veu a algunes de les seves fotografies, que pengen de les parets. Són principalment de les persones.
	Dan mira algunes de les seves fotos, que penja a les parets. Són principalment de persones.
	Dan mira algunes de les seves fotos, que es pengen sobre les parets. Són|Estan principalment de gent.
	Dan aspectes a alguns de les seves fotos, els quals pengen en les parets. Són

	principalment de persones.
	Dan wanders back towards the stool and sits.
	Dan vaga endarrere cap al stool i seu.
	Dan vaga altra vegada|de tornada cap al tamboret i seu.
	Dan es distreu tornada cap el Tamboret i s'asseu.
	Dan es desplaça de nou a la banqueta i s'asseu.
	Donen vaga cap al tamboret i s'asseu.
	She looks coolly at him.
	Ella es veu amb tranquil·litat a ell.
	Ella mira còmodament.
	Ella mira amb fredor cap a ell .
	El mira frescament.
	mira coolly a ell.
	He straightens his back as she continues to take pictures. She tilts her head to one side regarding him thoughtfully.
	Ell es redreça la seva esquena com ella continua per fer fotos. Ella s'inclina el cap cap a un costat sobre ell, pensatiu.
	Endreça el seu darrere|esquenacom continua fentquadres|fotografies. Inclina el seu cap a un costat quant a ellconsideradament|pensarosament.
	Es redreça l'esquena mentre ella continua prenent imatges . Ella inclina el cap cap a un costat respecte a ell , pensatiu - .
	Va arranja l'esquena mentre ella continua fer fotografies. Ella inclina el seu cap a un costat pel que fa a ell pensatiu.
	Ell straightens seu dóna suport a tan continua agafar fotografies. Ella tilts li es dirigeix a un costat que li considera thoughtfully.
	He raises them again flashing a smile. The photographer steps purposefully towards him and adjusts his tie. He looks up at her.
	Els alça una altra vegadaemetent|enfocant un somriure. El fotògraf passa resolutament cap a ell i ajusta el seulligam|corbata. Admira a ella.
	els aixeca una altra vegada centellejant un somriure. Els passos de fotògraf purposefully cap a ell i ajusta la seva corbata. Mira amunt a ella.
	Planteja tornar-los a parpellejar un somriure. El fotògraf passos intencionadament cap a ell i li ajusta la corbata. Es veu a ella.
	Aixeca de nou amb un somriure . El fotògraf passos deliberadament cap a ell i s'ajusta la corbata . Aixeca la mirada cap a ella .
	Eleva els tapajunts de nou un somriure. El fotògraf intencionadament els passos cap a ell i ajusta la corbata. Es mira a ella.

	She goes back to her camera and looks through the viewfinder at Dan. Then lifts her head to look directly at him.
	Ella va tornar a la seva càmera i es veu a través del visor de Dan. Llavors aixeca el cap per mirar directament a ell.
	Ella torna a la seva càmera i mira a través del visor en Dan . Després aixeca el cap per mirar directament als ulls .
	Torna a la seva càmera i fulleja el viewfinder a Dan. Llavorsascensors|impulsos ella cap per mirar-lo directament.
	torna a la seva càmera i aspectes a través del viewfinder a Dan. Llavors ascensors el seu dirigir-se a aspecte directament a ell.
	Ella es remunta a la seva càmera i es veu a través del visor de Dan. Llavors aixeca el seu cap de mirar directament a ell.
	He stands. She raises the camera on the tripod.
	està. Aixeca la càmera en el tripod.
	Es troba. S'aixeca la càmera en el trípode.
	Es posa dret . Ella aixeca la càmera en el trípode .
	Està dret. Alça la càmera en el trípode.
	Ell roman. Ella aixeca la càmera sobre el trípode.
	Dan's piercing eyes dart to one side then fall on the photographer, who meets his gaze and smiles softly, her eyes glistening.
	De Dan perforació ulls dard a un costat després cau sobre el fotògraf, que compleix la seva mirada i somriures suaument, ulls brillants.
	Els ulls penetrants de Dan es llancen a un costat llavors caure en el fotògraf, que troba la seva mirada i somriu suaument, resplendint els seus ulls.
	Dan ulls penetrants dart, a un costat, després de caure en el fotògraf, que compleixi amb la seva mirada i el somriure suau, els ulls brillants.
	Dan ulls penetrants dart cap a un costat llavors caiguda en el fotògraf, qui coneix seu mirar fixament i somriures suaument, els seus ulls que resplendeixen.
	Donen penetrants ulls dard a un costat i després caure en el fotògraf , que es troba amb la seva mirada i somriu suaument , els seus ulls brillaven .
	Her smile gone, she stands motionless, her eyes still fixed on him.
	El seu somriure anat, ella es troba immòbil, que els ulls encara fixos sobre ell.
	El seu somriure anat, es queda immòbil, els seus ulls encara fixaven en ell.
	El seu somriure anat, està immòbil, els seus ulls encara fixats en ell.
	El seu somriure va desaparèixer, ella roman immòbil , amb els ulls fixos en ell .
	El seu somriure ha anat, que destaca immòbil, els ulls encara fixa en ell.

Annex 4.3. Raw MT outputs

Source	A	B	C	D	E
A professional camera rests on its tripod. A woman peering down through the viewfinder lifts her head.	Unes restes de càmera professionals en el seu tripod. Una dona que mira atentament avall a través del viewfinder ascensors el seu cap.	Una càmera professional es basa en el trípode. Una dona mirant cap avall a través del visor aixeca el seu cap.	Una càmera professional es recolza al seu trípode. Una dona que fita avall el viewfinder alça el seu cap.	Un professional de la càmera es basa en el seu trípode . Una dona mirant cap avall a través del visor aixeca el cap .	Una càmera professional es basa en el seu trípode. Una dona mirant cap avall a través del visor aixeca el cap.
Dan sits stiffly on a stool in front of a screen. The beautiful photographe r turns away.	Dan seu stiffly en un stool davant d'una pantalla. Les voltes de fotògraf boniques fora.	Dan rígid asseu en un tamboret davant d'una pantalla. El fotògraf bella allunya.	Dan seu rígidament en un tamboret davant d'una pantalla. El fotògraf bonic s'allunya.	Dan està assegut rígidament en una cadira davant d'una pantalla . El fotògraf bella s'allunya .	Dan es troba stiffly en un tamboret davant d'una pantalla. La bella fotògraf es converteix distància.
Dressed all in black, Dan puts back his cigarette packet back in his jacket pocket and eyeing the photographe r, who is in her thirties, tall and slim, with a chiselled large-featured face. He sits back down. She studies him with a glint in her eye.	va Vestir tot en negre, Dan posa endarrere el seu paquet de cigarret endarrere en la seva butxaca de jaqueta i eyeing el fotògraf, qui és en el seu thirties, alt i esvelt, amb un chiselled gran-cara presentada. Seu endarrere avall. Li estudia amb un glint en el seu ull.	Vestida tota de negre, Dan posa altra vegada el seu paquet de cigarrets tornada a la seva butxaca de la jaqueta i mirant el fotògraf, que està en els seus trenta, allargades i estretes, amb una cara immensame nt gran presentat. S'asseu. Estudia-lo	Vestit completame nt amb negre,Dan de sa|endarrere ix el seu paquet de cigarrets altra vegada|de tornada a la seva butxaca de jaquetes i eyeing el fotògraf, que és|està e n les seves trentenes, altes i esveltes, amb una cara de manera gran presentada cisellada.	Vestida de negre , Dan torna a posar el seu paquet de cigarrets a la butxaca de la jaqueta i mirant al fotògraf , que és d'uns trenta anys , alt i prim , amb una cara gran equipat cisellat . Ell torna a seure. Ella ho estudia amb una brillantor en els seus ulls .	Vestida de negre, Dan posa de nou el seu paquet de tabac enrere a la butxaca de la jaqueta i eyeing el fotògraf, que en els seus trenta anys, alt i prim, amb una escarpada gran destacats cara. Ell és tornar a baixar. Estudia-lo amb un glint

		amb un glint al seu ull.	S'asseu avall. L'estudia amb un centellejar al seu ull.		en la seva mirada.
She smiles warmly.	somriu warmly.	Ella somriu calorosament.	Somriu de manera calenta.	Ella somriu amb gust .	Ella somriu càlida.
She nods. Dan stares steadily at her unsmiling. As she turns away again he gets to his feet and crosses the studio.	assenteix amb el cap. Dan Mira fixament fermament al seu unsmiling. Mentre gira fora una altra vegada aconsegueix als seus peus i travessa l'estudi.	Ella assenteix amb. Dan mira constantment al seu unsmiling. Com ella allunya nou ell arriba als seus peus i creua l'estudi.	Fa que sí. Dan fitafermament|constantment que inen. Mentre |Com s'allunya una altra vegada arriba als seus peus i creua l'estudi.	Ella assenteix amb el cap . Donen es queda mirant fixament a ella sense somriure . Com ella ho rebutja de nou es posa dret i creua l'estudi .	Ella assenteix amb el cap. Dan escales de manera constant en la seva unsmiling. Com ella torna de distància de nou es posa als seus peus i creua l'estudi.
Dan looks at some of her photos, which hang on the walls. They are mainly of people.	Dan aspectes a alguns de les seves fotos, els quals pengen en les parets. Són principalment de persones.	Dan mira algunes de les seves fotos, que penja a les parets. Són principalment de persones.	Dan mira algunes de les seves fotos, que es pengen sobre les parets. Són|Estan principalment de gent.	Donen es veu en algunes de les seves fotos , que pengen de les parets . Es tracta principalment de persones .	Donen veu a algunes de les seves fotografies, que pengen de les parets. Són principalment de les persones.
Dan wanders back towards the stool and sits.	Dan vaga endarrere cap al stool i seu.	Dan es distreu tornada cap el Tamboret i s'asseu.	Dan vaga altra vegada|de tornada cap al tamboret i seu.	Donen vaga cap al tamboret i s'asseu.	Dan es desplaça de nou a la banqueta i s'asseu.
She looks coolly at him.	mira coolly a ell.	Ella mira còmodament .	El mira frescament.	Ella mira amb fredor cap a ell .	Ella es veu amb tranquil·litat a ell.
He straightens his back as she continues to take pictures. She	Ell straightens seu dóna suport a tan continua agafar fotografies.	Va arranja l'esquena mentre ella continua fer fotografies. Ella inclina el	Endreça el seu darrere|esquenacom continua fentquadres|fotografies.	Es redreça l'esquena mentre ella continua prenent imatges . Ella	Ell es redreça la seva esquena com ella continua per

tilts her head to one side regarding him thoughtfully.	Ella tilts li es dirigeix a un costat que li considera thoughtfully.	seu cap a un costat pel que fa a ell pensatiu.	Inclina el seu cap a un costat quant a ellconsideradament|pensarosament.	inclina el cap cap a un costat respecte a ell , pensatiu - .	fer fotos. Ella s'inclina el cap cap a un costat sobre ell, pensatiu.
He raises them again flashing a smile. The photographer steps purposefully towards him and adjusts his tie. He looks up at her.	els aixeca una altra vegada centellejant un somriure. Els passos de fotògraf purposefully cap a ell i ajusta la seva corbata. Mira amunt a ella.	Eleva els tapajunts de nou un somriure. El fotògraf intencionadament els passos cap a ell i ajusta la corbata. Es mira a ella.	Els alça una altra vegadaemetent|enfocant un somriure. El fotògraf passa resolutament cap a ell i ajusta el seulligam|corbata. Admira a ella.	Aixeca de nou amb un somriure . El fotògraf passos deliberadament cap a ell i s'ajusta la corbata . Aixeca la mirada cap a ella .	Planteja tornar-los a parpellejar un somriure. El fotògraf passos intencionadament cap a ell i li ajusta la corbata. Es veu a ella.
She goes back to her camera and looks through the viewfinder at Dan. Then lifts her head to look directly at him.	torna a la seva càmera i aspectes a través del viewfinder a Dan. Llavors ascensors el seu dirigir-se a aspecte directament a ell.	Ella es remunta a la seva càmera i es veu a través del visor de Dan. Llavors aixeca el seu cap de mirar directament a ell.	Torna a la seva càmera i fulleja el viewfinder a Dan. Llavorsascensors|impulsos ella cap per mirar-lo directament.	Ella torna a la seva càmera i mira a través del visor en Dan . Després aixeca el cap per mirar directament als ulls .	Ella va tornar a la seva càmera i es veu a través del visor de Dan. Llavors aixeca el cap per mirar directament a ell.
He stands. She raises the camera on the tripod.	està. Aixeca la càmera en el tripod.	Ell roman. Ella aixeca la càmera sobre el trípode.	Està dret. Alça la càmera en el trípode.	Es posa dret . Ella aixeca la càmera en el trípode .	Es troba. S'aixeca la càmera en el trípode.
Dan's piercing eyes dart to one side then fall on the photographer, who meets his gaze and smiles softly, her eyes glistening.	Dan ulls penetrants dart cap a un costat llavors caiguda en el fotògraf, qui coneix seu mirar fixament i somriures suaument, els seus ulls que	De Dan perforació ulls dard a un costat després cau sobre el fotògraf, que compleix la seva mirada i somriures suaument, ulls brillants.	Els ulls penetrants de Dan es llancen a un costat llavors caure en el fotògraf, que troba la seva mirada i somriu suaument, resplendint	Donen penetrants ulls dard a un costat i després caure en el fotògraf , que es troba amb la seva mirada i somriu suaument , els seus ulls	Dan ulls penetrants dart, a un costat, després de caure en el fotògraf, que compleixi amb la seva mirada i el somriure

Her smile gone, she stands motionless, her eyes still fixed on him.	resplendeixen . El seu somriure anat, està immòbil, els seus ulls encara fixats en ell.	El seu somriure anat, ella es troba immòbil, que els ulls encara fixos sobre ell.	els seus ulls. El seu somriure anat, es queda immòbil, els seus ulls encara fixaven en ell.	brillaven . El seu somriure va desaparèixer, ella roman immòbil , amb els ulls fixos en ell .	suau, els ulls brillants. El seu somriure ha anat, que destaca immòbil, els ulls encara fixa en ell.

Annex 4.4. Contextualisation email for the participants

Hola a tothom,

Sou voluntaris per a un dels experiments que formen part de la meva tesi, centrada en l'estudi de l'audiodescripció i les tecnologies.

L'audiodescripció és la tècnica que s'utilitza per convertir la informació visual (les imatges) en informació oral. Es tracta de descriure el que passa en els espais de silenci, quan els personatges no parlen. La tecnologia que s'hi pretén aplicar és, en aquest cas, la traducció automàtica. Per traducció automàtica entenem la producció d'un text a partir d'un altre text en un altre idioma per mitjà de sistemes computacionals.

Aquest experiment en concret pretén avaluar la qualitat de cinc motors de traducció automàtica aplicats a l'audiodescripció de pel·lícules doblades al català. La vostra tasca se centrarà en la postedició (terme que es fa servir en l'àmbit de la traducció automàtica per referir-se a la revisió d'un text traduït automàticament) de la traducció de l'audiodescripció d'un clip de 3 minuts de durada extret de la pel·lícula *Closer* i en l'avaluació dels següents aspectes en relació amb cada segment:

· Necessitat de postedició: ha calgut revisar moltes coses perquè la traducció fos acceptable per publicar-se?
· Grau de dificultat de la postedició: fins a quin punt us ha semblat difícil dur a terme la postedició necessària?
· Adequació: fins a quin punt el contingut del text original s'expressa també en el text traduït?
· Fluïdesa: fins a quin punt la traducció és gramaticalment correcta i un parlant natiu la percebria com a llenguatge natural i intuïtiu (sense tenir en compte si el contingut del text traduït es correspon amb el del text original)?

Finalment, en un full de càlcul on presentaré cada segment original juntament amb les seves 5 corresponents traduccions automàtiques, haureu d'ordenar de millor a pitjor les versions dels diferents motors de traducció automàtica.

Fent clic al següent enllaç obtindreu tota la informació que necessiteu per començar el test, que previsiblement ha de tenir una durada de 4 hores, inclòs el temps de formació en l'eina de postedició. És molt important que us reserveu aquestes 4 hores seguides per a la prova i que eviteu al màxim qualsevol mena d'interrupcions. Del contrari, els resultats podrien sortir esbiaixats ja que no tos els participants realitzarien la prova en les mateixes condicions.

Si teniu cap dubte o voleu que us faci qualsevol aclariment mentre feu el test, podeu

posar-vos en contacte amb mi per correu electrònic a anna.torne@gmail.com o per Skype buscant l'usuari anna.fernandez.

M'hauríeu de tornar la prova abans de dijous que ve.

Gràcies per la vostra comprensió i col·laboració,

English translation

Dear all,

You are volunteers for one of the experiments that constitute my thesis, focused on the study of audio description and technologies.

Audio description is the technique used to turn visual information (images) into aural information. It involves describing what is happening in the silent gaps, when characters are not talking. The technology to be applied is, in this case, machine translation. We consider machine translation as the production of a text from another text in another language by means of computer systems.

This experiment in particular aims at assessing the quality of five machine translation engines applied to the audio description of films dubbed into Catalan. Your participation will be focused on the post-editing (term used in the machine translation domain to refer to the revision of a machine translated text) of the audio description translation of a 3-minute-long clip from the film *Closer* and on the assessment of the following aspects in relation to each segment:

· Post-editing necessity: did you have to correct many things in order for the translation to be acceptable to be published?
· Post-editing difficulty: to what extent did you find performing the necessary post-editing difficult?
· Adequacy: to what extent the content of the original text is also expressed in the translated text?
· Fluency: to what extent the translation is grammatically correct and a native speaker would perceive it as natural and intuitive language (without taking into account whether the translated text contents matches that of the original text)?

Finally, you will have to classify the versions of the different machine translation engines from better to worse in a spread sheet in which I will present each original segment together with their 5 corresponding machine translations.

If you click on the following link you will get all the information you need to start the test, which presumably will be four hours long, including the training for the post-

editing tool. It is very important that you reserve four hours in a row for the test and that you avoid inasmuch as possible any kind of interruptions. Otherwise, the results might be biased since not all participants would perform the test under the same circumstances.

If you have any doubts or you would like me to clarify you anything while conducting the test, you can contact me by email at anna.torne@gmail.com o per Skype (user: anna.fernandez).

The test should be finished by next Thursday.

Thank you for your understanding and collaboration,

Annex 4.5. Instructions

This test is estimated to last 4 hours. Please make sure you have as little interruptions as possible while carrying it out.

1. Read the fullinformació_participants MT engine selection file, fill it in, sign it if you agree to it and send it back to me by email to the following address: ana.fernandez.torne@uab.cat.

2. Download the tool you will use to perform the post-editing of the machine-translated audiodescription segments by clicking on the following link MT PET-v2.0.zip. This tutorial will help you familiarise with the main functions of the tool. You can also access the instructions in writing here if you like. The units included in the example are translated from English to Spanish, but remember the test will be done from English into Catalan.

3. Watch the clip. Please notice that the clip contains the dubbed version in Catalan and the English AD has been included in the silent gaps as subtitles (not audio), so don't expect to hear the English AD. If you are not seeing the AD subtitles, please enable them by pressing on the captions button on the bottom right hand corner of the displayer's screen

Please make sure you can access it at any time in case you need to go over it again.

4. You can also read the script of both the Catalan dialogues and the English AD. This will be the context of the segments you will be required to post-edit.

5. Open PET and start post-editing your jobs. You can use any resources that you deem necessary, for the revision, such as dictionaries, encyclopedias, etc.

Note: Please make sure you post-edit the files in the order you are told to according to the following chart. Choose the row containing your initials. This is extremely important for the development of the test.

MIA	**A**	**B**	**C**	**D**	**E**
CMJ	**B**	**D**	**E**	**C**	**A**
CCC	**C**	**E**	**B**	**A**	**D**
COB	**D**	**C**	**A**	**E**	**B**
EPR	**E**	**A**	**D**	**B**	**C**
APG	**A**	**C**	**B**	**E**	**D**
AMR	**B**	**A**	**E**	**D**	**C**

For the post-editing task, bear in mind that the translated segments are part of the audiodescription of a film. No synchronisation ("ajust") or time coding of the segments will be necessary. Please follow these guidelines:

- **Perform the minimum amount of editing necessary to make the AD translation ready for voicing retaining as much raw translation as possible**
- **Aim for grammatically, syntactically and semantically correct translation.**
- **Ensure that no information has been accidentally added or omitted.**
- **Ensure that the message transferred is accurate.**
- **Ensure that key terminology is correctly translated.**
- **Basic rules regarding spelling, punctuation and hyphenation apply.**

Please click on the right-hand column of the active unit <u>before starting to read the segment that you want to post-edit</u> as it is very important to calculate the time invested in the whole post-editing process.

6. Once you are done with the post-edition, rezip PET's main folder and send it back to me to the following address: anna.torne@gmail.com.

7. Open the spreadsheet "MT output ranking_XXX" that corresponds to your initials and rank the translations from best to worst, assigning numbers to each unit from 5 to 1 in the left colum.

- MT output ranking_MIA
- MT output ranking_CMJ
- MT output ranking_CCC
- MT output ranking_COB
- MT output ranking_EPR
- MT output ranking_APG
- MT output ranking_AMR

8. Fill in the Postquestionnaire form.

9. Fill in the MT engine selection pretest assessment form.

Thank you!

Annex 4.6. Information sheet

FULL D'INFORMACIÓ SOBRE L'EXPERIMENT

L'objectiu de l'estudi és avaluar la qualitat de diversos motors lliures de traducció automàtica en línia en la combinació lingüística anglès-català.

La vostra participació en l'experiment serà la següent: haureu d'avaluar 5 motors de traducció automàtica mitjançant qüestionaris i la postedició de 14 unitats descriptives seguint les instruccions que us donarem.

La vostra participació és totalment voluntària i us podeu retirar de l'estudi en qualsevol moment sense haver-ho de justificar de cap manera ni sense que això us afecti negativament de cap manera.

Les vostres dades seran totalment anònimes ja que al full escrit només us demanarem que hi escriviu el vostre sexe (home/dona), l'edat, els estudis previs i la professió.

No rebreu cap compensació econòmica per aquesta participació, la qual en cap cas es farà servir amb ànim de lucre.

Les investigadores que duran a terme l'experiment són l'Anna Fernández Torné (que participarà en les proves) i la Dra. Anna Matamala (que dirigeix la investigació). Aquesta darrera és la responsable i la podeu trobar a anna.matamala@uab.cat, al despatx K-2012 de la UAB o bé al telèfon 935811646.

Si voleu continuar informats de les evolucions de la investigació, només cal que ho demaneu a la investigadora principal al correu electrònic anterior.

MOLTES GRÀCIES per la vostra participació!

(Nota: a l'altra cara d'aquest full hi teniu el document de consentiment informat)

English translation

INFORMATION SHEET

The aim of this study is to assess the quality of various free online machine translation engines translating from English to Catalan.

Your participation in the experiment will be as follows: you will have to assess 5 machine translation engines by means of questionnaires and the post-editing of 14 descriptive units following the instructions we will give to you.

Your participation is absolutely voluntary and you may withdraw from the study at any time without further justification and without that negatively affecting you in any way.

Your personal details will be completely anonymous since in the written form you will only be asked to indicate your gender (man/woman), your age, your education and your profession.

You will receive no economic compensation for your participation, which will not be used for profit under any circumstance.

The researchers who will conduct the experiment are Anna Fernández Torné (who will take part in the tests) and Dr. Anna Matamala, who is the supervisor and responsible for this research. You can contact her at anna.matamala@uab.cat, or by phone at 935811646. She can also be found at the office K-2012 in the UAB.

If you would like to be informed of the research progress, please ask the principal investigator by contacting her at the previous email address.

Thank you very much for your participation!

(Note: you will find the informed consent form on the other side of this page)

Annex 4.7. Informed consent form

CONSENTIMENT INFORMAT

Títol de l'experiment: Avaluació subjectiva de motors de traducció automàtica

Jo, NOM I COGNOMS

- He llegit el full d'informació que m'han donat com també el de consentiment informat.
- He rebut prou informació sobre l'estudi i l'he entesa.
- He pogut fer preguntes sobre l'estudi.

He parlat amb: Anna Fernández Torné

- Entenc que la meva participació és voluntària
- Entenc que la meva participació és no remunerada
- Entenc que la meva informació serà confidencial
- Entenc que em puc retirar de l'estudi quan vulgui i sense haver de donar explicacions ni sense que hi hagi repercussions negatives.

Dono lliurament la meva conformitat per poder participar en l'estudi.

Data:

Firma de la investigadora Firma del participant

Anna Fernández Torné

(Nota: a l'altra cara d'aquest full hi teniu el document d'informació)

English translation

INFORMED CONSENT

Title of the experiment: Subjective assessment of machine translation engines

(NAME AND SURNAME) _____

- I have read the information sheet I have been given, as well as the informed consent form.
- I have received enough information about the study and I have understood it.
- I have been able to make questions about the study.

I have talked to: Anna Fernández Torné

- I understand that my participation is voluntary
- I understand that my participation is unpaid
- I understand that my personal details will be treated confidentially
- I understand that I can withdraw from the study whenever I want giving no particular reason for it and having no negative repercussions.

I freely consent to participate on this research study.

Date

Signature of researcher Signature of participant

Anna Fernández Torné

(Note: you will find the information sheet on the other side of this page)

Annex 4.8. PE tool instructions

Post-Editing Tool: PET

1. Download the tool by clicking on this link. Then go to File, Download. Please save the zipped file in the Desktop.
2. Double click on the zipped file and extract it in the Desktop as well.
3. Open the folder MT PET-v2.0 and double click the file run.bat. The main page appears.
4. In the main page, press User.

Insert your initials in capital letters (name, first surname, second surname) and press OK.

5. A list of pej files will appear in the New Jobs column. Choose example-job.pej and press Start. The annotation page appears.
6. At the right-hand side of the page there is the tool box, which includes the tool's action buttons.

Placing the pointer over a button for a second will show a tool tip text and information about shortcuts for that button. The one you need to remember is the down arrow (Next task), which will get you to the following unit.

The center of the annotation page contains the active unit in red, which is divided into two columns: the source text and the translation.

Once you click on the right column, the active unit turns yellow to indicate that the editing has started.

Note that the label at the top will change from “ready to edit" to “editing". The labels at the top also display the time spent on the current revision, the complete revisions and the total time spent on revisions.

IMPORTANT: Please click on the right-hand column of the active unit before starting to read the segment. This is essential for the right calculation of the time invested in the whole post-editing process.

7. To finish the edition press Next Task (down arrow).

An assessment page will be displayed after every active unit is completed.
You will have to rate each segment by how much you agree with the following statements, each of which is related to a different aspect:

- The MT text required no post-editing.

  How much post-editing was necessary to fix the translation.

- The MT text was easy to post-edit.

  How difficult the post-editing necessary to fix the translation was.

- All the information in the source text was present in the MT text.

  How much of the meaning expressed in the source was also expressed in the translation.

- The MT text is fluent Catalan.

  To what extent the translation was grammatically well formed and experienced as using natural or intuitive language by a native speaker without taking into account whether the information matches that of the source text or not.

8. Once you have entered all the assessments, click the Finish button on the upper right hand corner of the screen.

The post-editing tool logs the time spent to post-edit individual sentences. You can therefore pause between sentences, but you are asked to avoid pausing when possible.

If it is absolutely necessary, you can stop editing and even close the program. When reloading run.bat or resuming the post-editing, you will have to reenter your user (your initials in capital letters). The main page will display a .per file in the Results column. Select it and press Edit to continue editing a partial result.

Take your time until you are completely familiarised with the tool. Once you feel confident with it, you will be able to start with the actual post-edition.

Annex 4.9. Script of the Catalan dubbed dialogues together with the English AD script

00:09:55,900 --> 00:10:00,500
A professional camera rests on its tripod. A woman peering down through the viewfinder lifts her head.

TAKE #25 ---[CA-SE]-----
00:12:01

ANNA (G) Bé... / He de canviar el carret. Vas bé, (ON/OFF) de temps?
DAN (G)

00:10:05,780 --> 00:10:10,020
Dan sits stiffly on a stool in front of a screen. The beautiful photographer turns away.

DAN Et fa res, que fumi?
ANNA Si no te'n pots estar...
DAN Tampoc és això.
ANNA (E) Doncs no ho facis.

00:10:16,060 --> 00:10:27,300
Dressed all in black, Dan puts back his cigarette packet back in his jacket pocket and eyeing the photographer, who is in her thirties, tall and slim, with a chiselled large-featured face. He sits back down. She studies him with a glint in her eye.

TAKE #26
00:12:25

ANNA (G) / M'ha agradat, el llibre.
DAN Gràcies.

00:10:31,460 --> 00:10:32,500
She smiles warmly.

ANNA (RIALLA) / Quan surt publicat?
DAN (E/ON) L'any que ve. Com és, que l'has llegit?
ANNA (OFF) (AV.) El teu editor me'n va enviar una (OFF/ON) còpia. L'he llegit aquesta nit. / M'he quedat fins a les quatre.
DAN Quin compliment.

TAKE #27
00:12:44

ANNA (E) Està inspirada en algú que coneixes, la protagonista?
DAN En una noia que es diu Alice.
ANNA (E) (G) / (ON) I què li sembla, que li hagis robat la vida?
DAN Només ha sigut un préstec. / Li dedico el llibre. Està contenta.
ANNA (G)

00:10:57,860 --> 00:11:04,100
She nods. Dan stares steadily at her unsmiling. As she turns away again he gets to his feet and crosses the studio.

TAKE #28
00:13:05

DAN (E) Fas exposicions?
ANNA (E) De tant en tant. / L'any que ve en faig una.

00:11:08,980 --> 00:11:13,140
Dan looks at some of her photos, which hang on the walls. They are mainly of people.

DAN (OFF) De retrats?
ANNA (OFF) (G)
DAN De qui?
ANNA De desconeguts.

00:11:20,580 --> 00:11:22,900
Dan wanders back towards the stool and sits.

DAN I què els hi sembla, als desconeguts, que els hi robis la vida?

TAKE #29
00:13:27

ANNA És un préstec.
DAN Sóc un desconegut, (ON/OFF) jo?
ANNA No. / Ets un encàrrec.

00:11:33,380 --> 00:11:34,860
She looks coolly at him.

ANNA No t'encongeixis. Posa't dret.
DAN (G)

00:11:37,620 --> 00:11:42,900
He straightens his back as she continues to take pictures. She tilts her head to one side regarding him thoughtfully.

DAN L'has trobat obscè?

ANNA (E) El què?
DAN El llibre.
ANNA (G) / L'he trobat / encertat.
DAN En quin aspecte?
ANNA En el sexual. / En.. l'amorós.

TAKE #30
00:13:55
DAN (E) En quin sentit?
ANNA (G) L'has escrit tu.
DAN Tu l'has llegit,.. fins a les quatre.
ANNA No aixequis les celles. Se't veu molt tibat.

00:12:02,460 --> 00:12:09,060
He raises them again flashing a smile. The photographer steps purposefully towards him and adjusts his tie. He looks up at her.

DAN Però t'ha agradat?
ANNA Sí, però puc canviar d'idea. / (E) (G)

00:12:13,900 --> 00:12:18,940
She goes back to her camera and looks through the viewfinder at Dan. Then lifts her head to look directly at him.

ANNA (ON) Aixeca't.
00:12:21,220 --> 00:12:23,420
He stands. She raises the camera on the tripod.

DAN Alguna crítica?
ANNA El títol no em convenç.

TAKE #31
00:14:32
DAN Alguna proposta?
ANNA "L'aquàrium"?
DAN T'han agradat les marranades. / T'agraden els aquaris.
ANNA Els peixos són terapèutics.
DAN (T) (OFF) Hi vas molt, (OFF/ON) oi que sí?
ANNA (OFF/ON) Quan puc, sí.
DAN (OFF) A lligar (OFF/ON) amb desconeguts?
ANNA A fotografiar.. desconeguts.

00:12:48,620 --> 00:12:55,300

Dan's piercing eyes dart to one side, then fall on the photographer, who meets his gaze and smiles softly, her eyes glistening.

TAKE #32
00:14:57
DAN Acosta't.

00:12:58,660 --> 00:13:01,820
Her smile gone, she stands motionless, her eyes still fixed on him.

Annex 4.10. Post-questionnaire

* Necessari

Identificador *

Introdueix les teves inicials (la del nom, la del primer cognom i la del segon cognom) seguit de la teva edat. No hi deixis cap espai entremig.

Edat *

Sexe *

- o Home
- o Dona

Nivell d'estudis assolit *

- o Inferiors a la primera etapa d'educació secundària
- o Educació secundària, primer cicle
- o Segona etapa d'educació secundària
- o Formació professional de grau superior
- o Estudis universitaris de primer cicle (diplomatura, llicenciatura, enginyeria o grau)
- o Estudis universitaris de segon cicle (màster, postgrau o doctorat)

En cas d'haver cursat estudis superiors, digues quins.

Nivell professional *

- o Direcció de l'administració pública i de les empreses de 10 o més assalariats
- o Gerència d'empreses amb menys de 10 assalariats
- o Gerència d'empreses sense assalariats
- o Professions associades a titulació de segon i tercer cicle universitari i afins
- o Professions associades a titulació de primer cicle universitari i afins
- o Tècnics i professionals de suport

- o Empleats de tipus administratiu
- o Treballadors de serveis de restauració i personals
- o Treballadors de serveis de protecció i seguretat
- o Dependents de comerç i assimilats
- o Treballadors qualificats en agricultura i pesca
- o Treballadors qualificats de construcció, excepte operadors de màquines
- o Treballadors qualificats d'indústries extractives, metal·lúrgia, construcció de màquines i assimilats
- o Treballadors qualificats d'indústries d'arts gràfiques, tèxtil i confecció, elaboració d'aliments, ebenistes, artesans
- o Operadors d'instal·lacions industrials de maquinària fixa; muntadors i assembladors
- o Conductors i operadors de maquinària mòbil
- o Treballadors no qualificats en serveis (excepte transports)
- o Peons d'agricultura, pesca, alimentació, indústries manufactureres i transports
- o Forces armades
- o Persones que han deixat l'última feina fa més d'un any
- o Parats que busquen una primera feina

Professió en les teves paraules *

Anys d'experiència en l'àmbit de la traducció *

Especialització en l'àmbit de la traducció *

Has traduït mai un producte audiovisual (pel·lícules, sèries, documentals)? *

- o Sí
- o No

Si n'has traduït, digues si ha estat:

- o Com a traductor audiovisual professional
- o Com a estudiant de traducció audiovisual
- o Altres

Havies vist mai un producte audiovisual audiodescrit (pel·lícules, sèries, obres de teatre, etc.)? *

- o Sí
- o No

Si n'havies vist, de quin tipus?

(Pots marcar més d'una resposta)

- o Pel·lícules
- o Sèries
- o Dibuixos animats
- o Obres de teatre
- o Òperes

Has creat mai una audiodescripció? *

- o Sí
- o No

Si n'has creat, digues si ha estat:

- o Com a audiodescriptor professional
- o Com a estudiant d'audiodescripció
- o Altres

English translation

* Mandatory field

Identifier *

Enter your initials (first name initial, first surname initial and second surname initial) followed by your age. Do not leave any blank space in between.

Age *

Sex *

- Male
- Female

Level of studies reached *

- Lower than first stage of secondary school
- Secondary education, first stage
- Secondary education, second stage
- Advanced vocational education
- First cycle university education (diploma, degree, engineering or graduate studies)
- Second cycle university education (master, postgraduate or doctoral studies)

In case you have reached university education, please specify.

Occupation *

- Public administration management and management of companies with 10 or more wage earners.
- Management of companies with less than 10 wage earners
- Management of companies without wage earners
- Professions associated with 2nd and 3rd cycle university degrees and the like
- Professions associated with a 1st cycle university degree and the like
- Support technicians and professionals
- Administrative type employees
- Catering services workers and personal services workers
- Protection and security service workers
- Retail workers and the like
- Workers skilled in agriculture and fishing
- Skilled construction workers, except machinery operators
- Skilled workers in the extractive industry, metallurgy, construction of machinery and related trades.
- Skilled workers from the graphic arts, textile and tailoring, elaboration of food, cabinetmakers, craftspersons and other similar industries
- Fixed machinery and industrial installation operators; fitters and assemblers.
- Mobile machinery drivers and operators
- Unskilled workers in the service sector (except transports)
- Agriculture, fishing, construction, manufacturing industries and transport labourers.
- Armed forces
- Unemployed for longer than one year
- Unemployed, seeking a first job

Occupation in your own words *

Years of experience in the translation field *

Specialisation in the translation field *

Have you ever translated an audiovisual product (films, series, documentaries)? *

- Yes
- No

  If you have, indicate if you translated them:

  - As a professional audiovisual translator
  - As an audiovisual translation student

- o Other

Had you ever seen an audio described audiovisual product (films, series, theatre plays, etc.)? *

- o Yes
- o No

If you had, which kind were they?

(You can select more than one answer)

- o Films
- o Series
- o Cartoons
- o Theatre plays
- o Operas

Have you ever created an audio description? *

- o Yes
- o No

If you have, indicate if you created them:

- o As a professional audio describer
- o As an audio description student
- o Other

Annex 4.11. MT pre-test assessment form

* Necessari

Identificador *

Introdueix les teves inicials (la del nom, la del primer cognom i la del segon cognom) seguit de la teva edat. No hi deixis cap espai entremig.

T'ha costat gaire aprendre a fer servir l'eina de postedició? *

- o Sí
- o No

Indica aproximadament quanta estona has estat aprenent com funcionava. *

- o Menys de 15 minuts
- o Entre 15 i 30 minuts
- o Entre 30 i 45 minuts
- o Entre 45 i 60 minuts
- o Més de 60 minuts

T'ha semblat clara la manera en què s'han plantejat les preguntes d'avaluació dels segments? *

Per exemple: The translation text required no post-editing.

- o Sí
- o No

T'ha semblat clara la manera en què s'han plantejat les respostes a les preguntes d'avaluació dels segments? *

Strongly agree, Agree, Neutral, Disagree, Strongly disagree

- o Sí
- o No

Creus que les quatre preguntes eren necessàries? *

How much post-editing was necessary to fix the translation, How difficult the post-editing necessary to fix the translation was, How much of the meaning expressed in the source was also expressed in the target translation, To what extent the translation was grammatically well formed and experienced as using natural/intuitive language by a native speaker without taking into account the correctness of the information

- o Sí
- o No

En cas que no, indica de quina o quines prescindiries.

Pots marcar més d'una casella.

- o How much post-editing was necessary to fix the translation
- o How difficult the post-editing necessary to fix the translation was
- o How much of the meaning expressed in the source was also expressed in the target translation
- o To what extent the translation was grammatically well formed and experienced as using natural/intuitive language by a native speaker without taking into account the correctness of the information

T'hauries estimat més que les respostes es presentessin en forma de categories amb etiquetes concretes per a cada pregunta? *

Per exemple: How much post-editing was necessary to fix the translation? 4. No modification was performed 3. A little post editing wasneeded to fix small problems 2. A lot of post editing was needed (but it was still quicker than retranslation) 1. It required complete retranslation

- o Sí
- o No

Creus que presentar les preguntes d'avaluació després de cada segment fa que et desconcentris a l'hora de fer la postedició? *

Per exemple: How much post-editing was necessary to fix the translation? 4. No modification was performed 3. A little post editing was needed to fix small

problems 2. A lot of post editing was needed (but it was still quicker than retranslation) 1. It required complete retranslation

- o Sí
- o No

T'hauries estimat més que les preguntes es presentessin quan haguessis acabat la postedició sencera? *

- o Sí
- o No

Tenies por de ser massa taxatiu en les respostes i preferies ser menys extremista? *

- o Sí
- o No

En cas afirmatiu, per què?

La prova t'ha semblat massa llarga? *

- o Sí
- o No

Altres comentaris o observacions que vulguis fer.

English translation

* Mandatory field

Identifier *

Enter your initials (first name initial, first surname initial and second surname initial) followed by your age. Do not leave any blank space in between.

Was it difficult for you to learn how to use the post-editing tool? *

- o Yes
- o No

Indicate approximately how long did it take you to learn how to use it. *

- o Less than 15 minutes
- o Between 15 and 30 minutes
- o Between 30 and 45 minutes
- o Between 45 and 60 minutes
- o More than 60 minutes

Do you think the way in which the questions were formulated for the segments' assessment was clear? *

For example: The translation text required no post-editing.

- o Yes
- o No

Do you think the way in which the answers to the questions for the segments' assessment were formulated was clear? *

Strongly agree, Agree, Neutral, Disagree, Strongly disagree

- o Yes
- o No

Do you think that all four questions were necessary? *

How much post-editing was necessary to fix the translation, How difficult the post-editing necessary to fix the translation was, How much of the meaning

expressed in the source was also expressed in the target translation, To what extent the translation was grammatically well formed and experienced as using natural/intuitive language by a native speaker without taking into account the correctness of the information

- o Yes
- o No

In case you think they were not, indicate which one or which ones you would omit.

You can select more than one answer.

- o How much post-editing was necessary to fix the translation
- o How difficult the post-editing necessary to fix the translation was
- o How much of the meaning expressed in the source was also expressed in the target translation
- o To what extent the translation was grammatically well formed and experienced as using natural/intuitive language by a native speaker without taking into account the correctness of the information

Would you prefer that the answers would have been presented in the form of categories with concrete labels for each question? *

For example: How much post-editing was necessary to fix the translation? 4. No modification was performed 3. A little post editing wasneeded to fix small problems 2. A lot of post editing was needed (but it was still quicker than retranslation) 1. It required complete retranslation

- o Yes
- o No

Do you think that presenting the assessment questions after each segment makes you lose concentration when post-editing? *

For example: How much post-editing was necessary to fix the translation? 4. No modification was performed 3. A little post editing was needed to fix small

problems 2. A lot of post editing was needed (but it was still quicker than retranslation) 1. It required complete retranslation

- o Yes
- o No

Would you prefer that questions would have been presented once you had finished the whole post-editing? *

- o Yes
- o No

Were you afraid of being too restrictive when answering and preferred to be less radical? *

- o Yes
- o No

In case you were, why?

Do you think the test was too long? *

- o Yes
- o No

Other comments or observations that you would like to add:

Annex 5: Documents related to the MT main experiment on the comparison of AD creation, translation and PE efforts

Annex 5.1. Clips' AD scripts

CLIP A

1. Dan is waiting in the lobby of an opera house after the performance has started. A man and a woman rush in late. Another man rushes out clutching a hanky to his mouth as he coughs. Dan paces on the red carpeted floor, which is illuminated by a brightly lit chandelier. He looks down over a circular balcony and sees Anna rushing in.
2. She hurries across the foyer and through a door. Then arrives upstairs in the balcony. She and Dan face each other across the circular void. Then move in separate directions to try and meet up. Dan stops and Anna comes to greet him.
3. Anna and Dan enter the bar.
4. She strokes Dan's butt.
5. He stops studying his program and smiles at her.
6. Dan gives her an uncertain smile and pecks her on the lips.
7. Anna looks pensive and uneasy. Dan pays for the drinks.
8. Dan carries the drinks over to a table. He and Anna sit down.
9. She looks triumphant. Dan beams and takes her hand.
10. They kiss.
11. Dan kisses the palm of her hand.
12. Holding her hand he gazes at her adoringly. He then gets up.
13. She pulls him back by the hand and offers her lips. He bends and kisses her tenderly.
14. Dan gives her a lingering look. Then letting her fingers slip from his, he turns and heads out. Anna beams and watches him go. Her smile fades when he's left the room and she remembers her lunch with Larry.

CLIP B

1. Dan leads Alice into Anna's studio. He grabs his jacket and pulls it on as they find her in the kitchen.
2. Anna steps awkwardly into the doorway clutching a cup.
3. Alice turns and swaggers off. Dan, hands in pockets, fixes his gaze on Anna and shifts from side to side.
4. Frustrated, he glances towards the loo.
5. Dan encircles an arm around Alice.
6. Anna looks pointedly at Dan, who turns to Alice.
7. He drags on a cigarette and steps forward to shake Anna's hand.
8. He looks lingeringly at her and swallows. Finally he turns and walks away. Anna looks thoughtful as Dan strides hurriedly out.
9. Alice studies some of Anna's photographs, her hands on her hips.
10. Anna ducks back into the kitchen putting down the cup.
11. Anna glances over at Alice, as both women move around in different areas of the studio.
12. Anna looks round at Alice.
13. Alice is looking at a black and white photo of an elderly couple. She puts it down and tucking a loose strand of hair behind her ear she goes to sit down on the stool in front of the screen. She sits with her back straight, her hands resting primly in her lap. Anna stands close and starts to snap some photographs.
14. Alice regards Anna steadily.

CLIP C

1. Leaving the building.
2. Dan raises a hand hailing a cab.
3. He rubs her back. Then, as the cab pulls up, he opens the door for her. She looks back at him plaintively as she goes to get in.
4. She gives him a kiss reluctant to part. Then gets into the back of the cab.
5. Dan closes the door and waves briefly as it pulls away.

6. Another taxi pulls up at Dan, still staring after the cab Alice left in. Distracted, he glances towards the cabby.
7. He looks down to the ground.
8. The taxi goes away. Back inside Dan hovers at the entrance to the exhibition.
9. Anna looks up and sees Dan. She wanders over.
10. His holdall. He drags on a cigarette.
11. He drops his fag stubbing it out with his foot.
12. Inside the exhibition hall Larry is talking to a man and a woman. He peers past the woman seeing Dan and Anna at the top of the stairs talking.
13. He watches them with a scowl. Anna glances round. She and Dan pace.
14. Anna looks anxiously over her shoulder. Larry has his surly gaze fixed on them as they continue to talk. Dan moves in front of Anna.
15. Their faces are inches apart.
16. Her large eyes are full of anguish as Dan walks away. She stands with her back to him. He suddenly whips back round.
17. She glances sharply round at him. He stares at her through steely eyes. Then wanders away without looking back. Sauntering out of the building, he walks to the curve and hails a cab. Pensive he stares down at the pavement, then glances over his shoulder.

Annex 5.2. Pre-questionnaire

* Mandatory field

Identifier *

Introduce your initials (first name, first surname and second surname) followed by your age. Do not leave any blank space in between.

Rate the following tasks.

MT AD: machine translated audio description

Rate the tasks according to the effort you think they will involve to you, from least effort (1) to most effort (10). *

	1	2	3	4	5	6	7	8	9	10
AD creation	○	○	○	○	○	○	○	○	○	○
AD translation	○	○	○	○	○	○	○	○	○	○
MT AD postediting	○	○	○	○	○	○	○	○	○	○

Rate the tasks according to how much you think they will impair creativity, from least creativity impairment (1) to most creativity impairment (10). *

	1	2	3	4	5	6	7	8	9	10
AD	○	○	○	○	○	○	○	○	○	○

	1	2	3	4	5	6	7	8	9	10
creation										
AD translation	○	○	○	○	○	○	○	○	○	○
MT AD postediting	○	○	○	○	○	○	○	○	○	○

Rate the tasks according to how much you think they will be boring, from least boring (1) to most boring (10). *

	1	2	3	4	5	6	7	8	9	10
AD creation	○	○	○	○	○	○	○	○	○	○
AD translation	○	○	○	○	○	○	○	○	○	○
MT AD postediting	○	○	○	○	○	○	○	○	○	○

Rate the tasks according to how much you think they will convey calques, from least calque conveying (1) to most calque conveying (10). *

	1	2	3	4	5	6	7	8	9	10
AD creation	○	○	○	○	○	○	○	○	○	○
AD translation	○	○	○	○	○	○	○	○	○	○

	1	2	3	4	5	6	7	8	9	10
MT AD posteditng	○	○	○	○	○	○	○	○	○	○

Rate the tasks according to the quality you think they will reach, from best quality (1) to worst quality (10). *

BE CAREFUL! Best quality should be rated 1!

	1	2	3	4	5	6	7	8	9	10
AD creation	○	○	○	○	○	○	○	○	○	○
AD translation	○	○	○	○	○	○	○	○	○	○
MT AD posteditng	○	○	○	○	○	○	○	○	○	○

Rate each task by how much you agree with the following statements and explain why.

Translating ADs created in other languages is useful. *

	1	2	3	4	5	
Strongly disagree	○	○	○	○	○	Strongly agree

In which sense do you think it would or would not be useful? *

Machine translating ADs created in other languages and postediting them conveniently is useful. *

	1	2	3	4	5	
Strongly disagree	○	○	○	○	○	Strongly agree

In which sense do you think it would or would not be useful? *

You can add any further comments here (optional)

Annex 5.3. Post-questionnaire

* Mandatory field

Identifier *

Introduce your initials (first name, first surname and second surname) followed by your age. Do not leave any blank space in between.

Rate the tasks and explain why.

MT AD: machine translated audo description

Rate the tasks according to the effort they have involved to you, from least effort (1) to most effort (10). *

	1	2	3	4	5	6	7	8	9	10
AD creation	○	○	○	○	○	○	○	○	○	○
AD translation	○	○	○	○	○	○	○	○	○	○
MT AD postediting	○	○	○	○	○	○	○	○	○	○

Why do you think so? *

Rate the tasks according to how much you think they have impaired creativity, from least creativity impairment (1) to most creativity impairment (10). *

	1	2	3	4	5	6	7	8	9	10
AD creation	○	○	○	○	○	○	○	○	○	○
AD translation	○	○	○	○	○	○	○	○	○	○
MT AD postediting	○	○	○	○	○	○	○	○	○	○

Why do you think so? *

Rate the tasks according to how much you think they are boring, from least boring (1) to most boring (10). *

	1	2	3	4	5	6	7	8	9	10
AD creation	○	○	○	○	○	○	○	○	○	○
AD translation	○	○	○	○	○	○	○	○	○	○
MT AD postediting	○	○	○	○	○	○	○	○	○	○

Why do you think so? *

Rate the tasks according to how much you think they convey calques, from least calque conveying (1) to most calque conveying (10). *

	1	2	3	4	5	6	7	8	9	10
AD creation	○	○	○	○	○	○	○	○	○	○
AD translation	○	○	○	○	○	○	○	○	○	○
MT AD postediting	○	○	○	○	○	○	○	○	○	○

Why do you think so? *

Rate the tasks according to the quality you think they have reached, from best quality (1) to worst quality (10). *

BE CAREFUL! Best quality should be rated 1!

	1	2	3	4	5	6	7	8	9	10
AD creation	○	○	○	○	○	○	○	○	○	○
AD translation	○	○	○	○	○	○	○	○	○	○
MT AD postediting	○	○	○	○	○	○	○	○	○	○

Why do you think so? *

You can add any further comments here (optional)

Annex 5.4. AD creation post-task questionnaire

* Mandatory field

Identifier *

Introduce your initials (first name, first surname and second surname) followed by your age. Do not leave any blank space in between.

Rate the task by how much you agree with the following statements.

The clip was easy to audio describe. *

	1	2	3	4	5	
Strongly disagree	○	○	○	○	○	Strongly agree

Why do you think so? *

You can add any further comments here (optional)

Annex 5.5. AD translation post-task questionnaire

* Mandatory field

Identifier *

Introduce your initials (first name, first surname and second surname) followed by your age. Do not leave any blank space in between.

Rate the task by how much you agree with the following statements and explain why.

The source text was easy to translate. *

	1	2	3	4	5	
Strongly disagree	○	○	○	○	○	Strongly agree

Why do you think so? *

The clip was easy to audio describe departing from the original AD. *

	1	2	3	4	5	
Strongly disagree	○	○	○	○	○	Strongly agree

Why do you think so? *

Are there any elements you have had to adapt from the English AD taking into account the Catalan AD style? *

○ Yes

○ No

Which ones?

- [ ] Amount of information
- [ ] Length of descriptions
- [ ] Frequency of descriptions
- [ ] Number of incomplete sentences (with no verb)
- [ ] Register (too formal or too colloquial)
- [ ] Other:

You can add any further comments here (optional)

Annex 5.6. AD post-editing post-task questionnaire

* Mandatory field

Identifier *

Introduce your initials (first name, first surname and second surname) followed by your age. Do not leave any blank space in between.

Rate the task by how much you agree with the following statements.

The machine translated text required no postediting. *

	1	2	3	4	5	
Strongly disagree	○	○	○	○	○	Strongly agree

The machine translated text was easy to postedit. *

	1	2	3	4	5	
Strongly disagree	○	○	○	○	○	Strongly agree

All the information in the source text was present in the machine translated text. *

	1	2	3	4	5	
Strongly disagree	○	○	○	○	○	Strongly agree

The machine translated text was fluent Catalan. *

1 2 3 4 5

Strongly disagree ○ ○ ○ ○ ○ Strongly agree

The clip was easy to audio describe departing from the machine translated AD. *

1 2 3 4 5

Strongly disagree ○ ○ ○ ○ ○ Strongly agree

Why do you think so? *

Are there any elements you have had to adapt from the machine translated AD taking into account the Catalan AD style? *

○ Yes

○ No

Which ones?

☐ Amount of information

☐ Length of descriptions

☐ Frequency of descriptions

☐ Number of incomplete sentences

☐ Register (too formal or too colloquial)

☐ Other:

You can add any further comments here (optional)

Annex 5.7. Information sheet

FULL D'INFORMACIÓ SOBRE L'EXPERIMENT

L'objectiu de l'estudi és avaluar l'esforç de creació d'una audiodescripció en català, de traducció d'una audiodescripció anglesa al català i de postedició de la traducció automàtica d'una audiodescripció anglesa al català.

La vostra participació en l'experiment serà la següent: haureu d'audiodescriure un clip en català, traduir l'audiodescripció d'un segon clip de l'anglès al català i posteditar la traducció automàtica al català de l'audiodescripció anglesa d'un tercer clip seguint les instruccions que us donarem.

La vostra participació és totalment voluntària i us podeu retirar de l'estudi en qualsevol moment sense haver-ho de justificar de cap manera ni sense que això us afecti negativament de cap manera.

Les vostres dades seran totalment anònimes ja que al full escrit només us demanarem que hi escriviu el vostre sexe (home/dona), l'edat, els estudis previs i la professió.

No rebreu cap compensació econòmica per aquesta participació, la qual en cap cas les farà servir amb ànim de lucre.

Les investigadores que duran a terme l'experiment són l'Anna Fernández Torné (que participarà en les proves) i la Dra. Anna Matamala (que dirigeix la investigació). Aquesta darrera és la responsable i la podeu trobar a anna.matamala@uab.cat, al despatx K-2012 de la UAB o bé al telèfon 935811646.

Si voleu continuar informats de les evolucions de la investigació, només cal que ho demaneu a la investigadora principal al correu electrònic anterior.

MOLTES GRÀCIES per la vostra participació!

(Nota: a l'altra cara d'aquest full hi teniu el document de consentiment informat)

English translation

INFORMATION SHEET

The aim of this study is to assess the effort of creating an audio description in Catalan, of translating an English audio description into Catalan and of post-editing the Catalan machine translation of an English audio description.

Your participation in the experiment will be as follows: you will have to audio describe a clip in Catalan, to translate the audio description of a second clip from English into Catalan and to post-edit the Catalan machine translation of the English audio description of a third clip following the instructions we will give to you.

Your participation is absolutely voluntary and you may withdraw from the study at any time without further justification and without that negatively affecting you in any way.

Your personal details will be completely anonymous since in the written form you will only be asked to indicate your gender (man/woman), your age, your education and your profession.

You will receive no economic compensation for your participation, which will not be used for profit under any circumstance.

The researchers who will conduct the experiment are Anna Fernández Torné (who will take part in the tests) and Dr. Anna Matamala, who is the supervisor and responsible for this research. You can contact her at anna.matamala@uab.cat, or by phone at 935811646. She can also be found at the office K-2012 in the UAB.

If you would like to be informed of the research progress, please ask the principal investigator by contacting her at the previous email address.

Thank you very much for your participation!

(Note: you will find the informed consent form on the other side of this page)

Annex 5.8. Informed consent form

CONSENTIMENT INFORMAT

Títol de l'experiment: Avaluació subjectiva dels esforços de creació, traducció i post-edició d'audiodescripcions

Jo, (NOM I COGNOMS) _____

- He llegit el full d'informació que m'han donat com també el de consentiment informat.
- He rebut prou informació sobre l'estudi i l'he entesa.
- He pogut fer preguntes sobre l'estudi.

He parlat amb: Anna Fernández Torné

- Entenc que la meva participació és voluntària
- Entenc que la meva participació és no remunerada
- Entenc que la meva informació serà confidencial
- Entenc que em puc retirar de l'estudi quan vulgui i sense haver de donar explicacions ni sense que hi hagi repercussions negatives.

Dono lliurament la meva conformitat per poder participar en l'estudi.

Data:

Firma de la investigadora Firma del participant

Anna Fernández Torné

(Nota: a l'altra cara d'aquest full hi teniu el document d'informació)

English translation

INFORMED CONSENT

Title of the experiment: Subjective assessment of the audio description creation, translation and post-editing efforts

(NAME AND SURNAME) _____

- I have read the information sheet I have been given, as well as the informed consent form.
- I have received enough information about the study and I have understood it.
- I have been able to make questions about the study.

I have talked to: Anna Fernández Torné

- I understand that my participation is voluntary
- I understand that my participation is unpaid
- I understand that my personal details will be treated confidentially
- I understand that I can withdraw from the study whenever I want giving no particular reason for it and having no negative repercussions.

I freely consent to participate on this research study.

Date

Signature of researcher Signature of participant

Anna Fernández Torné

(Note: you will find the information sheet on the other side of this page)