Audio Subtitles or Spoken Subtitles/Captions: An ecological media accessibility service

Pilar Orero, Universitat de Barcelona, Mario Montagud, i2CAT Foundation & Universitat de València Jordi Mata, CCMA, Enric Torres, Anglatècnic, Anna Matamala, Universitat Autònoma de Barcelona,

Abstract

Subtitles are the most common and versatile access media service across platforms, languages, and technologies. Subtitles, which traditionally were between languages, have now expanded to same language subtitles, subtitles for the deaf and hard of hearing, and verbatim transcriptions. The production of subtitles has also changed greatly from the manual production with a subtitle editor, to respeaking, stenography, velotype, and the many possibilities offered by machine translation and crowdsourcing. Subtitles can be read on the main or on a secondary screen, in a wearable, as the mobile phone, intelligent glasses or a head mounted device for XR (eXtended Reality) environments. Finally, subtitles can become a hybrid access service when are generated with Easy-to-Read guidelines, or are read aloud. This last case is the objective of study in this article: spoken or audio subtitles (AST).

In the first section, an introduction will serve as a state of the art regarding audio subtitle literature in the field of Audiovisual Translation or Media Accessibility. Most research has focused at audio subtitles for movies from a descriptive approach, with a growing interest in multilanguage movies. Little has been written regarding the technology towards the generation of audio subtitles and the many workflows allowing for a versatile audio subtitle service in many presentations. This will be the content of the second part of the article. After looking at some possible use cases, it will look at existing solutions for three very different situations. The third part of the article will summarise existing requirements looking to widen the use cases, taking advantage of existing technology. The final section will look at possible future applications, and avenues for research.

Keywords: Spoken subtitles, Audio subtitles, subtitles, accessibility, HbbTV

Introduction

Audio subtitling (AST), or spoken subtitles/captions) is defined by ISO/IEC 20071-25 as "captions/subtitles that are read aloud over the audio in a video". Audio subtitling is an ecological access service, since it gives another use and recycles existing subtitling assets. According to the European Broadcasting Union (EBU) published Access Service Survey (EBU 2016), audio subtitles is the least deployed access service in Europe (see Fig. 1). The difference in quantity between subtitling and audio subtitling is remarkable.



Figure 1. Four access services on offer across 36 public broadcasters in Europe (EBU 2016)

One of the features defining Europe is the many languages spoken at official (24) and unofficial levels (over 200). When Europeans consume foreign media content, that is media content in a different language from that spoken domestically, they do so through three very different translation modalities: dubbing, subtitling, and voice-over (see Fig.2). The reasons and impact for the choice of translation modality vary from political (Kilborne 1993, McLaughlin and Múñoz-Basols 2016), or as explained in (Rupérez, Bris and Banal-Estañol 2015: 2)

dubbing tended to be adopted in larger countries, those whose national languages were widely used internationally, and those with dictatorial regimes. Smaller, democratic countries tended to adopt subtitling. No OECD country has moved from one to the other mode since WWII.

The different translation modalities also have a direct implication in education, with subtitling countries having better English skills than dubbing countries, and "subtitling, through better English skills, increases foreign direct investment and exports." (Rupérez, Bris and Banal-Estañol 2015: 17). The choice of translation modality is also related to media genres. For example, documentaries are translated with voice-over and off-screen dubbing (Matamala 2009, Franco, Matamala and Orero 2012). The type of audience also impacts on the translation modality chosen. For example, children programmes must be dubbed in order to secure access to children under reading age. This is a legal requirement in Europe from the Audiovisual Media Directive (AVMSD)¹. It is also a requirement from AVMSD to secure access to content to all citizens, and when content is in a different language with subtitles, those who cannot read are denied access. A simple solution to secure accessible subtitling is to turn written text from subtitles into spoken language, that is spoken or audio subtitles. Transforming subtitles into audio subtitles is a technology based solution, whose cost has yet to be determined. Still, financial burden of making subtitles into audio subtitles is outnumbered by the social impact. A European map of audiovisual translation should be the departing point to link audio subtitling to subtitling countries, as can be seen in Fig. 2.

¹ <u>https://ec.europa.eu/digital-single-market/en/revision-audiovisual-media-services-directive-avmsd</u>



Figure 2. Map of audiovisual translation modality in Europe (Media Consulting Group 2008)

The countries in yellow in the map from Fig.2 should correspond to the countries offering audio subtitling, as identified in Figure 3 from the EBU 2016 survey (EBU 2016).



Figure 3. European countries where broadcasters use audio subtitling (EBU 2016)

Four countries challenge the correlation between subtitles and audio subtitles: Spain, Greece, Portugal and the United Kingdom (UK). The first is Spain (see Fig.4), a dubbing country, where the Catalan public broadcaster TVC offers audio subtitling. This exceptional case can be read later in the section of this article dedicated to solutions. The other three are subtitling countries: Greece did not take part of the survey, and Portugal (RTP broadcaster) and UK (BBC broadcaster) do not offer AST yet.



Figure 4. Number of countries offering audio subtitling (EBU 2016)

Another interesting data gathered from the EBU survey (EBU 2016) related to AST is related to the delivery platforms. All broadcasters deliver content in real time, and also on demand (VoD), via their website. To date most access services delivered over broadcast are not offered for VoD, even though the content is already produced and paid for. This situation is soon to be changed in Europe, as it was changed in 2015 in USA as a result from the NAD vs NETFLIX law suit². In USA the Federal disability laws, as the EU AVMSD, follows a comprehensive update to keep pace with the digital world. The Americans with Disabilities Act (ADA) Section 508 of the Rehabilitation Act got an Information and Communications Technologies (ICT) refresh in 2015³ to reflect best practices from the World Wide Web Consortium (W3C) standard for Web Accessibility and Inclusive Design (WCAG 2018)⁴. Hence, it will be interesting to see how the data from Fig. 5 evolves in two years from now.

² <u>https://www.3playmedia.com/2015/07/23/nad-v-netflix-ada-lawsuit-requires-closed-captioning-on-streaming-video/</u>

³ <u>https://www.3playmedia.com/2017/12/21/the-ultimate-section-508-refresh-checklist/</u>

⁴ <u>https://www.3playmedia.com/2015/07/23/nad-v-netflix-ada-lawsuit-requires-closed-captioning-on-streaming-video/</u>

2 broadcasters	Audio subtitling services availability by platform
have made	OWN WEBSITE
audio subtitles	2
available via an	SOCIAL NETWORKS
online	0
platform.	CONNECTED TV SERVICES (HbbTV)
plationin	1
3 broadcasters are working on providing this in 2016.	MOBILE DEVICES
	NOT AVAILABLE ONLINE
	1
	INTENDS TO PROVIDE ONLINE AUDIO SUBTITLING IN 2016
	3
	Based on 6 broadcasters.
(1) (2)	
EURIOIVISION	ACCESS SERVICES SURVEY 2016

Figure 5. Platforms for audio subtitle delivery (EBU 2016)

It is also interesting to note that not all the content subtitled is later audio subtitled, as can be seen in Fig. 6.





This state of the art will change in the near future due to the publication of the revised EU AVMSD in 2018, which creates the media framework for the 21st century. This revised legal framework establishes new obligations along reporting conditions. In Europe, media accessibility will now move fast from "raising awareness" to full implementation. Among the many must carry obligations, linear and non-linear access services will have to be offered by both public and commercial broadcasters. It also extends to video sharing platforms with services such as NETFLIX, YouTube or social media like Facebook. This brings to the fore the fact that while some traditional accessibility services, such as subtitling or sign language interpreting, are well known perhaps due to its visibility on the screen, other services such as audio description (AD) or audio/spoken subtitles are still in their infancy.

To conclude, it is also interesting to highlight that multilingual productions are increasing and subtitles may be more present (Heiss 2004). Therefore, a situation that can easily take place

is an original production, or its dubbed version, with partial subtitling. In these cases, spoken subtitles should also be considered.

Audio subtitling research in Audiovisual Translation

In the field of Audiovisual Translation, audio subtitles (AST) have been studied since 2006 when Orero (2007) reports of using this service for live opera, inspired by Verboom et al. (2002). Opera at Liceu Opera House in Barcelona (Spain) has been audio described for over a decade (Orero et al 2019). The audio description has been provided by the research group TransMedia Catalonia at the Universitat Autònoma de Barcelona (UAB). This direct intervention allowed for testing different audio description styles and approaches (Matamala 2005 and 2007, Matamala and Orero 2007, Orero 2007, Orero and Matamala 2007, Puigdomènech et al 2017 and 2008, Cabeza and Matamala 2008, Cabeza 2011, Corral and Lladó 2011, Miquel-Iriarte et al 2012, Oncins et al 2012).

Opera is rarely sang in translation. Most opera productions are in the language it was written, requiring translation to the language spoken where is represented. It is also the case that sung opera is hard to understand, even if represented in the same language as the language performed (Orero and Matamala 2007, Matamala and Orero 2008). Most opera houses display the translation as surtitles or subtitles (Oncins 2012). The first audio subtitling test in Liceu was an orchestra version of Donizetti's Roberto Devereux. The audio subtitles were read by a human describer. The regular audio description audience had been informed of this new change in style, and the feedback before the performance was not positive. Persons with sight loss pointed out the negative effect of hearing the voice of the describer on top of the opera signing, with the effect of voice-over. After the performance, a satisfaction questionnaire was administered, and results showed full acceptance of this new service. Most blind and partially sighted audience interviewed were not aware of the subtitling service at Liceu. Their way to attend an opera performance was with audio description, but they missed the text of the lyrics. They did not know they could enjoy the translation of what was being sung into their own language. After Roberto Devereux, a new AD style (Puigdomènech et al 2008) was taken into consideration for Liceu, including AST complementary to AD. It was found that apart from giving access to the surtitled text to blind and partially sighted audiences, audio subtitling has many other advantages for opera. It mitigates the split attention effect of looking at one screen for the subtitles while the action takes place in a different view range. It also offers voiced subtitles for those with reading issues. A final advantage is that it can be delivered by a synthetic voice. Some years later, the same UAB team suggested AST to be produced at the client's end through a smartphone (Miguel-Iriarte et al 2012, Oncins 2012). AST through a companion screen were finally tested in 2016 (Orero et al 2019).

Away from live opera audio subtitles, Braun and Orero (2010) studied the practice of AD in combination with AST as a localisation strategy to make foreign films accessible. Audio subtitling was characterised as being situated at the interface of subtitling, audio description and voice-over, also with some features from dubbing. Through an analysis of a limited body of movies, some AST strategies were discussed such as "different methods of assigning voices to ASTs (and the accompanying AD narrative), different techniques of delivery (various types of voiceover), an integration of ASTs with the AD narrative and an adaptation of the AD narrative itself in order to explicate, for example, who is speaking and/or addressed." (ibid 186). Braun and Orero (2010) open the line of research for many other studies on the same topic with different methodology and languages such as: Korean (Orero 2011), German (Benecke 2012), Dutch (Remael 2012, Reviers and Remael 2015), and English (Matamala 2014). In the same line of multilingualism movies, Iturregui-Gallardo (2018) shows the different possibilities that ASTs offer when revealing the multilingual reality of audiovisual contents. More specifically, it reproduces for AST the categorisation created by Szarkowska, Zbikowska, and Krejtz (2013) for the subtitles for the deaf and hard of hearing (SDH). In all previous descriptive studies AST are adapted and delivered by humans.

Another line of research directly linked to AST is that of testing the reception of synthetic voices. This has been tested in AD by Szarkowska and Jankowska (2011), Fernández and Matamala (2015) and Matamala and Ortiz-Boix (2018). And finally Iturregui-Gallardo (2019) will read his PhD on the reception of audio subtitles using biometrics. More specifically, he investigates the emotional arousal produced by two different voicing strategies, namely AST with a dubbing effect and AST with a voice-over effect. While most studies previously mentioned focused on the reception by the blind and partially sighted audiences, Gambier (2012: XX) entry in The Routledge Handbook of Translation Studies widened the audience beyond the sight loss community "Audio subtitling is useful for dyslexic people, elderly people, the partially sighted and anybody who cannot read fast enough".

Against this theoretical framework AST can be delivered without any human intervention transforming subtitle assets into speech in different ways and workflows.

Scenarios and potential solutions for audio subtitles

Previous studies have focused on audio subtitling for opera or films, but from a descriptive theoretical perspective. In addition, there are many other scenarios which may be worth pointing out, since they may require the availability of new workflows, converting this into an interesting research topic worth to explore. In general, the potential scenarios depart from a situation where a person watching audiovisual content cannot understand the language spoken: subtitles are offered. Likewise, other potential scenarios are related to people with reading problems, ranging from visual loss, to age, or issues with reading skills. These scenarios may go from children who cannot read yet --watching content that is not considered to be "children programmes"-- to subtitles delivered so fast that the most proficient reader cannot follow. Relevant situations in which audio subtitles become useful are:

- a. Within news programmes, some interventions are in the original language with subtitles. Unless you can read subtitles, you do not understand what they are saying. An example could be Theresa May explaining the Brexit agreement in Catalan TV.
- b. Watching the news with subtitles over catch up TV or in start over mode.
- c. Watching different language/same language content in a smartphone where subtitles are too small.
- d. Watching different language/same language audiovisual content in a public screen too far away. In this case, subtitles are not legible. Similar examples are live situations, when for example going to a public conference, or in a museum.
- e. Watching a partially subtitled film, that is a film in its original or dubbed version with some subtitled dialogues, as another language is spoken. This may be especially relevant in multilingual productions.

Likewise, technical requirements for AST have been studied in other works. On the one hand, a key challenge consists of interpreting the text from subtitles when they are burnt on the video, by means of Optical Character Recognition (OCR) techniques, like in Jønsson and Bothe (2007). On the other hand, many other works have focused on providing solutions for Text to Speech (TTS) conversion (Bjorn 2010, Folke 2005, Hanzlícek, Matoušek, and Tihelka 2008, Schröder and Trouvain 2003, Verboom et al 2002, Nielson and Bothe 2007, Derbring, Ljunglöf and Olsson 2009, Ljunglöf, Derbring, and Olsson, 2012, Miesenberger, Klaus, and Zagler 2002, Mihkla et al 2014).

Another issue to be taken into consideration regarding technology, beyond TTS or OCR, is the place where the synthetic audio is generated. This may be at the broadcaster (or server, in case of VoD platforms) or at the client side. In general, two solutions can be applied to all previously described cases:

1. Audio is generated at the broadcast/server side. In such a case, the audio can be sent via a broadcast (multiplexed) audio channel (e.g., audio description channel) or via an additional audio channel (e.g., by using a parallel broadband connection,

taking advantage of the Hybrid Broadcast Broadband TV (HbbTV) standard (HbbTV standard). Apart from the audio generation and delivery processes, the mix can be also performed at the origin/server or receiver/client sides of the chain.

- 2. Audio is generated at the client side. In such a case, the audio can be generated from:
 - OCR of burned-in subtitles.
 - Streamed text with subtitles (i.e., separate subtitle tracks).
 - File with audio segments.

The workflow for all these cases may be any of the following.

WORKFLOW A. TTS of existing subtitle files

This workflow is the most commonly used, and it takes advantage of existing subtitle text files to generate the associated audio signal that will be read aloud on the end user device (either on main or companion devices). The TTS conversion can be done at the broadcast or at the client side. An example of the former situation is the YLE solution described in the next section.

WORKFLOW B. OCR for burnt in subtitles

In this case, the use of OCR technology serves to transform burnt in text into a text file that can be read aloud by a TTS engine, like in Workflow A. An example of this is the NPO solution described in the next section.

WORKFLOW C. Use of screen readers

Screen reader features are e.g. commonly available in smartphones, and can be applied to text on screen or subtitle files.

However, to our knowledge, this solution has not been developed for AST yet.

WORKFLOW D. Use of Timed Text Markup Language 2 (TTML2)

Timed Text Markup Language 2 (TTML2) is a new content type for timed text media which includes features to embed audio audio resources, with the associated time codes and metadata. TTML2 file format can have a big potential for both AD and AST.

WORKFLOW E. Use of HbbTV standard

Features from HbbTV standard can be used to deliver either extra audio tracks for TTS generation or already generated TTS at the broadcast side to the targeted clients (either main or companion devices).

WORFLOW F. Automatic translation plus TTS

It consists of performing an automatic translation of text subtitles and then generating the AST via TTS.

The appropriateness of combining the above described workflows, like D and E, or F with other ones, is also worth to explore.

In addition, it should be remarked that the above workflows targeted at the generation of AST must not replace existing workflows for AD, but must complement, and inter-operate with them.

Three existing solutions for broadcast

This section provides three examples of implementations for audio subtitles from three public broadcasters in Europe, retrieving the text from burnt text on the image or from separate files. The first solution is from The Netherlands, a subtitling country. In 2001, the Dutch public broadcasting organization, called NPO, developed a way to convert 'burnt in' subtitles in the video into a separate data stream containing the subtitle text, by using an OCR technique (Verboom et al 2002). This allowed for carrying out / obeying the requirement/pressure of providing 95% of the programs on NPO's premium channels with a subtitle data stream. This automated spoken subtitling service is a cost-effective way of serving the sight disabled audience, compared to audio description that is used in other countries and needing manual authoring before broadcast. Starting with a pilot in 2011, NPO has deployed this OCR-based solution on all three public premium television channels since March 2012. Subtitle data is distributed through a private data feed and teletext page 889 is used to provide client side visual aids with an appropriate signal. In addition, the subtitle data are used as source for a broadcast- mixed speech synthesized audio track that is distributed with the HD channels. Off-air, the output data can also be used for indexation purposes. Soon, NPO concluded that the subtitle data generated by the OCR-based solution system had enough quality to become the only spoke subtitle source in service, as the target audience was very enthusiastic and satisfied with the new deployed service.

The second solution is from the Finish public broadcaster, YLE, who subtitles most of its foreign content. YLE identified the barrier this situation presented to persons who were visually impaired, and developed the complementary accessibility service where subtitles were spoken out loud by a TTS voice synthesizer. The solution takes advantage of the fact that the implementation of these techniques is simpler in digital TV than in analogue systems, and to the improved performance and quality of modern voice synthesizers. YLE has developed a system that can create an additional soundtrack to TV programmes through voice synthesis. This requires a pre-produced subtitling file for the programme. The new synthetized AST track is created by using a parallel process at the broadcast, being also mixed with the original soundtrack at this part of the chain.

The third solution is from the Catalan public broadcaster CCMA, mainly a dubbing broadcaster. In 2015, CCMA decided to change the translation modality in their news programme when the person speaking uses a different language from Catalan. The usual situation is when a politician is doing some declaration in the original language or in an interview. Voice-over was discarded for subtitling, now we can hear the original voice with subtitles in Catalan. This situation generated an issue for the audience who had reading issues. The adopted solution consisted of creating a synthetic voice from the voice used for all documentaries and a some current affairs programmes on a regular reader. This synthetic voice is nowadays used to read subtitles. The solution went beyond accessibility, offering with this personalised voice a TVC brand for AST, different from any other channel in Spain. The final effect to the audience is that the same person who voices documentaries, I also voicing the subtitles with the same effect: a voice-over.

Conclusions

The growing legal pressure in Europe to offer more accessibility in the audiovisual domain has been greeted with interest by stakeholders of the value chain. New services means new costs, and while the reviewed Audiovisual Media Services Directive AVMSD (2018) sets requirements, it does not point to who is responsible to meet them: producers, distributors, or exhibitors. The cost was also identified as the first barrier for media access service implementation by EBU (2016).

The major challenges foreseen are:

1 2 3 4	Cost Engagement of on-line service providers with accessibility Putting access services online Switch-over to digital
Di	fficulties for delivering access services in the next
ye	ars are foreseen to come from financial cuts with 17
br	oadcasters (55%) indicating it will be very or
ex	tremely difficult.

Figure 7. Major challenges for the deployment of access services (EBU 2016)

Subtitling is the accessibility service with highest presence across screens, broadcasters, and VoD distributors. In fact, it is considered that 85% of Facebook media content is consumed with subtitles (Patel 2016). The generation of an automatic accessible service epigonic from subtitling is common sense and cost effective.

This article has shown different ways to generate the service, and also new avenues to deploy audio subtitling. It also closes with the warning about using audio subtitling as a replacement for audio description. This solution will do a disservice to the visual loss community. Audio subtitling should be made available with audio description, as a complement, never as a substitution.

Future work will be targeted at exhaustively analysing the presented workflows for AST generation, and adopt the most appropriate solution within the umbrella of EU H2020 ImAc⁵ (Immersive Accessibility) and EasyTV⁶ projects

Acknowledgment

This article is related to the research carried out in the European funded projects ImAC (GA: 761974) and EasyTV (GA: 761999). Anna Matamala and Pilar Orero are TransMedia Catalonia members (2017SGR113).

References

- 1. Björn, B. 2010. HTML text to speech (TTS) API specification. W3c editor's draft, W3C. Available at <u>https://lists.w3.org/Archives/Public/public-xg-</u>htmlspeech/2010Nov/att-0036/htmltts-draft.html [access January 2019]
- 2. Benecke, B. 2012. "Audio description and audio subtitling in a dubbing country: Case studies" in E. Perego (ed.) *Emerging topics in translation: Audio description.* Trieste, EUT Edizioni Università di Trieste, pp. 99-104.
- 3. Braun, S. and P. Orero. 2008. "Audio Description Research: State of the Art and Beyond". *Translation Studies in the New Millennium* 6, pp. 14–30.
- Cabeza, C. and A. Matamala. 2008. "La audiodescripción de ópera: una nueva propuesta del Liceo" in A. Pérez-Ugena and R. Vizcaino-Laorga (eds) Ulises y la Comunidad Sorda. Madrid: Observatorio de las Realidades Sociales y de la Comunicación, pp. 95-106.

⁵ <u>http://www.imac-project.eu/</u>

⁶ https://easytvproject.eu

- Cabeza, C. 2011. "Opera Audio Description at Barcelona's Liceu Theatre' in Díaz-Cintas, J.; A. Matamala and J. Neves (eds) New Insights into Audiovisual Translation and Media Accessibility. Amsterdam: Rodopi, pp. 227-237.
- 6. Corral, A. and R. Lladó 2011. "Opera Multimodal Translation: Audio Describing Karol Szymanowski's Król Roger for the Liceu Theatre." *JosTrans* 5, pp. 163-179.
- Derbring, S., P. Ljunglöf and M. Olsson 2009. "SubTTS: Light-Weight Automatic Reading of Subtitles." *Nordic Conference on Computational Linguistics*, May 14-19, Odense, Denmark, pp. 272-274.
- European Broadcasting Europe. 2016. Access Services Paneuropean Survey, available https://www.ebu.ch/files/live/sites/ebu/files/Publications/Presentations/EBU%20 Access%20Services%20Survey%202016.pdf f[access January 2019].
- 9. Fernández, A. and A. Matamala. 2015. "Text-to-speech vs human voiced audio description: a reception study in films dubbed into Catalan." *Jostrans. Journal of Specialised Translation* 24, pp. 61-88.
- 10. Filmer, D. 2018. "Voicing diversity? Negotiating Italian identity through voiceover translation in BBC broadcasting." *Perspectives. Studies in Translation*, pp. 299-315.
- 11. Folke, E. 2005. IT i praktiken slutrapport. Hjälpmedelsinstitutet, Sweden.Available at <u>https://www.regionhalland.se/PageFiles/35236/Slutrapport%20och%20utvärder</u> ing%20Hjälpmedel%20möter%20brukare.pdf [access January 2019]
- 12. Franco, E.; A. Matamala, and P. Orero. 2010. Voice-over Translation: An overview. Bern: Peter Lang.
- 13. Gambier, Y. 2012. "The position of audiovisual translation studies," in Millán, C. and F. Bartrina (eds) *The Routledge Handbook of Translation Studies*. Abingdon: Routledge, pp. 45-59.
- 14. Hanzlícek, Z.; J. Matoušek, and D. Tihelka. 2008. "Towards automatic audio track generation for Czech TV broadcasting: Initial experiments with subtitles-to-speech synthesis." ICSP '08, 9th International Conference on Signal Processing, Beijing, China.
- 15. Heiss, C. 2004. "Dubbing Multilingual Films: A new Challenge" *Meta* 49 (1), pp. 208–220.
- 16. HbbTV standard specification, available at https://www.hbbtv.org/resource-library/specifications/ [last Access January 2019]
- 17. Iturregui-Gallardo, G. 2018. "Rendering multilingualism through audio subtitles: shaping a categorization for aural strategies". *International Journal of Multilingualism*. Available at DOI: 10.1080/14790718.2018.1523173 [last accessed November 2019]
- 18. Iturregui-Gallardo, G. 2019. Audio subtitling: Voicing strategies and their effect on emotional activation. Universitat Autònoma de Barcelona, Unpublished PhD thesis.
- 19. Jønsson, M. and H.H. Bothe. 2007. "OCR- algorithm for detection of subtitles in television and cinema", in CVHI'07, 5th Conference and Workshop on Assistive Technology for People with Vision and Hearing Impairments, Granada, Spain.
- 20. Ljunglöf, P., S. Derbring and M. Olsson. 2012. "A free and open-source tool that reads movie subtitles aloud." in *Proceedings of the SLPAT '12 Proceedings of the Third Workshop on Speech and Language Processing for Assistive Technologies*, pp. 1–4.
- 21. Ofcom. 2006. Guidelines on the provision of television access services, available at <u>http://www.ofcom.org.uk/tv/ifi/guidance/tv_access_serv/guidelines/guide</u> <u>lines.pdf</u>. [access January 2019]
- 22. Kilborn, R. 1993. "Speak my language: current attitudes to television subtitling and dubbing". *Media, Culture and Society*, 15(4), pp. 641–660.
- 23. Matamala, A. 2005. "Live Audio Description in Catalonia". *Translating Today* 4, pp. 9-11.
- 24. Matamala, A. 2007. "La audiodescripción en directo", in C. Jiménez (ed.) *Traducción y accesibilidad. Subtitulación para Sordos y audiodescripción para*

ciegos: nuevas modalidades de Traducción Audiovisual. Frankfurt: Peter Lang, pp. 121-132.

- 25. Matamala, A. and P. Orero. 2007a. "Accessible Opera in Catalan: Opera for All" in J. Díaz Cintas, P. Orero and A. Remael (eds) *Media for All: Subtitling for the Deaf, Audio Description and Sign Language*. Amsterdam: Rodopi, pp. 201-214.
- 26. Matamala, A. and P. Orero. 2007b. "Accessible opera: overcoming linguistic and sensorial barriers". *Perspectives. Studies in Translatology*, 15(4), pp. 262-277.
- 27. Matamala, A. 2009. "Translating documentaries: from Neanderthals to the Supernanny". *Perspectives: Studies in Translatology* 17(2), pp. 93-107.
- 28. Matamala, A. 2014. "Audio describing text on screen", in A. Maszerowska, A.Matamala, and P. Orero (eds) *Audio description: New perspectives illustrated.* Amsterdam: John Benjamins Publishing Company, pp. 103-120.
- 29. Matamala, A. and C. Ortiz-Boix. 2018. "Text-to-speech voice-over? A study on user preferences in the voicing of wildlife documentaries". *SKASE. Journal of Translation and Interpreting 1*, pp. 24-38.
- 30. McLaughlin, M. and J. Muñoz-Basols. 2016. "Ideology, censorship and translation across genres: past and present." *Perspectives*, 24 (1), pp. 1-6.
- Media Consulting Group. 2008. Study on dubbing and subtitling needs and practices in the European audiovisual industry, available at: <u>http://www.ltinnovate.org/sites/default/files/documents/1342-</u> <u>Study%20on%20dubbing%20and%20subtitling%20needs%20and%20practice</u> <u>s%20in%20the%20European%20audiovisual%20industry%20(2008).pdf</u> [last Access January 2019].
- 32. Miesenberger, K., J. Klaus, and W. Zagler. 2002. "Computers helping people with special needs." 8th international conference, ICCHP 2002. Linz, Austria: Springer.
- 33. Mihkla, M., I. Hein, I. Kiissel, A. Räpp, R. Sirts, R and T. Valdna. 2014. "A System of Spoken Subtitles for Estonian Television" in A. Utka et al (eds) *Human Language Technologies the Baltic Perspective*, pp 19-26.
- 34. Miquel-Iriarte, M., A. Vilaro, P. Orero, J. Serrano and H. Delgado. 2012. "Entitling: A Way Forward for Accessibility", in E. Perego (ed.), *Eyetracking in audiovisual translation*. Roma, Italy: ARACNE Editrice, pp. 259-276.
- 35. Nielson, S. and H.H. Bothe. 2007. "SubPal: A device for reading aloud subtitles from television and cinema" in CVHI'07, 5th Conference and Work- shop on Assistive Technology for People with Vision and Hearing Impairments, Granada, Spain.
- 36. Oncins, E. 2012. "The tyranny of the tool: Surtitling live performances". *Perspectives. Studies in Translatology* 23 (1), pp. 42-61.
- Oncins, E., O. Lopes, P. Orero, J. Serrano, J. Carrabina. 2012. "Multi language and multi system mobile application to make accessible live performing arts: All Together Now". *JosTrans* 20, pp. 147-164.
- 38. Orero, P. 2007. "Audiosubtitling: A Possible Solution for Opera Accessibility in Catalonia". *TradTerm* 13, pp.135-149.
- 39. Orero, P. 2011. "The audio description of spoken, tactile and written language in Be with me', in A. Matamala, J.M. Lavaur, and A. Serban (eds), *Audiovisual translation in close-up*. Bern: Peter Lang, pp. 239–255.
- 40. Orero, P. and A. Matamala. 2007. "Accessible Opera: Overcoming Linguistic and Sensorial Barriers". *Perspectives. Studies in Translatology* 15 (4), pp. 262-278.
- 41. Orero, P.; G. Iturregui-Gallardo, I.C. Permuy Hércules de Solás, J. Bestard and A. Matamala 2019. "Opera Accessibility in the 21st Century: New services, new possibilities". *TRANS* 23, pp. 243-255.
- 42. Patel, S. 2016. " 85 percent of Facebook video is watched without sound" available at:

43. Puigdomènech, L., A. Matamala and P. Orero. 2010. "Audio description of Films: State of the Art and a Protocol Proposal', in L. Bogucki & K. Kredens

https://digiday.com/media/silent-world-facebook-video/ [last Access January 2019].

(eds), *Perspectives on Audiovisual Translation*, Lodz Studies in Language 20, Frankfurt: Peter Lang, pp. 27-44.

- Puigdomènech, L., P. Orero and A. Matamala. 2008. "The Making of a Protocol for Opera Audio Description," in Pegenaute, L., De Cesaris, J., Tricás, M. and E. Bernal (eds) *La traducción del futuro: mediación lingüística y cultural en el siglo XXI*. Barcelona: PPU, pp. 381-392.
- 45. Remael, A. 2012. "Audio Description with Audio Subtitling for Dutch Multilingual Films: Manipulating Textual Cohesion on Different Levels". *Meta: Translator's Journal* 57(2), pp. 385-407.
- 46. Reviers, N. and A. Remael. 2015. "Recreating multimodal cohesion in audio description: A case study of audio subtitling in Dutch multilingual films." *New Voices in Translation Studies*, 13(1), pp. 50–78.
- 47. Rupérez, A., A. Bris and A. Banal-Estañol. 2015. "TV or not TV? The Economic Impact of Subtitling and English Skills" available at http://albertbanalestanol.com/wp-content/uploads/Subtitles_final.pdf [Last accessed November 2019]
- 48. Schröder, M. and J. Trouvain. 2003. "The German text-to-speech synthesis system MARY: A tool for research, development and teaching." *International Journal of Speech Technology* 6, pp. 365–377.
- 49. Szarkowska, A. and A. Jankowska. 2011. "Text-to-speech audio description: Towards wider availability of AD." *Jostrans. The Journal of Specialised Translation* 15 (1), pp. 81–98.
- Szarkowska, A., J. Zbikowska, and I. Krejtz. 2013. "Subtitling for the deaf and hard of hearing in multilingual films." *International Journal of Multilingualism*, 10 (3), pp. 292–312.
- Verboom, M., Crombie, D., Dijk, E., & Theunisz, M. (2002) 'Spoken Subtitles: Making Subtitled TV Programmes Accessible', pp. 295-302, in K. Miesenberger, J. Klaus, & W. Zagler (Eds.), Computers Helping People with Special Needs. 8th International Conference, ICCHP 2002, July 15-20, Linz, Austria.
- 52. W3C Web Content Accessibility Guidelines (WCAG) Standard, https://www.w3.org/WAI/standards-guidelines/wcag/ Last Access in January 2019.
- 53. W3C Timed Text Markup Language 2 (TTML2). https://www.w3.org/TR/2018/REC-ttml2-20181108/ Last Access in January 2019.

Authors

Pilar Orero, PhD (UMIST, UK) works at Universitat Autònoma de Barcelona (Spain). Member of the research group TransMedia Catalonia works on topic related to media accessibility and alternative communication, focusing on vulnerable groups. Member of the Spanish UNE working group on accessibility. Leader and participant on numerous EU funded projects. E-mail: <u>pilar.orero@uab.cat</u>

Mario Montagud (PhD in Telecommunications) is a Senior Researcher at i2CAT Foundation (Barcelona, Spain) and a Part-Time Professor at the University of Valencia (Spain). His topics of interest include Computer Networks, Interactive and Immersive Media, Synchronization and QoE (Quality of Experience). Mario is (co-)author of over 70 scientific and teaching publications, contributes to standardization, and is member of the Organization Committee of the many international workshops and conferences. He is currently involved in three EU H2020 projects. E-mail: <u>mario.montagud@i2cat.net</u>

Jordi Mata is the head of engineering R&D at the Public Catalan Broadcaster CCMA. He has worked in numerous EU funded projects related to media broadcast and media accessibility such as: DTV4ALL, TVRing, HBB4ALL. He is currently involved in H2020 ImAc and EasyTV. E-mail: <u>imata.n@ccma.cat</u>

Enric Torres CEO at Anglatècnic (http://www.anglatecnic.com) Head of the team of telecommunications engineers and computing engineers. Experts in media accessibility

solutions for broadcasting IT sectors. He is currently involved in H2020 ImAc. E-mail: <u>enric@anglatecnic.com</u>

Anna Matamala, BA in Translation (UAB) and PhD in Applied Linguistics (UPF), is an associate professor at Universitat Autònoma de Barcelona. She currently leads TransMedia Catalonia research group and the Erasmus + project EASIT, and is involved in the organisation of scientific events such as ARSAD. She participates in other European projects such as ImAc or ADLAB PRO, and in standardisation work. E-mail: anna.matamala@uab.cat