

**This is a pre-print copy of the following book chapter:**

Matamala, Anna (2017) "Mapping audiovisual translation investigations: research approaches and the role of technology". M. Deckert (ed) *Audiovisual translation- research and use*. Frankfurt: Peter Lang, 11-28.

Please visit: <https://www.peterlang.com/view/product/80659>

## **Mapping audiovisual translation investigations: research approaches and the role of technology**

Anna Matamala  
Universitat Autònoma de Barcelona

### **Abstract**

This article maps audiovisual translation research by analysing in a contrastive way the abstracts presented at three audiovisual translation conferences ten years ago and nowadays. The comparison deals with the audiovisual transfer modes and topics under discussion, and the approach taken by the authors in their abstracts. The article then shifts the focus to the role of technology in audiovisual translation research, as it is considered an element that is impacting and will continue to impact both research and practice in this field. Relevant research in audio-related, text-related and image-related technologies applied to audiovisual translation is summarised. The last section briefly discusses how technological tools can also help audiovisual translation professionals, users and researchers.

### **0. Introduction**

Research in audiovisual translation (AVT) is thriving. A lot has been done in recent years, and a lot can be done in the near future. The continuous transformation of a society where audiovisual content is ubiquitous, technology is paramount and citizens are becoming netizens impacts directly on AVT practices and, by extension, on AVT research. The interest in audiovisual transfer modes such as dubbing, subtitling or voice-over has expanded to include media accessibility related modalities such as audio description (AD), audio subtitling, subtitling for the deaf and hard of hearing (SDH) or sign language interpreting in the media. Long-established descriptive approaches are supplemented by applied investigations, by technologically-based research and by empirical studies on the process and on the end-users' reception. Individual scholars are giving way to teams of researchers with different backgrounds joining efforts at an international level in order to approach AVT from an interdisciplinary perspective and gain a better understanding of the field. And all this research often impacts directly on society, because researchers are more and more involved in knowledge transfer activities.

In this article I aim to take stock of what has been done in recent years in AVT research and suggest future avenues, focusing specifically on the role of technology. First, I will try to elucidate whether the optimistic tone of the first paragraph correlates with reality. This is why I will begin the article by looking back into the history of AVT research and comparing the picture I obtain to the present situation. This broad image will give way to a close-up in which the focus will be AVT research in relation to technology, as I consider it can have a remarkable impact on the field.

### **1. Mapping audiovisual translation research**

To gain an overview of the evolution of the field, the abstracts from three seminal conferences on AVT held about ten years ago and three recent conferences were collected. The events under analysis were the following:

- In So Many Words, 2004, University of Surrey (London) (ISMW).

- Between Text and Image. Updating Research in Screen Translation, 2005, University of Bologna at Forlì (BTM).
- Media for All conference, 2005. Universitat Autònoma de Barcelona (M4A).
- Intermedia, 2016, University of Lodz (INTER).
- Linguistic and cultural representation in audiovisual translation, 2016, Sapienza Università di Roma & Università degli Studi di Roma Tre (LCRAV).
- 6<sup>th</sup> International Conference Media for All. Audiovisual Translation and Media Accessibility: Global Challenges, 2015. University of Western Sydney (M4A-6).

A relevant and long-running conference such as Languages and the Media, which takes place in Berlin every two years, could not be considered due to lack of available materials within the time-frame this analysis was carried out.

The abstracts were analysed manually in order to shed light on the main transfer modes, approaches, and topics presented at the conferences from a contrastive point of view. The fact that abstracts may not always accurately reflect the final presentation and the fact that the analysis was carried out by one individual scholar are some of the limitations of the present study, but it is my believe that this tentative mapping of the field can still provide interesting insights into the evolution of AVT research.

First of all, the audiovisual transfer modes were considered. Table 1 reflects the percentages of abstracts dealing only with just one transfer mode, dealing with more than one at the same time or simply not specifying the transfer mode.

AVT MODES	ISMW	BTM	M4A	TOTAL	M4A-6	LCRAV	INTER	TOTAL
Dubbing	20.99	40.91	17.14	26.3	7.27	36.21	11.76	18.41
Subtitling	37.04	31.82	17.14	28.67	45.45	31.03	23.53	33.34
Dubbing and subtitling	13.58	9.09	5.71	9.46	1.82	1.72	5.88	3.14
Voice-over	6.17	0	0	2.06	5.45	3.45	2.94	3.95
Audio description	3.70	0	14.29	6	9.09	0	17.65	8.91
SDH	2.47	0	14.29	5.58	9.09	1.72	5.88	5.57
SDH & AD	0.00	0	5.71	1.9	1.82	3.45	2.94	2.74
Sign language	0.00	0	0	0	5.45	0	0	1.82
Interpreting	1.23	0	0	0.41	1.82	0	0	0.61
Not specified	14.81	18.18	25.71	19.57	12.73	22.41	29.41	21.52

Table 1. Audiovisual transfer modes in conferences

Ten years ago, dubbing and subtitling gathered approximately the attention of the same percentage of papers (26.35% for dubbing, 28.67% for subtitling), totalling more than 50% of the presentations. A remarkable 9.46% dealt with both modalities, probably offering contrastive studies, and voice-over presence was anecdotal (2.06%). Audio description and subtitling for the deaf and hard-of-hearing concentrated around 13% of the papers, whilst around 20% did not mention or deal with a specific modality. When looking at each conference individually, one can observe that the presence of 40% of papers on dubbing in BTM impacts directly on the total percentage of dubbing, whilst the relative relevance of AD and SDH in M4A compensates for the low percentages of these modalities in the other two conferences. As far as sign language and media interpreting, they do not seem to find a good forum for dissemination in AVT conferences.

When looking at current data, a drop in dubbing research and in comparative studies including both dubbing and subtitling is observed, whilst interest in subtitling increases globally. This trend would be even more striking if data from the conference in Rome (LCRAV), with a remarkable 36% of papers on dubbing, was not considered. Voice-over remains a rather neglected mode, even in voice-over countries, despite a slight increase in numbers. And, as far as media accessibility modes are concerned, there is a slight increase overall, but lower than expected. This is probably due to the effect of the first M4A conference, where the focus was on media accessibility and a high number of papers were already presented on these modalities, whilst in its sixth edition (M4A-6) the approach was not so specific and subtitling for hearing audiences was given more attention. Similarly to ten years ago, sign language and

media interpreting present low numbers, and a significant number of abstracts do not specify any transfer mode.

Summing up, the previous data show that in 2004-2005 the focus of interest were traditional modes (subtitling, dubbing and voice-over totalled 66.54%) and media accessibility related modes amounted 13.48%. More recently, interest in media accessibility has increased up to almost 20%, and research presented at conferences dealing with traditional modes has decreased (less than 60%). The previous data also show some regional variation, with a stronger interest in dubbing in the Italian conferences.

The second step in the analysis was to highlight the approach taken, following a top-down approach in which four categories were pre-established: theoretical, descriptive, reception and technological papers. Despite the fact that most studies are built upon a theoretical framework that is generally made explicit in the abstracts and despite the fact that reception studies can be the result of a descriptive analysis, the aim was to choose the most relevant aspect in the abstract. Results are presented in Table 2.

Approach	ISMW	BTM	M4A	TOTAL	M4A-6	LCRAV	INTER	TOTAL
Theoretical	8.86%	13.64%	2.86%	8.45%	3.64%	3.45%	6.06%	4.38%
Descriptive	75.95%	50%	80%	68.85%	72.73%	84.48%	63.64%	73.62%
Reception	5.06%	27.27%	8.57%	13.64%	16.36%	8.62%	21.21%	15.40%
Technology and tools	10.13%	9.09%	8.57%	9.26%	7.27%	3.45%	9.09%	6.6%

Table 2. Approach to audiovisual translation research

The previous data show, first of all, a drop in theoretical approaches to AVT. A theoretical framework usually sustains research but developing theoretical proposals seems to attract less interest nowadays, at least in AVT conferences. One wonders whether other fora, like non-specific translation conferences, are seen as a more suitable venue for this type of contribution. Secondly, and contrary to my initial expectations, descriptive approaches to translation have not decreased but have actually increased, from 68.85% to 73.62%. This has been probably influenced by the conference in Rome, where an overwhelming 84% of the abstracts presented descriptive research. Similarly, reception research has increased but not as much as was expected, moving from 13.64% to 15.40%, and papers dealing only with technology or tools have surprisingly decreased. This is probably due to the characteristics of LCRAV but also to the fact that technology nowadays permeates all the other categories.

Finally, regarding topics, a qualitative analysis was preferred, reaching the following conclusions: all conferences present a wide array of topics, showing the diversification of interests in the field, but the focus on cultural aspects, on training, and on describing specific practices maintains in similar values across conferences. Linguistic approaches still capture the interest of many researchers, but not as much as in the past, and new topics such as fan translation, crowdsourcing and multilingualism experience a considerable rise.

When going back to the views expressed in the first lines of this paper, one could say that the tone was a bit optimistic, but they correlate very well with the structure of the last conference under analysis, *Intermedia 2016*, where around 6% of the papers put the focus on theoretical approaches, around 60% are descriptive, around 20% are reception studies, and almost 10% deal with technology and tools.

## 2. Past, present and future: the role of technology in AVT

In the previous data research on technology as applied to AVT had a lower presence than expected, but this was attributed to the fact that technology can be a tool for all types of research. Another possible explanation is that very often this type of research is presented in other conferences with a stronger focus on technology. In spite of this situation, technology is and is expected to be key in many aspects of AVT: in the process of creating translations, in the process of delivering them, in the process of receiving them, and also in the process of researching AVT. This is why in this section I will shift the focus of the article and I will concentrate on technology. I will deal with some technologies which have played a significant role recently and point to other technologies which may have an impact in the short term, briefly discussing research which links technologies and AVT but also stressing out the role of technology in both the translation and research processes. This section does not pretend to be an exhaustive and detailed

overview but a general approach that can help some readers become familiar with relevant research and, hopefully, find some inspiring ideas for the future.

## **2.1 Research on technologies applied to AVT**

Technologies have been distributed in three main categories: technologies related to the audio (including speech), technologies related to the text, and technologies related to the visuals.

### **2.1.1 Speech and audio-related technologies**

Various types of speech technologies have been researched in AVT such as speech recognition or text-to-speech, but also audio alignment or clean audio. Concerning speech recognition, respeaking has become the standard system for live intralingual subtitling and is defined by Romero-Fresco (2011:1) as a technique in which a respeaker listens to the original audio of a live programme or event and respeaks it, including punctuation marks and some specific features for the hearing impaired audience, to a speech recognition software, which turns the recognized utterances into subtitles on screen. Romero-Fresco (2011) presents a wide overview of this modality, which is an established practice in certain countries and is still under development in others, mainly due to the availability of proficient sound speech recognition system in some languages. Research on respeaking has focused on describing and analysing its features (Remael et al. 2014, Van Waes et al. 2013) but also testing its reception through various methodological tools, from questionnaires to eye-trackers (Romero-Fresco 2011). Works on creating metrics to analyse the quality of respoken subtitles have also been published (Dumouchel et al. 2011): from the NERD (Romero-Fresco 2011) to the NER model (Romero-Fresco and Martínez 2015). Nowadays, the skills required by the respeakers, as compared to interpreters, and respeaking for interlingual subtitles are two hot research topics (Szarkowska et al. 2015, 2016), and newer applications of respeaking such as the transcription of non-fictional content are being researched (Daniluk et al. 2015).

Speech recognition has also been used to automatically generate subtitles, a function included by popular tools such as YouTube and Google. Research-wise, the SAVAS project has been one of the most relevant in the field: focusing on the broadcast news domain, it has not only developed a remote respeaking system for collaborative subtitling (S.Respeak!), but it has also created a speaker independent transcription and subtitling application (S.Scribe!), which automatically transcribes pre-recorded audio and video files into time-aligned enriched subtitles, and an online subtitling system (S.Live!), which automatically transcribes live audio into configurable and well-formatted subtitles (Álvarez et al. 2015). The generation of the automatic subtitles is based on the application of Large Vocabulary Continuous Speech Recognition plus other technological components. In fact, speech recognition is often combined with other technologies and components to carry out relevant tasks, such as long audio alignment, speech and speaker recognition or language tracking (Álvarez et al. 2014).

In a different domain, the educational one, the project Translectures (Silvestre-Cerdà et al. 2012) implemented transcription (and translation) systems for online video lectures and, more recently, a new step has been taken in the EMMA project to transcribe (and translate) massive open online courses (Brouns et al. 2015). As explained by Valor-Miró et al. (2015), these transcription systems not only can be used by non-native speakers or deaf and hard-of-hearing viewers needing subtitles, but also for other uses beyond AVT such as lecture content searches, content summarisation or information retrieval.

Automatic transcription through speech recognition has also been applied together with a speaker diarization process to automatically extract AD transcripts (Delgado et al. 2015). Still, the acoustically noisy conditions of fiction films audios, the fact that they do not constitute a restricted domain, and the presence of spontaneous and overlapping speech are challenges to overcome.

When looking at the opposite, i.e. text converted into speech, investigations have focused mainly on the voicing of AD. Szarkowska (2011) led a project in Poland in which text-to-speech audio descriptions (TTS AD) were assessed in various contexts: in a monolingual feature film in Polish, in a dubbed educational TV series for children, in a foreign fiction film with voice-over, in a no-fiction film with audio subtitling and in a dubbed feature film. Results indicate that most participants accept TTS AD although it is not the preferred solution. Similar results were found in Catalan by Fernández-Torné and Matamala (2015). Extensive research in the field is reported by Kobayashi et al. (2009), who describe the application of TTS AD in online videos on websites.

TTS is also used successfully in audio subtitling, especially in countries such as Sweden, Denmark or the Netherlands (Verboom et al. 2002). Research in this field, though, is more limited, and mainly focuses on technical solutions (Hanzlicek et al. 2008, Ljunglöf et al. 2012) or, to a lesser extent, on user reception experiments (Thrane 2013). Newer applications may be found in voice-over, where Ortiz-Boix and Matamala (forthcoming) have researched the application of TTS in wildlife documentaries, and the BBC has started a pilot on multilingual voiceover (<http://www.bbc.co.uk/mediacentre/latestnews/2015/bbc-virtual-voiceover-translations>).

Finally, noise reduction and speech enhancement technologies such as clean audio (Fuchs & Oetting 2014, Orero & Permuy Hércules de Solás, forthcoming) are also being researched as tools to improve access to audiovisual content for hearing impaired audiences, and will probably yield interesting results in the near future. Possibilities are manifold, and the previously mentioned technologies and many others could undoubtedly have an impact in AVT research. As an example, speech emotion recognition (Koolagudi and Rao 2012) could be explored to automatically transfer emotions encoded in speech into written tags for those who cannot access the audio. Similarly, audio technology is omnipresent in dubbing processes but there is a considerable gap with research in the AVT field in this regard.

### **2.1.2. Text- related technologies**

Technologies can be used to alter texts in various ways: by transferring it from one language to another, by assisting in this transfer or by making formal or linguistic changes to it. Translation memories, for instance, are extensively used in technical, scientific, financial and legal translation to control terminological consistency and speed up the translation of repetitive segments. Still, AVT has often been viewed as a modality that was too creative and too dependent on the image for such a tool to be useful, and research in this area has been limited to a few papers. Pérez Rojas (2014), for instance, describes how to automatically create translation memories for subtitling using translated books adapted into films. And Hanouille et al. (2015a, 2015b) focus on documentary films and prove the usefulness of terminology-extraction software within translation memory tools as a support to audiovisual translators. Overall, though, as Díaz-Cintas (2015: 633) rightly points out, “little attention has been paid so far to the role that computer-aided translation (CAT) tools can play in subtitles or to the potential that translation memories and machine translation can yield in this field, although the situation is changing rapidly”.

Indeed, an area in which the situation has changed rapidly is machine translation (MT), although the interest has generally been on subtitling (Popowich et al. 2000, Piperidis et al. 2004, Melero et al. 2006, Armstrong et al. 2006, De Sousa et al. 2011). Exhaustive research has been carried out within the SUMAT project (Del Pozo 2014), in which MT was implemented in seven language pairs using a corpus trained with both professional and amateur subtitles. Large-scale evaluations measured the output quality and productivity gain (Etchegoyhen et al. 2014). Beyond the media environment, but still dealing with audiovisual content, the EMMA project is currently researching the implementation of MT in translating educational video content (Brouns et al. 2015). Some research, though, put the focus on the end user and analyse the impact of the automatised processes: in this regard, Matamala et al. (2015) prove that the usefulness of automatic subtitles correlates with the knowledge of the original language by end-users: whilst intralingual subtitles (involving speech recognition but not MT) prove useful for users with a B2 level of English, they do not affect positively users with higher proficiency; on the contrary, interlingual subtitles (involving both speech recognition and MT) may affect comprehension negatively in highly proficient participants due to a distracting effect.

Beyond the realm of the written word of subtitles, the ALST project (Matamala 2015) has researched at a smaller scale the inclusion of MT in the AD of fiction films, on the one hand, and in the voice-over of wildlife documentaries, on the other. The focus was been on the effort involved in post-editing, as it has been considered that MT without a revision is not fit for broadcasting. Ortiz-Boix and Matamala (2016) compare the effort involved in translating two short wildlife documentary excerpts and the effort involved in post-editing their MT. Results show that post-editing is faster and involves less technical and cognitive effort. Ortiz-Boix and Matamala (forthcoming) also carry out a quality analysis at various levels, and observe that there are not striking differences between the overall quality of post-edited and translated wildlife documentary films. Similarly, Fernández-Torné and Matamala (2016) compare the effort involved in creating an AD from scratch, in translating it manually, and in post-editing its MT version. Results show that the objective effort in post-editing is lower, but the subjective effort perceived by participants is higher, a finding that highlights the need to consider not only productivity measures but also the views of professionals.

Apart from MT and translation memories, other text-related technologies have been researched in audiovisual content, although not so extensively and often outside the audiovisual translation research community: text simplification (Daelemans et al. 2004, Shardlow 2014), text compression (Aziz et al. 2012, Luotolahti and Ginter 2015) or text segmentation (Scaiano et al. 2010, Álvarez et al. 2016) are three worth mentioning. The integration of speech and text technologies has also been investigated in speech-to-speech translation projects such as EU-Bridge (<http://project.eu-bridge.eu/>).

### **2.1.3. Image-related technologies**

Beyond the field of text and audio, image-related technologies open a wide array of possibilities in the audiovisual world, where audio and visual inputs interact. Whilst the previous technologies focused their research mostly in the subtitling domain, image-related technologies combined with natural language processing tools are opening new research grounds in the audio description field. In this regard, automatic description of images is being explored as a tool to improve accessibility to audiovisual content by blind and visually impaired audiences.

In the field of static images, social media are already implementing different services, both launched in 2016: Facebook's automatic alternative text which provides an automatic text description of a photo using object recognition technology (Voykinska et al. 2016) or Twitter's alternative text system, which adopts a manual approach. Concerning moving images, challenges are higher, but some research has already been developed in this area. Torabi et al. (2015) present a dataset of annotated videos based on audio descriptions, and Rohrbach et al. (2016) go a step further to propose a novel dataset with full length movies aligned with AD and movie scripts which are used to investigate a still challenging task: automatically describing scenes in natural language. The road ahead is a long one, but audiovisual translation scholars should keep an eye on this type of research, so that the real implementation of this technology in the future can be adequately assessed.

Other interesting fields related to AVT are automatic sign language translation and signing avatars (Wolfe et al. 2015) or automatic lip synch (Hoon et al. 2014), to name just two. Still, very often this type of image-related research is discussed in computational fields and does not reach the AVT research community. Bigger projects involving researchers from different research grounds are needed to aim for technologies which have an impact on the end-user.

## **2.2. Tools to create and receive content**

Apart from technology as the focus of research, technology is also the basis of tools to translate or adapt content and tools to consume content. Various researchers (Matamala 2005, Cerezo Merchán et al. 2016:131) have described the so-called audiovisual translation workstation, a workstation that changes continuously and is moving towards cloud-based solutions.

Regarding subtitling, for instance, there is a wide array of both professional and amateur software: free software (for instance, Aegisub, VisualSubSync, Subtitle Workshop, Gaupol, Jubler Subtitle Editor, Subtitle Creator, Subtitle Edit, Subtitle Processor, Gnome Subtitles, DivXLand Media Subtitler, Open Subtitle Editor) and professional software (WinCaps, Softwel Swift Create, Spot, EZTitles, FAB or SoftNI, among others). Their specifications vary and, whilst some allow to create subtitles by respeaking, automatically detect shot changes, provide an audio level indication waveform and propose automatic timing of subtitling, among other features, others only offer basic functionalities. Some subtitling software such as SRT translator even supports machine translation (Athnasiadi 2015). But, as said before, the workstation of the professional is moving to the cloud: platforms such as ZOO subs or iMediaTrans (Díaz-Cintas 2015: 637), for instance, offer professional subtitling. And outside the professional world, crowdsubtitling platforms such as Amara take a different approach and provide user-friendly cloud-based interfaces where volunteers are given the time-coded subtitles and focus on the translation.

Options are more limited for revoicing software addressed to translators, either for dubbing or voice-over (see, for instance, Dialogspotting 2 by Zioaudio), because many professionals still rely on word processing software, at least in countries like Spain (Cerezo Merchán et al. 2016:132). It is in dubbing studios where spotting tools or recording technology is used, but this does not generally reach the translator. In fandubbing contexts, the focus is also put on tools that allow to record and synchronize the new language track (see, for instance, Dubroo), and automatic services are also offered by websites such

as VideoDubber. Concerning audio description, on the contrary, the software offer is higher: from professionals tools such as Softel Swift ADePT or Fingertext to free tools such as MAGpie2, Livedescribe or YouDescribe, often for user-generated content and collaborative environments.

Research on user preferences or comparative analyses of tools do not abound. Effects of subtitling software on the process are analysed, for instance, by Aulavuori (2008). And Athnasiadi provides an overview of translation memories in AVT, showing how translation memories such as Transit NXT and cloud-based localisation platforms such as Transifex or XTM support subtitling files. Regarding AD, Vela Valido (2007) compares existing software in Spain and the USA. And Oncins et al. (2012) present an overview of existing subtitling software used in theatre and opera houses, and propose a universal solution to live media access which would include subtitling, AD and audio subtitling, among other features. Oncins et al.'s proposal links nicely the production and consumption processes. Indeed, at the other side of chain, tools for receiving content are expanding. Cinema and television screens have been complemented, and sometimes surpassed, by second screens in mobiles, tablets and even glasses. A myriad of apps providing access to subtitles, audio description and dubbed tracks are now available. A non-exhaustive list would include apps such as MovieReading, WhatsCine, ArtAccés or MyLingoApp, but what is really needed is research on the usability of these apps (see, for instance, Walczak 2016) or second screens (Graham et al. 2015) by audiences.

The presence of so many technical solutions opens the door to many research questions. It remains to be seen the actual usage of many tools and functionalities both in professional and amateur environments, both at the production and the reception ends. And it remains to be seen how new developments will be integrated into existing tools to adapt to the needs of audiovisual content.

### **2.3. Technology for audiovisual translation research**

To conclude this general overview, it should be stressed out that technology is not only a research object but it is also improving the way scholars are carrying out their investigations.

The rise in process and reception research in AVT studies has widened the choice of devices to include eye-trackers (Perego 2012, Kruger et al. 2015) and tools that monitor electrophysiological reactions such as electroencephalography (Szarkowska et al. 2015), heart rate (Ramos 2016) or electrodermal activity (<http://pagines.uab.cat/nea/>). Keylogging has also been used (Fernández-Torné and Matamala 2016, Ortiz-Boix and Matamala 2016), as well as screen recording software (Masse and Jud 2015). Still, the usage of these tools, except for the eye-tracker, is very recent.

The analysis of multimodal corpora has also implied the use of textual and multimodal corpus analysis tools such as ELAN (Matamala and Villegas 2016), both quantitative and qualitative. And needless to say that software for managing statistical information, and for creating and gathering data are commonplace these days. What is not so recurrent in the field of AVT is the use of online platforms to crowdsourcing microtasks, a practice found in other fields such as MT quality evaluation or transcription.

## **3. Conclusions**

This article has two different sections: on the first one, an overview of the field ten years ago and nowadays through the analysis of six conference abstracts has allowed to identify changes in research approaches, although they have not been as striking as initially expected. Indeed, the results of the analysis raise some interesting questions, such as the type of research that is presented in AVT conferences and the profile of the participants attending these conferences as well as the impact of regional interests on the research agenda. It would be interesting to analyse what sort of AVT-related investigations are presented in general translation conferences but also in more technological venues, and it would be useful to propose truly interdisciplinary events, a task which is not easy but which would undoubtedly be enriching for the field. The correlation of the results of my analysis with actual publications in journals would also shed some light on the dissemination strategies of AVT researchers.

Although technology was not as present as initially hypothesised in the conferences under analysis, it was observed that it permeates across various types of research proposals. Therefore, in the second part of the article, the focus has been put on technology, as it has been considered an element that is impacting and will continue to impact AVT in terms of research and practice. Technologically-based research linked to audiovisual transfer modes has been summarised, and some suggestions about future implementations

have been put forward. Whilst some technologies are mature enough and have found their way into the professional arena, others are just beginning to be researched and are expected to yield interesting results in the near future. Still, this will only be possible if AVT scholars are involved in this type of research in truly interdisciplinary teams that not only cater for technological developments but also take into account the impact on processes, professionals, products and end-users.

### Acknowledgements

I would like to thank the conference organisers for providing the book of abstracts. Anna Matamala is a member of TransMedia Catalonia research group, funded by the Catalan government under funding scheme 2014SGR0014. This paper is part of the research project NEA (FFI2015-64038-P, MINECO/FEDER, UE).

### References

- Álvarez, A., Arzelus, H., & Etchegoyhen, T. (2014). Towards customized automatic segmentation of subtitles. In *Advances in Speech and Language Technologies for Iberian Languages* (pp- 229-239). Berlin-Heidelberg: Springer.
- Álvarez, A., Arzelus, H., & Ruiz, P. (2014). Long audio alignment for automatic subtitling using different phone-relatedness measures. *IEEE. International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 6280-6284). IEEE.
- Álvarez, A., Balenciaga, M., Del Pozo, A., Arzelus, H., Matamala, A., & Martínez-Hinarejos, C. (2016). Impact of automatic segmentation on the quality, productivity and self-reported post-editing effort of intralingual subtitles. In *Proceedings of the Tenth Int. Conference on Language Resources and Evaluation (LREC 2016)*. Paris: ELRA.
- Armstrong, A., Way, A., Caffrey, C., Flanagan, M., Kenny, D., & O'Hagan, M. (2006). Improving the quality of automated DVD subtitles via example-based machine translation. In *Proceedings of Translating and the Computer 28*. London: ASLIB.
- Athnasiadi, R. (2015). *The applications of machine translation and translation memory tools in audiovisual translation: a new era?* MA thesis. University College London.
- Aulavuori, K. (2008). *Effects of the digital subtitling software on the subtitling process: a survey among Finnish professional television subtitlers*. MA thesis. University of Helsinki.
- Aziz, W., de Sousa, Sheila C.M., & Specia, L. (2012). Cross-lingual sentence compression for subtitles. In *Proceedings of the 16<sup>th</sup> EAMT Conference* (pp 103-110).
- Brouns, F., Serrano Martínez-Santos, N., Civera, J., Kalz, M., & Juan, A. (2015). Supporting language diversity of European MOOCs with the EMMA platform. In *Proceedings of the European MOOC Stakeholder Summit 2015* (pp. 157-165).
- Cerezo Merchán, B., Chaume, F., Granell, X., Martí Ferriol, J.L., Martínez Sierra, J.J, Marzà, A., & Torralba Miralles, G. (2016) *La traducción para el doblaje en España. Mapa de convenciones*. Castelló: Universitat Jaume I.
- Daelemans, W., Höthker, A., & Tjong Kim Sang, E. (2004). Automatic sentence simplification for subtitling in Dutch and English. In *Proceedings of LREC 2014* (pp 1045-1048). ELRA: Paris.
- Daniluk, L., Matamala, A., & Romero-Fresco, P. (2015) Transcribing documentaries. Can respeaking be used efficiently? Paper presented at 5th International Symposium Respeaking, Live Subtitling and Accessibility. Rome.
- Delgado, H., Matamala, A., & Serrano, J. (2015) Speaker diarization and speech recognition in the semi-automatization of AD: an exploratory study on future possibilities. *Cadernos de tradução* 35(2), 308-324.
- De Sousa, S., Aziz, W., & Special, L. (2011). Assessing the post-editing effort for automatic and semi-automatic translations of DVD subtitles. In *Proceedings of Recent Advances in Natural Language Processing* (pp. 97-103).
- Del Pozo, A. (2014). *SUMAT final report*. Donostia, Vicomtech-IK4.
- Díaz-Cintas, J. (2015) Technological strides in subtitling. In S-W. Chan (ed), *Routledge Encyclopedia of translation Technology* (pp. 632-643). Ondon: Routledge.



- Dumouchel, P., Boulianne, G., & Brousseau, J. (2011). Measures for quality of closed captioning. In A. Serban, A. Matamala & J.M. Lavour (eds), *Audiovisual translation in close-up* (pp. 151-172). Bern: Peter Lang.
- Etchegoyhen, T., Bywood, L., Fishel, M., Georgakopoulou, P., Jiang, J., & van Loenhout, G. et al. (2014). Machine translation for subtitling: a large-scale evaluation. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation LREC 2014* (pp. 46-53). Paris: ELRA.
- Fernández-Torné, A., & Matamala, A. (2015). Text-to-speech vs. human voiced audio description: a reception study in films dubbed into Catalan. *Journal of specialised translation* 24, 61-88.
- Fernández-Torné, A., & Matamala, A. (2016). Machine translation in audio description? Comparing creation, translation, and post-editing efforts. *Skase* 9(1), 64-85.
- Fuchs, H., & Oetting, D. (2014). Advanced clean audio solutions: dialogue enhancement. *Motion Imaging Journal* 123(5): 23-27.
- Graham, R., Romero-Fresco, P., Lambourne, A., Holmes, R., Middleton, R. (2015). *Stagetext: CaptionCue. Research and Development Report*. Available online at: [http://artsdigitalrnd.org.uk/wp-content/uploads/2014/06/Stagetext\\_Project\\_Report.pdf](http://artsdigitalrnd.org.uk/wp-content/uploads/2014/06/Stagetext_Project_Report.pdf) (last access: 19 July 2016).
- Hanouille, A., Hoste, V., & Remael, A. (2015a). The efficacy of terminology-extraction systems for the translation of documentaries. *Perspectives. Studies in Translatology* 23(3), 359-374.
- Hanouille, A., Hoste, V., & Remael, A. (2015b). The translation of documentaries: can domain-specific, bilingual glossaries reduce the translators' workload? *New Voices in Translation Studies* 23, 25-49.
- Hanzlicek, Z., Matousek, J., & Tihelka, D. (2008). Towards automatic track generation for Czech TV broadcasting: initial experiments with subtitles-to-speech synthesis. In *9th International Conference on Signal Processing* (pp. 2721-2724). IEEE.
- Hoon, L., Chai, W., & Rahman, K. (2014). Development of real-time lip sync animation framework based on viseme human speech. *Archives of Design Research* 27(4), 19-29
- Kobayashi, M., Fukuda, K., Takagi, H., & Asakawa, C. (2009). Providing synthesized audio description for online videos. In *ASSETS '09: Proceedings of the 11<sup>th</sup> International ACM SIGACCESS Conference on Computers and Accessibility* (pp. 249-250). New York: ACM.
- Koolagudi, S.G., & Rao, X. (2012). Emotion recognition from speech: a review. *International Journal of Speech Technology* 15(2), 99-117.
- Kruger, J.L, Szarkowska, S., & Krejtz, I. (2015). Subtitles on the moving image: an overview of eye tracking studies. *Refractory. A Journal of Entertainment Media*. Available online at: <http://refractory.unimelb.edu.au/2015/02/07/kruger-szarkowska-krejtz/> (last access: 19 July 2016).
- Ljunglöf, P., Derbring, S., Olson, M. (2012). A free and open-source tool that reads movie subtitles aloud. In *NAACL-HLT Workshop on Speech and Language Processing for Assistive Technology* (pp. 1-4). Montreal: ACL.
- Luotolahti, J., Ginter, F. (2015). Sentence compression for automatic subtitling. In *Proceedings of the 20<sup>th</sup> Nordic Conference of Computational Linguistics (NODALIDA 2015)* (pp. 135-143). Vilnius: Linköping University.
- Matamala, A. (2005). La estación de trabajo del traductor de productos audiovisuales: herramientas y recursos. *Cadernos de tradução* 16, 251-268.
- Matamala, A. (2015). The ALST project: technologies for audiovisual translation. In *Proceedings of the 37th Conference Translating and the Computer* (pp. 79-89). Geneva: Tradulex.
- Matamala, A., & Villegas, M. (2016). Building an audio description multilingual multimodal corpus: the VIW project. In J. Edlung, D. Heylen & P. Paggio (eds) *Multimodal Corpora: Computer vision and language processing (MCC 2016)* (pp 29-32). Portoroz, Slovenia.
- Matamala, A., Oliver, A., Álvarez, A., & Azpeitia, A. (2015). The reception of intralingual and interlingual automatic subtitling: an exploratory study within the HBB4ALL project. In *Proceedings of the 37th Conference Translating and the Computer* (pp. 12-17). Geneva: Tradulex.
- Massey, G., & Jud, P. (2015). Teaching audiovisual translation with products and processes: subtitling as a case in point. In L. Bogucki & M. Deckert (eds) *Assessing audiovisual translation* (pp. 99-116). Frankfurt: Peter Lang.

- Melero, M., Oliver, A., & Badia, T. (2006). Automatic multilingual subtitling in the eTITLE project. In *Proceedings of Translating and the Computer 28*. London: ASLIB.
- Oncins, E., Lopes, O., Orero, P., Serrano, J., & Carrabina, J. (2013). All together now: a multi-language and multi-system mobile application to make live performing arts accessible. *Journal of Specialised Translation* 20, 147-164.
- Orero, P., Permy Hércules de Solás, I.C. (forth.). Clean audio: from an emerging access service to a mature reality.
- Ortiz-Boix, C., & Matamala, A. (2016). Post-editing wildlife documentary films: a new possible scenario? *Journal of Specialised Translation* 26, 187-210.
- Ortiz-Boix, C., & Matamala, A. (forth) Text-to-speech Voice-over? A study on user preferences in the voicing of wildlife documentaries.
- Ortiz-Boix, C., & Matamala, A. (forth.) Assessing the quality of post-edited wildlife documentaries. *Perspectives. Studies in Translatology*.
- Perego, E. (ed) (2012). *Eye-tracking in audiovisual translation*. Trieste: Aracne.
- Pérez Rojas, K. (2014). Automatically building translation memories for subtitling. In E. Torres-Simón & D. Orrego-Carmona (eds), *Translation research projects 5* (pp. 51-62). Tarragona: Intercultural Studies group.
- Piperidis, S., Demiros, I., Prokopidis, P., Vanroose, P., Hoethker, A., Daelemans, W., Sklavounou, E., Konstantinou, M., & Karavidis, Y. (2004). Multimodal multilingual resources in the subtitling process. In *Proceedings of the Fourth International Conference on Language Resources and Evaluation LREC-2004* (pp. 205-208). Paris: ELRA.
- Popowich, F., McFetridge, P., Turcato, D., & Toole, J. (2000). Machine translation of closed captions. *Machine Translation* 15(4), 311-341.
- Ramos, M. (2016). Testing audio narration: the emotional impact of language in audio description. *Perspectives. Studies on Translatology*. Published online.
- Remael, A., Van Waes, L., & Leijten, M. (2014). Live subtitling with speech Recognition: how to pinpoint the challenges. In D. Abend-David (ed), *Media and translation: an interdisciplinary approach* (pp. 131-148). New York: Bloomsbury Academic.
- Rohrbach, A., Torabi, A., Rohrbach, M., Tandon, N., Pal, Ch., Larochele, H., Courville, A., Schiele, B. (2016) Movie Description. Available online at: <https://arxiv.org/abs/1605.03705> (last access 19 July 2016).
- Romero-Fresco, P. (2011). *Subtitling through speech recognition: Respeaking*. Manchester: St. Jerome.
- Romero-Fresco, P., & Martínez, J. (2015). Accuracy rate in live subtitling- the NER Model. In R. Baños & J. Díaz-Cintas (eds) *Audiovisual translation in a global context* (pp. 28-50). London: Palgrave Macmillan.
- Scaiano, M., Inkpen, D., Laganière, R., Reinhartz, A. (2010). Automatic text segmentation for movie subtitles. *Proceedings of the 23<sup>rd</sup> Canadian conference on Advances in Artificial Intelligence* (pp. 295-298). Berlin-Heidelberg: Springer.
- Shardlow, M. (2014). A survey of automated text simplification. *International Journal of Advanced Computer Science and Applications, Special Issue on Natural Language Processing* 2014, 58-70.
- Silvestre-Cerdà, J.A., Albert, J., Del Agua, M., Garcés, G., Gascó, G., Giménez-Pastor, A., Martínez, A., Pérez González de Martos, A., Sánchez, I., Serrano Martínez-Santos, N., Spencer, R., Valor-Miró, J.D., Andrés-Ferrer, J., Civera, J., Sanchís, A., & Juan, A. (2012) transLectures. In *Proceedings of IberSPEECH 2012* (pp. 345-351). Madrid.
- Szarkowska, A. (2011). Text-to-speech audio description: towards wider availability of AD. *Journal of Specialised Translation* 15, 142-162.
- Szarkowska, A., Dutka, L., Chmiel, A., Brocki, L. (2015). Are interpreters better respeakers? An exploratory study on respeaking competences. 5th International Symposium Respeaking, Live Subtitling and Accessibility. Rome.
- Szarkowska, A., Krejtz, K., Dutka, L., & Pilipczuk, O. (2016). Cognitive load in intralingual and interlingual respeaking – a preliminary study. *Poznań Studies in Contemporary Linguistics* 52(2), 209-233.
- Thrane, L.K. (2013). *Text-to-speech on Digital TV. An Exploratory Study of Spoken Subtitles on DRISyn*. MA Thesis, Univ. Of Copenhagen.

- Torabi, A., Pal, Ch., Larochelle, H., & Courville, A. (2015). Using descriptive video services to create a large data source for video annotation research. Available online at: <https://arxiv.org/abs/1503.01070> (last access 19 July 2016).
- Valor Miró, J.D., Silvestre-Cerdà, J.A., Civera, J., Turró, C., & Juan, A. (2015). Efficiency and usability study of innovative computer-aided transcription strategies for video lecture repositories. *Speech Communication* 74, 65-75.
- Van Waes, L., Leijten, M., & Remael, A. (2013). Live subtitling with speech causes and consequences of text reduction. *Across Languages and Cultures* 14 (1), 15-46.
- Vela Valido, J. (2007). *Software de audiodescripció: anàlisi comparatiu de los programas existentes en España y EE.UU.* MA thesis. Universitat Autònoma de Barcelona.
- Verboom, M., Crombie, D., Dijk, E., & Theunisz, M. (2002). Spoken subtitles: making subtitled TV programmes accessible. In K. Miesenberger, J. Klaus & W. L. Zagler (eds), *Proceedings of Computers Helping People with Special Needs, 8th International Conference, ICCHP 2002* (pp. 295-302). Berlin-Heidelberg: Springer.
- Voykinska, V., Azenkot, S., Wu, S., & Leshed, H. (2016) How blind people interact with visual content on social networking services. Computer-supported cooperative work and social computing (CSCW) 2016. Available online at [https://sites.coecis.cornell.edu/leshed/files/2016/03/1584\\_voykinska-2iq8538.pdf](https://sites.coecis.cornell.edu/leshed/files/2016/03/1584_voykinska-2iq8538.pdf) (last access 19 July 2016).
- Walczak, A. (2016) Cinema4All: Audio Description on Smartphones. Intermedia Conference. Łódź, Poland.
- Wolfe, R., Efthimiou, E., Glauert, J., Hanke, T., McDonald, J., Schnepp, J. (eds) (2015). *Universal Access in the Information Society. Special issue: recent advances in sign language translation and avatar technology*. Berlin-Heidelberg: Springer.