

Chapter 18. Technological Advances in Audiovisual Translation

Jorge Díaz Cintas
University College London

Serenella Massidda
University of Roehampton, London

[T]echnology is no longer something that stands over and against human beings; it structures human life and social relations from within. Technical forms of organization, technical networks and equipment, the entire material aspect of technology, have penetrated into the depths of who we are, into our bodies and our communicative relations, in the process bringing about deep changes to the way our societies cohere.

(Böhme, 2008/2012: 7)

Abstract

This chapter maps out the field of audiovisual translation (AVT) in conjunction with technology by investigating emerging trends and discussing some critical aspects of the increasing pervasiveness of digital accelerationism and the globalized (r)evolution that has affected the entertainment industry in the last decades. By adopting a diachronic perspective, this chapter opens with an historical trajectory that spans from the invention of cinema to the rise of the Web 2.0 in the new millennium. In order to take stock of the impact that technological advances have had on AVT practices, the most prominent areas of the field, i.e. subtitling and revoicing, are analysed to unveil the specific technologies, architectures, and software programs developed to enhance and optimize translation tasks as well as global localization workflows. In the last section, a set of conclusions highlights the implications of technological innovation in the professional practice of audiovisual translators.

Keywords

Audiovisual translation, subtitling, dubbing, revoicing, cloud-based platforms, workflows

Introduction

The increasing pervasiveness and democratization of technology together with the intensification of globalization processes have made possible the fluid and immediate circulation of information and cultural messages. The entertainment and multimedia industry, through the various audiovisual translation (AVT) practices available, plays a crucial role in this dynamic mediascape. Not only has technology revolutionized the sphere of mass media, but it has also aggrandized its potential with the introduction of new distribution channels aiming to reach wider and ever more global audiences. The easy availability of technology, the seduction of multimodal productions and the affordances of AVT have all acted as triggers in the rapid spread of audiovisual materials through television, computers, tablets, smart phones and silver screens.

Technology has come to be an omnipresent reality that infiltrates not only the social life of the individual but also the way in which the external environment is being moulded. For authors like Schwab (2016: 7), '[o]f the many diverse and fascinating challenges we face today, the most intense and important is how to understand and shape the new technology revolution, which entails nothing less than a transformation of humankind'. To be successful in this ecosystem, translators need to adapt and adjust to the new changes so that they can harness state-of-the-art technologies to their advantage rather than risking being replaced by them.

In today's society, driven by technological accelerationism (Beckett 2017), we have become dependent on the affordances offered by digitization to share information and communicate with our peers and, consequently, exchanges have become more multimedia

and audiovisual than ever before. Screens are now an omnipresent feature of social interaction, having colonized the private and public spheres and having become the main tool to inform, mislead, entertain, keep in touch, educate and sell goods and services. The creation and exchange of multimodal productions that combine the audio and the visual planes, and in which written texts, (moving) images and speech are semiotically integrated, is an everyday occurrence at virtually everyone's fingertips.

The prevalence of the feature film as the audiovisual genre par excellence has waned in recent decades as other formats gain in popularity and recognition. In addition to more traditional formats like documentaries, news, TV series and children's programmes, the audiovisual mediascape is now richly populated with docudramas, political broadcasts, chat shows, user-generated videos, stand-up comedy, slow programmes, corporate videos, edutainment productions, video games and commercials to name but a few. The media industry is not only thriving in terms of content variety and high volume of productions, but also in its financial significance as well as its technological metamorphosis. Gone are the days when the Betamax and the VHS tapes occupied the place of honour in the distribution of audiovisual materials. The switching off of analogue technology and the establishment of digitization during the last decade of the 20th century saw the emergence of new media storage and distribution formats, like the CD-ROM, the DVD, the VCD and the Blu-ray, as well as the diversification of the distribution channels beyond the cinema and the television, with the consolidation of the internet and the upsurge of social media networks and video on demand (VOD) platforms.

The boom in popularity of over-the-top (OTT) services like Netflix, Amazon Prime, and Hulu, along with the explosion of audiovisual serialized productions have led to a growing demand for content localization across the world. According to a study conducted in 2017 on behalf of the MESA Europe Content Localization Council (Green 2018), the volume of work will reach \$2.5 billion annual spend before 2020 in Europe, Middle East and Africa (EMEA).

This audiovisualization of communication has in turn propitiated an exponential growth in the volume of AVT that is commissioned and consumed, making this specialization one of the most dynamic and vibrant branches within translation studies. AVT is here understood as a hypernym that subsumes a series of practices usually activated to deal with the translation of audiovisual material for audiences who do not understand the language of the original production, be it by means of revoicing (dubbing, voiceover, narration, interpreting) or timed text (subtitling, surtitling). It also encompasses forms of linguistic mediation aimed at making audiovisual programmes accessible to audiences with sensory impairments, such as subtitling for the deaf and the hard-of-hearing (SDH) and audio description (AD) for the blind and the partially sighted (see Remael and Reviers in this volume).

Language service providers (LSPs) and translation vendors have been exploring for years the possibilities offered by new technology and by practices such as crowdsourcing and microtasking, in their attempt to keep up with higher demand and faster turnarounds, while maintaining their budgets and competitors on check. AVT's ample financial and social potential has also been acknowledged by technology developers and manufacturers, who have invested time and effort in enhancing the functionality and efficiency of the tools used in this domain, and academia has of late veered from the study of linguistic issues to the scrutiny of the role played by technology (Díaz Cintas 2015, Matamala 2017).

Tools for Subtitlers in the Digital Era

After the invention of cinema at the end of the 19th century and the advent of television in the 1950s, the arrival of the internet in the 1990s can be hailed as one of the most significant turning points in the history of human communication. Indeed, the web and the social networks have had a transformative impact on the way in which we interact and communicate with each other, mainly due to the attractiveness of the audiovisual format in which audio and visuals coalesce, enhancing the semiotic possibilities of the composite message and its potential to improve comprehension and retention.

Although AVT practices have been in existence since the appearance of cinema, technologically speaking their evolution was relatively slow for many decades. In the field of AVT, the transition from analogue to digital technology was symbolically marked by 'the death of VHS and the upsurge of the DVD at the end of the last millennium' (Díaz Cintas 2015: 632). From a technical perspective, this shift has led to an ever-increasing commercial interest surrounding the design and manufacturing of powerful software programs. The first editors designed exclusively for the creation of subtitles started being commercialized in the mid-1970s. At the time, subtitlers needed a desktop computer, an external video player in which to play the VHS tapes with the material to be translated, and a television monitor to watch the audiovisual productions. The computer would have a word processor as well as a basic subtitling program, which made it possible to simulate the subtitles against the images on screen. Since then, the situation has changed significantly and these days with a PC or laptop, a digital copy of the video to be translated, and a subtitling program, subtitlers can perform all pertinent tasks in front of a single screen: watch the original video, type their translation, decide the in and out times of each subtitle, take due care of shot changes, monitor the display rate and length of their subtitles, decide on the positioning and colour of the text, and spell check their translation.

The capability and functionality of most professional subtitling programs have been improved at an incredibly fast pace in recent decades, with some of the leading manufacturers being EZTitles (www.eztitles.com), FAB (www.fab-online.com), Screen Systems (www.screensystems.tv), Spot (www.spotsoftware.nl), and TitleVision (<http://titlevision.dk>). Alternatively, some subtitling vendors have developed their own proprietary software programs like Deluxe's EddiePlus (<http://bydeluxemedia.com>) and SDI Media's Global Titling System (www.sdimedia.com). The functionalities provided by commercial software programs are plentiful and varied. During the subtitling production phase, users can rely on automatic backup, type horizontal and vertical subtitles, originate subtitles with the help of the audio wave form bar, avoid crossing shot changes thanks to the auto-detected shot changes function, and enhance the accuracy of subtitle synchronization by employing audio scrubbing, i.e. by enabling the user to drag a cursor across a segment of a waveform to hear it. After the translation, subtitlers can lock the timecodes and/or subtitles to prevent external editing, merge subtitle files, resynch the timing to suit different video formats, export and convert the subtitle file into virtually any format available on the market, and compare source, translated and revised files for quality control (QC). The subtitle file can go through automatic or manual quality checks to ensure that the layout, spelling and timing of the output are correct. Commercial subtitling editors check the maximum number of characters per line, raised subtitles, spaces and rows, safe area, font size and position; while, through the spell check, the software performs an analysis of (customizable) linguistic errors and typos, and ensures that a file complies with style guidelines and/or client rules. The timing check analyses minimum and maximum duration of subtitles, timing overlaps, shot changes, display rates, gaps between chained subtitles and many other parameters. All in all, this functionality helps subtitlers boost their productivity and reach high levels of consistency. From a commercial perspective, and to compete with the rise of freeware, a transition has taken place in the industry from capital expenditure (CAPEX), i.e. investing in the purchase of tools, to operating expenditure (OPEX), whereby users rent a piece of software on a pay as you go basis.

Since its inception, the Web 2.0 has unleashed an unprecedented series of epochal changes (Perrino 2009). A slippery concept, one of the earliest definitions is the one provided by O'Reilly (2006: online), in which he highlights that rather than a mere digital revolution, the Web 2.0 was a business revolution in the computer industry powered by 'the move to the internet as platform'.

By harnessing collective intelligence, O'Reilly (2006: online) suggests that Web 2.0 applications should be built with the users in mind so that they can 'get better the more people use them'. Such an approach, in turn, has led to the rise of the 'prosumer' (Toffler 1980) or 'produser' (Bruns 2008) as one of the new, main actors in the cyberworld in which user experience and user agency are the driving force. Another defining trait of the Web 2.0 refers to the open source (OS) nature of the software. Every program has a source code, i.e. a set

of instructions in the form of text to be performed by a machine, and while proprietary software comes in as a ready-to-run version, where the interface communicates with the source code, but the user has no access to it, OS software includes the source code that advanced users can freely access and legally modify. If proprietary software has a bug, the developer will need to fix it by releasing an upgraded version, which can be a financially onerous and time-consuming process. If OS software has a bug, the pool of developers in the community will fix it for free. The latter is the approach championed by companies like Netflix, with the creation of their Open Source Software Center (<https://netflix.github.io>). The fact that professional subtitling software has traditionally been rather expensive and out of reach for many translators has encouraged some to favour the development of free, OS subtitling editors such as Aegisub (www.aegisub.org), DivXLand Media Subtiter (www.divxland.org/en/media-subtiter), Subtitle Edit (www.nikse.dk/SubtitleEdit), and Subtitling Workshop (<http://subworkshop.sourceforge.net/index.php>).

This democratization of technology can be said to have initially materialized in the availability of easy-to-use subtitling freeware and video editing suites which, together with the potential opened by faster internet connections and peer-to-peer computing, gave rise in the early 1990s to the amateur practice known as fansubbing (Díaz Cintas and Muñoz Sánchez 2006), and which in recent years has metamorphosed into a multitude of cybersubtitling variants that breed in hyperspace (Díaz Cintas 2018). In the olden days, fansubbers' drive was confined to the popularization of Japanese anime but their remit and scope have widened to cover a vast range of audiovisual programmes, including the latest US films and TV series (Massidda 2015), and raising some serious ethical considerations. Indeed, fansubs are technically illegal as they are not officially licensed and, therefore, infringe the copyright of the owners of the audiovisual programme (Kelion 2013).

Taking advantage of the potential offered by OS technology, for over two decades prosumers have been developing near-professional subtitling freeware capable of creating, editing, and converting text-based subtitle files in myriad of formats. OS subtitling software tends to be user-friendly, multi-language, with customizable interfaces and shortcuts, and provides advanced functions that are pretty similar to the ones embedded in professional programs, such as automatic calculation of ideal duration of each subtitle, spell check in different languages, frame per second conversion, and automatic detection of subtitle errors, both temporal and linguistic. Similarly to commercial programs, most subtitling free editors also support style and colour tags and inform subtitlers about the number of characters per second being used in any given subtitle. As for the spotting phase, it can be performed with the added benefit of the soundwave bar that allows to modify the timing by dragging the subtitle blocks created and following the audio waveform representation. Some programs, such as Subtitle Edit version 3.5.7, also integrate an auto translate tool powered by Google.

The technological possibilities unleashed by the new forms of immersive content, like 360° video, virtual reality and augmented reality, are having an impact on the way in which subtitles are being conceptualized and displayed so that they can add value to the immersive experience rather than detract. While in conventional cinema and television settings, people look at the screen in front of them and the subtitles are mostly displayed within the image, usually at the bottom of the screen, immersive environments allow viewers to look in any direction, thus raising the question of where best to position the subtitles. Experiments are already being conducted with members of the audience (Brown and Patterson 2017) in an area of vast potential and bound to expand in the near future.

Tools for Revoicing

In contrast to the far-reaching technical developments witnessed in the field of subtitling, revoicing practices like dubbing and voiceover seem to have been bypassed by the latest technology (Baños 2018). In the industry, professional translators are not usually required to work with dedicated software and they can fulfil their tasks by using a simple text editor and a video player. It is later, in the recording studio, where dubbing directors and sound engineers make use of specialist applications. As an exception to this rule, and in countries like France,

translators who also take care of lip-synching are frequently asked to insert their translation in a track known as *bande rythmo*, which is added onto the working copy of the audiovisual production, at the bottom of the screen, and includes the target dialogue as well as the necessary symbols for the dubbing actors (Chaume 2012). Freeware programs like Capella (www.cappella.tv), and professional ones like VoiceQ (<https://voiceq.com>), have been designed to allow translators and those in charge of the synchronization of the target dialogue to create the *bande rythmo* for the benefit of the voice talents.

Audio description for the blind and the partially sighted is an access service that consists of providing an additional aural narration describing any visual or audio information that will help visually-impaired individuals to follow the plot of the story, such as the body language and facial expressions of the characters, the surrounding landscape, the source of certain sounds, the actions taking place on screen and the characters' sartorial style. From the very beginning, specialized software packages have been available to facilitate the production and recording of AD scripts, which tend to be very similar to subtitling stations. These days, one of the leading companies specialising in this sector is Starfish Technologies (www.starfish.tv/audio-description-video-description), and initiatives like YouDescribe (<http://youdescribe.org>), which make it very easy to create AD tracks for YouTube videos, have helped popularize this practice.

As stakeholders' interest on making the most of speech technologies is growing, new applications are being developed and having a noticeable impact on the profession. Nowadays that video material is produced in vast quantities and it rarely comes accompanied with dialogue lists, the need to transcribe the speech heard in the soundtrack has become more pressing than ever before. In this respect, automatic speech recognition (ASR) technology has known a vertiginous progress in transforming original speech into written text, which can then be used to create intralingual subtitles or templates for translation into multiple languages (see Ciobanu and Secară in this volume).

The other professional practice that makes full use of ASR to produce live subtitles is respeaking. It is a process whereby a person listens to the original utterances of a live programme and dictates the speech, i.e. respeaks it, including punctuation marks and some specific features for the deaf and the hard-of-hearing audience, to a microphone connected to a computer, to an ASR app and to a subtitling program, which then displays subtitles on screen (Romero-Fresco 2011). The fact that ASR apps can only work with certain major languages, simply because more research and capital investment has been put into them, makes their application impractical in the case of some minoritized languages; at least for the time being.

Though live subtitling was mainly used in the production of intralingual SDH in the past, it is nowadays making inroads into the field of interlingual subtitling. This is due to the multilingual nature of some live TV shows, in which a foreign guest may have to be subtitled interlingually, or to the need to subtitle live high-profile foreign language broadcasts such as the inaugural speeches of presidents or Nobel Prize winners. Interlingual live subtitling is also becoming increasingly popular at conferences and other public events, where it can replace simultaneous interpreting. In interlingual live subtitling another challenge is added to the already complex respeaking practice: the audio input must be translated/interpreted live and communicated to the ASR software, while the respeakers listen to the new audio input that is to follow and monitor their own speech as they are reproducing the previous input. As a result, the delay or latency, which can be disturbing in intralingual live subtitling, often further increases in interlingual live. To train professionals in this job, the European project Interlingual Live Subtitling for Access (ILSA, www.ilsaproject.eu) is working in the development of the competences needed to successfully operate in this field and has also designed a specialized training course for live interlingual subtitling.

Another effective application of ASR is represented by automatic subtitling based on machine-learning algorithms, a functionality retrieved in web-as-participation-platforms such as YouTube. Similarly, the project SAVAS – Live Subtitling and Captioning Made Easy (https://cordis.europa.eu/project/rcn/103572_en.html), run from 2012 until 2014, had as one of its key aims the development of speaker independent ASR technology for the production of multilingual subtitling in the broadcasting domain. As documented by Aliprandi *et al.* (2014),

three subtitling systems were developed: S.Scribe! for the automatic transcription of audio and video files into time-aligned subtitles, as well as an online version called S.Live! Their third tool, S.Respeak!, is a respeaking system that can be used remotely for collaborative live subtitling.

At the other end of the spectrum, text-to-speech technologies (TTS) have been the focus of research studies devoted to AD, such as the one carried out by Szarkowska (2011). A set of TTS audio descriptions for feature films were exploited to test viewers' preferences when it comes to human versus synthetic Polish voices, with the result that while the majority of participants preferred natural speech, many of them found synthetic speech acceptable, a conclusion corroborated a few years later by the experiment conducted by Fernández-Torné and Matamala (2015) in the case of Catalan. TTS technologies applied to audio subtitling are also widely employed in Denmark, Sweden and the Netherlands to make subtitle-translated programmes accessible to the visually-impaired by creating a service which automatically produces talking subtitles using synthetic speech (Verboom *et al.* 2002).

One of the research projects that has looked into the potential of speech technologies (recognition and synthesis) and translation technologies (machine translation) is Linguistic and Sensorial Accessibility: Technologies for Voiceover and Audio Description (ALST), conducted between 2013 and 2015. Using OS resources, its main aim was to investigate whether certain technological solutions can have a positive impact when providing access to audiovisual content to viewers who do not understand the original language as well as to those who cannot hear the audio or see the visual content (Matamala 2016). Part of the testing assessed whether speech recognition could be used to automatically transcribe an AD narration in a source language for which no written script was available. This material was subsequently machine translated into a different language and then post-edited by a human agent to enhance the quality of the raw material. The results, though encouraging, prove to be inconclusive and the authors call for further experiments with MT engines trained with specific corpora. The scholars' second objective centred on appraising how visually-impaired users experienced text-to-speech synthetic voices as compared to natural ones when listening to an AD script and corroborates previous research that show viewers' acceptance of artificial voices, though natural voices are their preferred ones. From a commercial perspective, VOD giant Netflix have recently experimented with synthesized AD for two documentary series, *Shot In the Dark* and *Fastest Car*, featuring a male and female computer voice respectively (Scott 2018), thus signalling a shift from traditional practices to potential future directions.

Machine Translation and Translation Memory as Tools

For almost two decades, computational linguists, recognising the potential of new technologies to escalate productivity, have explored the interaction between machine translation (MT) and interlingual subtitling. MT can be classified into three main categories. Rule-based MT (RBMT) has its foundations on written rules and vocabulary lists, a type of architecture popular in the 1970s and 1980s, while statistical MT (SMT) is based on statistical probabilities resulting from parallel corpora, and neural MT (NMT), launched in 2016, is founded on deep learning in neural networks. Many scholars in AVT have investigated the application of these MT systems to interlingual subtitling with an emphasis on the quality of the raw output, the role of post-editing, and the potential gains in productivity for subtitlers using MT as opposed to human translation (Nyberg and Mitamura 1997, Popowich *et al.* 2000, Melero *et al.* 2006, Armstrong *et al.* 2006, De Sousa *et al.* 2011, Sawaf 2012, Athanasiadi 2017).

Research projects have also been carried out since the early 2000s on the development of MT engines for interlingual subtitling. One of the pioneering projects in this field, MUSA (Multilingual Subtitling of Multimedia Content, <http://sifnos.ilsp.gr/musa>), 2002-2004, had the rather ambitious purpose at the time of creating 'a multimedia system that would convert the audio stream of audiovisual programmes into text transcriptions with the help of a speech recognition system' (Díaz Cintas 2015: 639). Harnessing the power of ASR and RBMT, combined with translation memories (TMs), the ensuing output was to be condensed into

subtitles subsequently translated into either English, French, or Greek. The disappointing results were mainly due to the low efficiency of the MT system, which was still underdeveloped in those years to attain the project's goals.

To enhance the democratization of information and reach out to people with hearing disabilities or who do not speak other languages, YouTube's accessibility services started in 2008 with the development of a caption feature allowing users to add subtitles to their audiovisual content. After a short pilot study, the auto-captioning component was officially integrated in 2010, a technology which relies on Google's Voice Search and that works best with a clearly spoken track. YouTube also provides an auto-translate feature powered by Google Translate, a popular MT system providing real-time translation of video captions by simply clicking on the CC button and selecting the language of the user's choice from a list (Harrenstien 2009).

Another ground-breaking initiative exploring the automation of subtitling is SUMAT (Subtitling by Machine Translation, https://cordis.europa.eu/project/rcn/191741_en.html), which ran from 2011 until 2014 (Etchegoyhen *et al.* 2014). A partnership of subtitling companies and technical partners, the project employed cloud-based SMT engines to translate subtitles in seven bi-directional language pairs thanks to a corpus of monolingual and parallel subtitles that had been collected by the participating subtitling companies (Bywood *et al.* 2017: 496). The output from the automated process was followed by human post-editing, in an attempt to optimize the subtitling workflow and the resulting quality. According to some of the experiments conducted, subtitlers considered that 56.79% of the subtitles 'required little to no post-editing to be of a publishable standard' (*ibid.*). Yet, despite these positive results, issues pertaining to the copyright of the subtitles used in the building of the engine means that the system has never been commercialized.

Also funded by the European Commission, from 2015 until 2018, TraMOOC (Translation for Massive Open Online Courses, <http://tramooc.eu>) is a collaborative project aimed at providing high quality MT services based on neural translation architecture for online courses, in 11 language pairs. By employing OS technology and crowdsourcing, TraMOOC managed to build a platform enabling the integration of any MT system into a specific educational environment. Along similar lines, and sharing the same technology that powers poliTrans (<https://politrans.upv.es/?page=faq>), an online platform for automated and assisted multilingual media subtitling, the European Multiple MOOC Aggregator project (EMMA, <https://platform.europeanmoocs.eu>), has also probed the value of ASR and MT in the transcription, translation and subtitling of educational, online video content (Brouns *et al.* 2015).

The already mentioned ALST project (Matamala 2016) has also explored the potential of MT in the production of AD for fictional works as well as in the translation of documentaries, using off-screen dubbing and voiceover. In the latter case, given that the raw output obtained from the automated process was not fit for broadcasting, the researchers tried to measure the effort invested in the post-editing of said output, by comparing the effort involved in the translation of two short wildlife documentary excerpts with the effort and time spent in post-editing the MT material. Their findings suggest that post-editing is a faster activity, that demands less technical and cognitive effort on the part of the individual. According to the researchers, no substantial differences are noticeable between the overall quality of the post-edited version and the one translated by a human without the assistance of MT.

The usefulness of CAT tools for the translation of audiovisual programmes is still relatively underexplored. For instance, the support of TMs and glossaries, especially for the translation of specific audiovisual genres such as corporate videos, technical and scientific documentaries and any audiovisual text with a high percentage of linguistic repetition, could speed up the subtitling process substantially. In the case of other genres, such as traditional linear and non-linear procedural dramas with an episodic format and made up of many seasons, CAT tools could help strengthen cohesion across episodes when it comes to the translation of idiosyncratic expressions, certain taboo words, or proper names of people and places that are often repeated. Paradoxically, experiments exploring the potential of TMs and other CAT tools in the field of AVT have started to be carried out rather recently, well after numerous trials with MT had already been conducted.

In one of the earliest articles on the use of TMs and MT in interlingual subtitling, O'Hagan (2003: 14) embarked on an exploration to gauge 'the applicability of TM in a context of serialized films where the subtitler could leverage the subtitles produced for the previous episode' as well as to test 'the usability of freely available MT for creating subtitles mainly by non-professional subtitlers'. The results showed that both the TM tool and the raw output from the MT engine were of rather limited benefit to the respective intended users due to the lack of repetitions between film sequels for TM and the low quality of MT. However, a potential use of TM was suggested in the form of advanced bilingual concordance features while there was a slight difference in MT quality between the two different films tested, warranting further investigation.

A decade later, Pérez Rojas (2014) describes a methodology to automatically create TMs that could help in the production of interlingual subtitles. Her primary objective is to test whether the data contained in novels that have been adapted into films can be utilised as reference material for the translation of the subtitles. To this effect, the author makes use of a corpus of twenty film-adapted novels originally written in English and their translation in Spanish, exploiting, in the extraction of the texts, certain extra-linguistic markers to differentiate characters' dialogical interventions from narration. The aligned bi-texts were subsequently used as TMs in the subtitling of the adapted films. The results are rather mixed and while the success rate of the dialogue extraction from the novels is of 95% and 85% for the English and Spanish works respectively, with an overall 90% alignment accuracy, the efficiency of the translation memory was rather modest as hits between 70% and 100% matches only accounted for 15% of the corpus.

Working with a different audiovisual text type and translation mode, which the authors call 'off-dubbing', Hanoulle *et al.* (2015) focus on the potential offered by terminology-extraction systems for the translation of natural science documentaries; a genre characterized by the prolific use of technical terms and domain-specific terminology as well as a high frequency level of repetitions. With the help of SDL Multiterm Extract 2011 Trados, based on a statistical approach, and Similis, a hybrid system to perform automatic term detection, the scholars set out to investigate the extent to which the integration of bilingual glossaries into the translation process can reduce the translators' workload and/or the number of terminological errors they make. By evaluating the performance of a number of postgraduate students, the authors conclude that specialised glossaries do reduce the students' process time significantly but not the number of terminological errors.

In recent years, developers like Star AG (www.star-uk.co.uk) have adapted their TM system, Transit NXT, to work as a tool for subtitlers, providing video in context via a built-in media player, whilst at the same time allowing translators access to their TM resources for subtitling. Along the same line, MemoQ video preview (<https://help.memoq.com/8-5/en/memoQ-video-preview-tool/memoq-video-preview-tool.html>) supports the rendering of videos within the interface of the translation editor as well as the translation of subtitles in Excel or .srt format. Wordbee's online subtitle translation feature (<https://www.wordbee.com/blog/wordbee-features/new-video-module-for-subtitle-preview-in-wordbees-cat-editor>) permits translators to preview video clips when translating so that they can benefit from the visual context in which the translation will fit, while at the same time enjoying Wordbee's TM, terminologies and MT features.

Migrating to the Cloud

The underlying philosophy of cloud computing is to move away from the development of applications that reside on a client's computer or a server and create a virtual space between the devices, where the apps are shared. The ubiquitous and mysterious cloud can be conceived as a metaphor for the World Wide Web. Devices such as laptops, smartphones, PCs, television sets and tablets can connect to the cloud from anywhere and everywhere, allowing users to perform a wide range of activities: store data (Dropbox, Google Drive), translate (Memsource, XTM Cloud, MateCat), subtitle (Oona, Amara), play online computer games, and stream videos (Amazon Prime, Hulu, Netflix, YouTube Premium), among many

others. Despite its many advantages, the shift to the cloud also increases companies' vulnerability and heightens threats like cyberpiracy, against which advanced encryption technology is normally used.

In the specific field of subtitling, cloud-based platforms have been exploited from some time now for collaborative initiatives. Indeed, when it comes to the production of subtitles, whether intralingual or interlingual, the traditional paradigm of a translation company that commissions a project from professional subtitlers and pays them for their work has ceased to be the only model in existence. In today's world, viewers are also bound to come across subtitles that nobody has commissioned or paid for, as in the already discussed fansubbing scenario. Yet, a third option are the subtitles requested from volunteers, who are not paid for their work, a proposition known as crowdsubtitling. As it happens, some of the first subtitling initiatives conducted in cloud-based environments revolved around crowdsubtitling.

In this context of what Díaz Cintas (2018) calls altruist subtitling, crowdsubtitling usually refers to the production of collaborative, not-for-profit subtitles initiated and powered by specific organizations or teams of volunteers. From a technical perspective, they often use online applications or platforms built for the specific purpose of subtitling and which are very easy to learn and use – a simplified version of the more sophisticated professional tools later developed –, as is the case of Amara (<https://amara.org>) or Dotsub (<https://dotsub.com>). Free, co-creative platforms are used by ventures like TED Talks (www.ted.com) to make their videos widely accessible around the world in umpteen languages. In their standard mode of operation, participants are not normally allowed to decide the timing of the individual subtitles and they are instead given a list of master subtitles to translate into their working language, thus concentrating on the purely linguistic transfer. In this way, the subtitling process is relatively fast and easy as the technical dimension has been performed by someone else and no software needs to be downloaded or installed on their PCs. From an ethical perspective, the main difference with fansubbing lies in the fact that in crowdsubtitling both clips and subtitles are shared on open websites like TED, Khan Academy (www.khanacademy.org) or Viki (www.viki.com), with the consent of the interested parties.

When it comes to the audiovisual industry, as highlighted by Georgakopoulou (2012: 1), one of the challenges is the 'ever-increasing demand for audiovisual translation services', while at the same time being 'forced to contend with the reduction of budgets as well as the contraction of timeframes in which these services need to be provided'. To survive under these conditions and to remain competitive, LSPs need to evolve and many have started exploring solutions in the cloud.

The increase in subtitling visibility, propelled by the rise in the demand for subtitles in the entertainment industry, is partly due to the advent of VOD streaming systems. At the National Association of Broadcasters show held in 2017, executives from some of the largest media producers and distributors shared their predictions for the future of the industry, focused on ubiquity, the potential offered by the cloud, the exponential growth of audiovisual content, and the need to release the productions around the world at the same time, in what is known as simulcasting (Krefetz 2017). As already discussed, the development of the Web 2.0 has revolutionised the traditional professional translation processes, challenging the concept of high-quality as consumers' preferences seem to have shifted towards immediacy, greater interactivity and lower costs. To meet some of these new demands, the industry has been experimenting with the automation and outsourcing of the translation process, which in turn has prompted companies specialising in translation solutions to move to the cloud.

The AVT profession is witnessing profound changes due to the very nature of the technology-driven *modus operandi* in place, which is increasingly being articulated around a global pool of localization teams connected to proprietary cloud-based platforms, with the ultimate goal of improving speed, efficiency and scalability. In this agile environment, subtitling has been the preferred translation mode chosen by most developers to test the new waters. Indeed, professional clouds subtitling is a recent trend that refers to the notion of subtitling on the cloud through collaboration among people based in different geographical locations. Essentially, clouds subtitling adopts a different holistic working cycle that resembles to some extent the typical chain of subtitling preparation followed by subtitling companies, but online.

These commercial providers usually employ professional subtitlers, rather than volunteers as in crowdsubtitling, and offer clients the possibility of choosing a particular subtitler to take care of their project.

Cloud-based subtitling and captioning platforms are browser-based systems that assets can access with their own (multiple) devices to carry out subtitling and post-production tasks. They offer a wide range of tools to perform the cueing of subtitles, to review and proofread other colleague's translation, to convert files into the most widespread formats, and to burn subtitles and images. They also come with an automated quality control tool for (semi-)automated checks of the technical and linguistic dimensions of subtitling, cloud encryption to ensure that the content stored online is safe, and a fully visible monitoring/managing system which can optimize the company's internal workflow, from a potential entry test to select and on-board new freelance professionals to an automatic invoicing process. Some of these cloud platforms are testing automatic alignment of text with audio, whilst still allowing for subtitle editing, with options on positioning and use of colours as well as various other technical attributes that can be set by the client or the vendor.

What crowdsubtitling also brings in is the potential for closer monitoring on the part of the clients themselves, who can actively participate in the workflow by placing orders and reviewing translations; the possibility of delivering the final product in different formats with greater ease; the option of archiving and reviewing the content; and the lowering of costs thanks to its streamlined workflow management, saving time and physical space in the editing, post-production and delivery stages of the process. The entire subtitling project is managed online, through a platform that usually incorporates a project management ecosystem as well as a subtitling editor with a user-friendly application that operates as subtitlers' workspace. One of the crucial advantages for subtitlers is that they can manage their projects without having to buy or download software themselves while, in the case of project managers, they can keep track of all projects through interactive workflows. In the absence of physical offices, cloud-based platforms become the ultimate virtual workspace.

Although reasonably recent, crowdsubtitling has made rapid inroads in the industry. In the first decade of the new millennium, the interest of tech and translation companies in the potential of crowdsubtitling soon became apparent, with the first web-based proprietary system in the localization industry launched by ZOO Digital (ZOOsubs, www.zoosubs.com), back in 2009. Since then, a wide range of web-based subtitling applications have been developed by companies throughout the world. Some of these products are proprietary and can only be used by the employees of the company, as in the cases of Nordisk Undertext's Plint (www.undertext.se/plint); IYuno Media Group, which owns iMediaTrans (www.imediatrans.com), and Netflix, which has created Subtitle Originator and QC 2.0 to liaise with their preferred vendors.

In other instances, the platform is available on demand to general users, as in the cases of eCaption (www.ecaption.eu) and Ooona's subtitle toolkit (<https://ooona.oonatools.tv>). The latter platform allows for a number of activities related to subtitling such as spotting, intralingual and interlingual subtitling, reviewing, converting formats, transcribing, burning and encoding. The toolkit is available in a basic as well as professional version, which offers advanced functionalities such as shot change detection, waveform representation and audio scrubbing. The final output of the process performed by the web-based architecture is not just the traditional subtitle file, but a video with embedded subtitles, something that only a few years ago would require an additional audio and video encoding piece of software.

In some of these toolkits, the transcription and linguistic transfer processes are assisted by the use of specialist CAT tools as is the case of CaptionHub (<https://captionhub.com>), a platform initially designed for the creation of closed captions and templates. It incorporates ASR software for transcribing dialogue, built-in machine-learning algorithms for auto-captioning and the creation of templates, and state-of-the-art neural machine translation for a post-editing focused workflow. Similarly, Unbabel (<https://unbabel.com/for-video>) provides a managed service that combines MT systems and human post-editing for the transcription and subtitling of audiovisual productions (see Vieira; Läubli and Green in this volume).

SubtitleNext, created by PBTEU (<https://pbteu.com>), is a hybrid solution that relies on a traditional desktop software program which communicates with the computerized version of the tool on the web, providing enhanced workflow flexibility for both companies and their assets. The system covers all subtitling related scenarios, from creation to distribution, including a project management tool to administer all tasks, and the possibility to work directly with online video distribution platforms such as YouTube and Facebook and with content where real-time subtitling is required.

As discussed, some of these toolkits incorporate machine-aided translation systems, usually available through integration with third party solutions, and hence not included in the system by default. Surprisingly, only a few make use of other tools like TMs, term bases and glossaries. This approach is more frequently adopted by LSPs, rather than subtitling companies, willing to integrate video content to their online workflow. One such example is subCloud connector, developed by the multimedia localization company hakromedia (www.memsource.com/blog/2016/02/11/hakromedia-subcloud-connector), which integrates their subtitle editing workbench subCloud with Memsource so that subtitlers can create and edit subtitles in the browser whilst benefiting from TMs, term bases and MT systems.

In the novel cloud-based ecosystem, manufacturers and developers have concentrated by and large on subtitling as their preferred translation practice. More recently, efforts are being directed to dubbing and voiceover (Green 2018), thanks to the many advances taking place on speech technology. In 2017, ZOO Digital launched ZOOdubs, a cloud-based tool capable of centralising, accelerating and globalising the dubbing process by creating empty dubbing audiotracks, the equivalent of subtitling master templates, of an audiovisual programme. Using these audiotracks, a pool of freelance translators and adapters work on the translation and adaptation of the dialogue into the target language, which is subsequently shared with the voice actors for the actual recording, all connected in real time to the online dubbing tool. In this manner, scripting, auditioning of future voice talents, casting, recording, editing and mixing can all be conducted online, as opposed to the traditional approach of performing these activities locally, in soundproof recording studios. The increase of dubbed and voiced-over content, for young and adult audiences, responds to companies' strategic decisions to differentiate their localization services in order to reach wider audiences, and such rise in volume has attracted the attention of other tech companies looking to carve their own space in the online management of the dubbing process, such as the dubbing manager being developed by Ooona (<http://ooona.net/dubbing-manager>) and MediaWen's STVHub platform (<http://stvhub.com>) to provide automatic dubbing and subtitling online solutions.

Concluding Remarks

In a constantly changing and globalized world commanded by technology, translation is not much different to other professions in the sense that its very nature is also subject to change and transformation. Activities and tasks common only a few years back might well disappear in the not-so-distant future, superseded by new ones or altered in such a way as to become intrinsically different.

Some of the technological developments described in this chapter attest to the empowerment of users, viewers, and prosumers who can easily create, manipulate, translate and distribute audiovisual content across the world, individually or as members of collaborative networks. The advantages and disadvantages of this evolution, from a professional perspective, have been discussed by Baños (2018), when she argues that for some this development is a real threat, that leads to deprofessionalization, while for others is a welcome paradigm that opens up the profession and bolsters the provision of more accessible and translated audiovisual content.

Technological accelerationism has certainly contributed to the democratization of internet access, has proved pivotal in the rapid spread of audiovisual productions, and has had a tremendous impact on the actual practice of translation and on translators' productivity. But its economic reverberations have also led to Orwellian dystopias among some

professionals, who decry the fact that technology is being developed purely for the sake of economic efficiency and fear futuristic scenarios of machines replacing human translation. As long as there has been a translation industry there has also been severe competition around craftsmanship, tools and technology, pushing the boundaries of the profession and altering the relationship among the various stakeholders. Yet, to mitigate anxiety, technology developers should concentrate more on how to assist professionals and facilitate their work than on making human efficiencies.

In such mercurial ecosystem, new entrants as well as experienced translators are expected to adapt and accommodate not only to the current technology but also to the new one still to emerge. Flexibility and adaptation in the face of innovation will be the key to professional success, bearing in mind that the path to success may not be so much in the technology itself, but rather in the innovative use the industry and the professionals therein make of it.

References

- Aliprandi, C., C. Scudellari, I. Gallucci, N. Piccinini, M. Raffaelli, A. del Pozo, A. Álvarez, H. Arzelus, R. Cassaca, T. Luis, J. Neto, C. Mendes, S. Paulo and M. Viveiros (2014) 'Automatic live subtitling: state of the art, expectations and current trends', *NAB Broadcast Engineering Conference*, 1-7. Available online: https://www.researchgate.net/publication/268069711_Automatic_Live_Subtitling_state_of_the_art_expectations_and_current_trends [last access 25 January 2019].
- Armstrong, A., A. Way, C. Caffrey, M. Flanagan, D. Kenny and M. O'Hagan (2006) 'Improving the quality of automated DVD subtitles via example-based machine translation'. *Proceedings of Translating and the Computer 28*, London: Aslib, 1-13. Available online: <http://www.mt-archive.info/Aslib-2006-Armstrong.pdf> [last access 25 January 2019].
- Athanasidi, R. (2017) 'Exploring the potential of machine translation and other language assistive tools in subtitling', in M. Deckert (ed.) *Audiovisual Translation – Research and Use*, Bern: Peter Lang, 11-28.
- Baños, R. (2018) 'Technology and audiovisual translation', in S. Chan (ed.) *An Encyclopaedia of Practical Translation and Interpreting*, Hong Kong: Chinese University Press, 15-41.
- Becket, A. (2017) 'Accelerationism: how a fringe philosophy predicted the future we live in', *The Guardian*, 11 May. Available online: <https://www.theguardian.com/world/2017/may/11/accelerationism-how-a-fringe-philosophy-predicted-the-future-we-live-in> [last access 25 January 2019].
- Böhme, G. (2008/2012) *Invasive Technification. Critical Essays in the Philosophy of Technology*, transl. by Cameron Shingleton, London: Bloomsbury.
- Brouns, F., N. Serrano Martínez-Santos, J. Civera, M. Kalz and A. Juan (2015) 'Supporting language diversity of European MOOCs with the EMMA platform', *Proceedings of the European MOOC Stakeholder Summit 2015*, 157-165. Available online: <http://dspace.ou.nl/bitstream/1820/6026/1/emooocs2015-emma-final.pdf> [last access 25 January 2019].
- Brown, A. and J. Patterson (2017) 'Designing subtitles for 360° content', *BBC, Research & Development*, 26 October. Available online: <https://www.bbc.co.uk/rd/blog/2017-03-subtitles-360-video-virtual-reality> [last access 29 December 2018].
- Bruns, A. (2008) 'The future is user-led: the path towards widespread produsage', *Fibreculture Journal* 11: 1-10.
- Bywood, L., P. Georgakopoulou and T. Etchegoyhen (2017) 'Embracing the threat: machine translation as a solution for subtitling', *Perspectives* 25(3): 492-508.
- Chaume, F. (2012) *Audiovisual Translation: Dubbing*, Manchester: St Jerome.
- De Sousa, S., W. Aziz and L. Specia (2011) 'Assessing the post-editing effort for automatic and semi-automatic translations of DVD subtitles', *Proceedings of Recent Advances in Natural Language Processing*, 97-103.

- Díaz Cintas, J. (2015) 'Technological strides in subtitling', in S. Chan (ed.) *Routledge Encyclopaedia of Translation Technology*, London: Routledge, 632-643.
- Díaz Cintas, J. (2018) 'Subtitling's a carnival: new practices in cyberspace', *The Journal of Specialised Translation* 30: 127-149.
- Díaz Cintas, J. and P. Muñoz Sánchez (2006) 'Fansubs: audiovisual translation in an amateur environment', *The Journal of Specialised Translation* 6: 37-52.
- Etchegoyhen, T., L. Bywood, M. Fishel, P. Georgakopoulou, J. Jiang and G. van Loenhout (2014) 'Machine translation for subtitling: a large-scale evaluation', *Proceedings of the Ninth International Conference on Language Resources and Evaluation LREC 2014*, Paris: ELRA, 46-53.
- Fernández-Torné, A. and A. Matamala (2015) 'Text-to-speech vs. human voiced audio description: a reception study in films dubbed into Catalan', *The Journal of Specialised Translation* 24: 61-88.
- Georgakopoulou, P. (2012) 'Challenges for the audiovisual industry in the digital age: the ever-changing needs of subtitle production', *The Journal of Specialised Translation* 17: 1-26.
- Green, S. (2018) 'How digital demand is disrupting dubbing', *M&E Journal*, March. Available online: <https://www.mesalliance.org/2018/03/15/journal-digital-demand-disrupting-dubbing> [last access 25 January 2019].
- Hanoulle, A., V. Hoste and A. Remael (2015) 'The efficacy of terminology extraction systems for the translation of documentaries', *Perspectives* 23(3): 359-374.
- Harrenstien, K. (2009) 'Automatic captions in YouTube'. Available online: <http://googleblog.blogspot.com/2009/11/automatic-captions-in-youtube.html> [last access 25 January 2019].
- Kelion, L. (2013) 'Undertexter subtitle translation site raided by police', *BBC News*, 10 July. Available online: <https://www.bbc.com/news/technology-23252523> [last access 25 January 2019].
- Krefetz, N. (2017) 'Amazon, Netflix, and more go global with the cloud', *Streaminmedia.com*. Available online: <http://www.streamingmedia.com/Articles/Editorial/Featured-Articles/Amazon-Netflix-and-More-Go-Global-With-the-Cloud-118024.aspx> [last access 25 January 2019].
- Massidda, S. (2015) *Audiovisual Translation in the Digital Age: The Italian Fansubbing Phenomenon*, Basingstoke: Palgrave Macmillan.
- Matamala, A. (2016) 'The ALST project: technologies for audio description', in A. Matamala and P. Orero (eds) *Researching Audio Description: New Approaches*, Basingstoke: Palgrave Macmillan, 268-284.
- Matamala, A. (2017) 'Mapping audiovisual translation investigations: research approaches and the role of technology' in M. Deckert (ed.) *Audiovisual Translation – Research and Use*, Bern: Peter Lang, 11-28.
- Melero, M., A. Oliver and T. Badia (2006) 'Automatic multilingual subtitling in the eTITLE Project', *Proceedings of Translating and the Computer* 28, London: Aslib, 1-18. Available online: <http://www.mt-archive.info/Aslib-2006-Melero.pdf> [last access 25 January 2019].
- Nyberg, E. and T. Mitamura (1997) 'A real-time MT system for translating broadcast captions', *Proceedings of the Sixth Machine Translation Summit*, 51-57. Available online: <http://www.lti.cs.cmu.edu/Research/Kant/PDF/mts6.pdf> [last access 25 January 2019].
- O'Hagan, M. (2003) 'Can language technology respond to the subtitler's dilemma? A preliminary study', *Translating and the Computer* 25, London: Aslib, 1-18. Available online: <http://mt-archive.info/Aslib-2003-OHagan.pdf> [last access 25 January 2019].
- O'Reilly, T. (2006) 'Web 2.0 compact definition: trying again', *Radar, Insight, Analysis and Research about Emerging Technologies*, 10 December. Available online: <http://radar.oreilly.com/2006/12/web-20-compact-definition-tryi.html> [last access 25 January 2019].

- Pérez Rojas, K. (2014) 'Automatically building translation memories for subtitling', in E. Torres-Simón and D. Orrego-Carmona (eds) *Translation Research Projects 5*, Tarragona: Universitat Rovira i Virgili, 51-61.
- Perrino, S. (2009) 'User-generated translation: the future of translation in a Web 2.0 environment', *The Journal of Specialised Translation* 12: 55-78.
- Popowich, F., P. McFetridge, D. Turcato and J. Toole (2000) 'Machine translation of closed captions', *Machine Translation* 15(4): 311-341.
- Romero-Fresco, P. (2011) *Subtitling through Speech Recognition: Respeaking*, Manchester: St Jerome.
- Sawaf, H. (2012) 'Automatic speech recognition and hybrid machine translation for high-quality closed-captioning and subtitling for video broadcast', *Proceedings of Association for Machine Translation in the Americas – AMTA*, 1-5. Available online: <https://pdfs.semanticscholar.org/2eeb/01f115f3418fc5ab752bea719b6883047284.pdf> [last access 25 January 2019].
- Schwab, K. (2016) *The Fourth Industrial Revolution*, Geneva: World Economic Forum.
- Scott, S. (2018) 'Netflix launch synthesised audio description'. *DoubleTap*. Available online: <http://doubletap.online/netflix-launch-synthesised-audio-description?platform=hootsuite> [last access 25 January 2019].
- Szarkowska, A. (2011) 'Text-to-speech audio description: towards wider availability of AD', *The Journal of Specialised Translation* 15: 142-162.
- Toffler, A. (1980) *The Third Wave*, London: Pan Books.
- Verboom, M., D. Crombie, E. Dijk and M. Theunisz (2002) 'Spoken subtitles: making subtitled TV programmes accessible', in K. Miesenberger, J. Klaus and W. Zagler (eds) *Computers Helping People with Special Needs. ICCHP 2002*, Berlin: Springer-Verlag, 295-302.

Further reading

- Baños, R. (2018) 'Technology and audiovisual translation', in S. Chan (ed.) *An Encyclopaedia of Practical Translation and Interpreting*, Hong Kong: Chinese University Press, 15-41.

This article focuses on the role played by technology in the field of audiovisual translation, both from an industry as well as research perspectives.

- Chan, S. (ed.) (2015) *Routledge Encyclopaedia of Translation Technology*, London: Routledge.

Divided into three parts, this book provides a comprehensive overview of the field of computer-assisted translation. In particular, the article by Díaz Cintas offers a pertinent discussion on the implications that technology has on the practice and research of subtitling.

Biographies

Jorge Díaz Cintas is Professor of Translation and founder director of the Centre for Translation Studies (CenTraS) at University College London. He is the author of numerous articles, special issues and books on audiovisual translation. He serves as the Chief Editor of the series *New Trends in Translation Studies* and is a member of the Language Industry Expert Group of the European Commission. He is the recipient of the Jan Ivarsson Award (2014) and the Xènia Martínez Award (2015) for invaluable services to the field of audiovisual translation.

Serenella Massidda is a Lecturer in Audiovisual Translation at Roehampton University and Honorary Research Associate at CenTraS, University College London. She holds a European Doctorate in Audiovisual Translation by the University of Sassari. She is the author of *Audiovisual Translation in the Digital Age – The Italian Fansubbing Phenomenon*, published by Palgrave MacMillan in 2015. She is also a professional translator and subtitler, and a member of ESIST (European Association for Studies in Screen Translation).