Audiovisual translation

Luis Pérez-González

Audiovisual translation focuses on the practices, processes and products that are involved in or result from the transfer of multimodal and multimedial content across languages and/or cultures. Audiovisual texts are multimodal inasmuch as their production and interpretation relies on the combined deployment of a wide range of semiotic resources or modes (Baldry and Thibault 2006), including language, image, music, colour and perspective. They are multimedial insofar as this panoply of modes is delivered to the viewer through various media in a synchronized manner (Negroponte 1991). Since the turn of the century, the combined effect of technological and statutory developments has prompted a significant and rapid expansion in the provision and study of assistive forms of audiovisual translation that aim to facilitate access to media content for sensory-impaired viewers. In assistive translations, the meaning conveyed in the source text through acoustic or visual signifying means is re-encoded in written or spoken language – in subtitles for the hard of hearing and audio description for the blind, respectively. Concerned with intersemiotic rather than interlingual transfers of meaning, accessibility-driven practices have significantly widened the scope of audiovisual translation as a field of professional practice and domain of scholarly enquiry.

Throughout the second half of the twentieth century, the most established forms of audiovisual translation, whose emergence and consolidation are inextricably entwined with the popularization of film for mass entertainment, were adopted by other media such as television and, since the 1980s, software applications and videogames. This increasing diversity of audiovisual content has been mirrored in the changing terminology used to designate this area of study (Chaume 2004). As subtitling and dubbing were born on the back of sound motion pictures, the terms film dubbing and film translation featured prominently in early scholarly work (Fodor 1976; Snell-Hornby 1988). The subsequent expansion of television broadcasting provided new avenues for the dissemination of translated audiovisual texts, with labels such as film and TV translation (Delabastita 1989) and media translation (Eguíluz et al. 1994) gaining visibility in the literature. Following the exponential growth in the volume of audiovisual texts produced by and for electronic and digital media, terms such as screen translation (Mason 1989; O’Connell 2007), multimedia translation (Gambier and Gottlieb 2001) and multimodal translation (Pérez-González 2014a) have emerged, illustrating the extent to which audiovisual translation has outgrown its core domain of enquiry and annexed neighbouring fields under an all-inclusive research agenda.
History and modalities

Even during the silent film era, exporting films to foreign markets involved some form of interlingual mediation. The turn of the twentieth century witnessed the incorporation of written language into the conglomerate of film semiotics in the form of intertitles (Ivarsson 2002; Pérez-González 2014b; O’Sullivan and Cornu 2018). The use of intertitles placed between film frames grew in parallel with the emergence of increasingly complex filmic narratives. Intertitles situated the action in a specific temporal and spatial setting, provided viewers with insights into the characters’ inner thoughts and helped them negotiate the discrepancies between screen time and real time during a period when filmic techniques were rudimentary (Panofsky 1934; Dick 1990). Removing the original intertitles and inserting a new set of target language texts back into the film was all that was required to exploit it commercially in a foreign market. But intertitles also served as the springboard for the development of new forms of audiovisual translation. In-house commentators were employed to fill the same gaps as the intertitles (Dreyer-Sfard 1965; Nornes 2007), although these entertainers often sought to enhance the viewer’s experience by spreading gossip about film stars or even explaining how the projector worked (Cazdyn 2004). The national industries of the USA and a number of European countries – the main players during the formative period of audiovisual translation as we know it today – thrived on this absence of linguistic barriers to film exports, until the aftermath of World War I hampered the financial capability of European industry to fund new projects. By the early 1920s, American films had come to secure a dominant market share throughout Europe, pushing some national film industries close to the brink of collapse (Novell-Smith and Ricci 1998).

According to Forbes and Street (2000), the advent of sound in the late 1920s put a temporary end to American domination of European film industries, as the big studios became suddenly unable to satisfy the demand of European audiences for films spoken in their native languages. Experimental attempts to appeal to local European sensibilities failed to earn the American industry its lost market share back, and it soon became obvious that new forms of audiovisual translation were required to reassert its former dominance. During the second half of the 1920s, technological developments made it possible to revoice certain fragments of dialogue or edit the sound in scenes that had been shot in noisy environments through post-synchronization (Whitman-Linsen 1992; Chaves 2000). Despite being conceived as a means of improving the quality of an original recording, post-synchronized revoicing was soon used to replace the source dialogue with a translated version, and is therefore acknowledged as the immediate forerunner of dubbing (Lewin 1931; Šurovičová 2003). Concurrent advances in the manipulation of celluloid films during the 1920s allowed distributors to superimpose titles straight onto the film strip images through optical and mechanical means (Ivarsson 2002). By the late 1920s it had become customary to use this evolved version of the primitive intertitles to provide a translation of the source dialogue in synchrony with the relevant fragment of speech, thus paving the way for the development of modern subtitling.
The perfection of these new techniques and their acceptance by European audiences ended the moratorium on American control of European and other international markets (Forbes and Street 2000), with American films regaining a market share of 70 per cent in Europe and Latin America by 1937 (Chaves 2000). This second wave of domination was regarded as a threat not only to the sustainability of Europe’s national film industries, but also to their respective languages, cultures and political regimes. The multiplicity of European interests and ideologies would soon lead each country to adopt its own protectionist measures and/or censorship mechanisms (Novell-Smith and Ricci 1998). In a number of cases, these took the form of licensing policies that restricted imports of American films, or established ratios of foreign-to-domestic film screenings. Protectionism and censorship also informed the choice of preferred modalities of audiovisual translation in each country, based on financial considerations and the degree of interference in the source text that each modality allowed for. Despite these efforts, and except for brief exceptional periods like that of World War II, these dynamics of domination were to remain unchanged until the advent of digital culture.

The increasing digitization of audiovisual commodities witnessed since the middle of the first decade of the twenty-first century led to the fragmentation of national audiences and the formation of geographically dispersed audienceships (Pérez-González 2014b). In this new media landscape, non-linear or decentralized models of distribution and consumption of audiovisual content are eroding the traditional dominance of Anglophone media industries and fostering global media flows from and into a widening range of source and target languages, often outside commercial circuits. As this relative devolution of power from media corporations to consumers continues to gather pace, viewers are engaging in networked co-creational practices, including collaborative volunteer translation of audiovisual texts, to resist the political, ideological and aesthetic homogenization that industrial audiovisual translation has traditionally favoured.

In terms of modality, subtitling consists of the production of snippets of written text (subtitles, or captions in American English) to be superimposed on visual footage, normally near the bottom of the frame. The literature distinguishes three types of subtitling: interlingual, intralingual and respeaking.

Interlingual subtitling provides a written rendition of the source speech in the target language. In communities where at least two languages co-exist, bilingual subtitles deliver two language versions of the same source fragment, one in each of the two constitutive lines of the subtitle (Gambier 2003). Each of the fragments into which subtitlers divide the speech for the purposes of translation must be delivered concurrently with its written rendition in the target language via the subtitle. The “technical constraints of shortage of screen space and lack of time” (O’Connell 1998:67) have been shown to “compromise the dynamics of dramatic characterization envisaged by the creator of the original audiovisual text” (Pérez-González 2014b:16). The linguistic condensation inherent to the subtitling process, which often disambiguates the pragmatic complexity and/or indirectness of the source speech, may change the way in which characters are perceived by the target audience.
Intralingual subtitling, where captions are composed in the same language as the source speech, plays an important role in fostering the integration of minorities such as immigrants in their host communities. The fact that it is of particular importance for sensory impaired viewers explains why the term intralingual subtitling has become almost synonymous with subtitling for the deaf and hard of hearing. This form of intralingual transfer provides a text display of the speech but also incorporates descriptions of sound features which are not accessible to hard-of-hearing viewers. To compensate for its higher lexical density (Wurm 2007), intralingual subtitling has been traditionally bound by idiosyncratic conventions in terms of timing, text positioning and use of colours (Neves 2005). But ongoing technological and regulatory advances are prompting a more user-centred conceptualization of subtitling for the hard of hearing. Neves (2018), for example, advocates a shift towards terms that place less emphasis on viewers’ impairments. To remove the stigma of disability and reflect the growing diversity of devices mediating the experiences of deaf and hard-of-hearing viewers, she proposes two terms: enriched subtitling, which “speaks for all the added elements that make subtitles relevant to specific users”, and responsive subtitling, which reflects how today’s subtitles “travel across platforms and media” (ibid.:84).

Although the use of interlingual and intralingual subtitles was for a long time restricted only to films and programmes recorded in advance, the development of real time or live subtitling technologies, ranging from the stenograph and stenotyping methods to speech recognition systems (Lambourne 2006; Romero-Fresco 2011), has increased the accessibility of live news, live chat shows and reality TV, among other audiovisual genres. Defined as “the production of [intralingual] subtitles by means of speech recognition” (Romero-Fresco 2018:107), respeaking involves repeating the original sound of a live programme into a microphone connected to a computer equipped with a voice-recognition software application. Having rapidly gained prominence as the preferred method for the production of live subtitles, respeaking has the potential to be used interlingually in future (ibid.:108).

Technically, revoicing (Luyken et al. 1991; Baker and Hochel 1998) is an umbrella term designating a variety of oral language transfer: voice-over, narration, audio description, free commentary, simultaneous interpreting and lip-synchronized dubbing. In practice, lip-synchronized dubbing, commonly referred to as dubbing, is dealt with as a distinct modality. Although all revoicing methods involve a greater or lesser degree of synchronization between soundtrack and on-screen images, the need for synchronization is particularly important in the case of dubbing.

Voice-over or ‘half-dubbing’ (Gambier 2003) involves pre-recorded revoicing: after a few seconds in which the original sound is fully audible, the volume is lowered and the voice reading the translation becomes prominent. This combination of realism (the original sound remains available in the acoustic background throughout) and almost full translation of the original text (Luyken et al. 1991) makes voice-over particularly suitable for interviews, documentaries and other programmes which do not require lip synchronization. Voice-over is also used today to translate feature films for some relatively small audiovisual markets in Europe and Asia because it is substantially
cheaper than dubbing (O’Connell 2007). It is executed differently in each of these target contexts (Naficy 2003; Szarkowska 2009; Franco et al. 2010; Matamala 2018), depending, for example, on whether the voice talent is expected to mimic the prosodic features of the original speech.

Narration aims to provide a summarized but faithful and carefully scripted rendition of the original speech, and its delivery is carefully timed to avoid any clash with the visual syntax of the programme. Although its use was first documented in the 1940s (Perego 2018), audio description, understood as a form of mostly intralingual narration incorporated into media content to facilitate the accessibility of audiovisual products for the visually impaired, has grown exponentially in the second decade of the twentieth century. An audio description is a spoken account of those visual aspects of a film which play a role in conveying its plot, rather than a translation of linguistic content. The voice of an audio describer delivers this additional narrative between stretches of dialogue.

Other forms of revoicing are performed on the spot by interpreters, presenters or commentators by superimposing their voices over the original sound. Free commentary, for example, involves adapting the source speech to meet the needs of the target audience, rather than attempting to convey its content faithfully (Gambier 2003). Commentaries are commonly used to broadcast high-profile events with a spontaneous tone. Simultaneous interpreting is typically carried out in the context of film festivals when time and budget constraints do not allow for a more elaborate form of oral or written language transfer. Interpreters may translate with or without scripts and dub the voices of the whole cast of characters featuring in the film (Lecuona 1994).

Finally, lip-synchronized (or lip-sync) dubbing is one of the two dominant forms of film translation, together with interlingual subtitling. In the field of audiovisual translation, dubbing denotes the re-recording of the original voice track in the target language using dubbing actors’ voices; the dubbed dialogue aims to recreate the dynamics of the original, particularly in terms of pace of delivery and lip movements (Luyken et al. 1991). Regarded by some as the supreme and most comprehensive form of translation (Cary 1969), dubbing “requires a complex juggling of semantic content, cadence of language and technical prosody … while bowing to the prosaic constraints of the medium itself” (Whitman-Linsen 1992:103-104). Since the 1980s, the importance of lip synchrony proper has been revisited, with new and more elaborate models of dubbing synchrony advocating the need to match other features of the original film which contribute to characterization or artistic idiosyncrasy (Fodor 1976; Whitman-Linsen 1992; Herbst 1994; Chaume 2004). The idiosyncrasy of dubbed language as an instance of prefabricated orality (Baños and Chaume 2009; Chaume 2012), the impact of dubbed voice and prosody on the reception of translated films (Bousseaux 2015) and the role of synthetic voices (Bousseaux 2018) have all attracted growing attention.
Research themes

Research on audiovisual translation follows a range of approaches, including comparative, process and causal models.

Comparative audiovisual translation research has focused primarily on the analysis of shifts and the scrutiny of corpora. Efforts to systematize the range of shifts or differences between source and target texts that arise during the translation process have allowed scholars to produce taxonomies of audiovisual translation strategies (Díaz Cintas and Remael 2007); account for types of semiotic transfer through which meaning is redistributed across visual and acoustic codes (Delabastita 1990; Lambert and Delabastita 1996); and describe rhetorical adjustments that bring the narrative configuration of the original film in line with the conventions favoured in the target culture (Cattrysse 2004). Similarly, the use of electronic corpora has enabled comparisons between spontaneous speech and translated dialogue as a means of establishing how naturalness is constructed in dubbed conversation (Pavesi 2009a, 2009b, 2018; Romero-Fresco 2006, 2009).

Process-oriented research informed by psycholinguistic models (d’Ydewalle and van de Poel 1999; Rajendran et al. 2013) is yielding insights into the impact of subtitling segmentation on the processing of language chunks that are pre-assembled as such in the brain. Cognitive models based on the notions of visual and narrative salience (Kruger 2012; Vandaele 2012; Vercauteren 2012) are also driving process studies of audio description that aim to identify the information describers should prioritize in their narrations, given that visually prominent information might not always be crucial for understanding how the plot unfolds. Likewise, neurolinguistic approaches are helping researchers explore the impact of changes in acoustic stimuli on the trajectories of viewers’ gaze as they make sense of audiovisual texts (Vilaró et al. 2012).

A causal logic seeking to establish why audiovisual translated texts exhibit certain features or are interpreted in particular ways underpins a large body of research. A plethora of studies has drawn on systems and norm-based theories to understand what dictates the choice between dubbing and subtitling in certain national markets (Delabastita 1990; Karamitrouglou 2000); or how the global dominance of American films might account for the emergence of dubbese – a variety of spoken language used only in films dubbed from English and characterized by the presence of many Anglicisms and syntactic calques, thus reflecting the subordination of the target film culture to its American counterpart. Also following a causal logic, some research has drawn on discourse analysis and pragmatics to show how translation alters the interpersonal dynamics of original filmic dialogue (Mason 2001; Remael 2003; Pérez-González 2007), and on gender studies and sociolinguistics to account for the neutralizing and standardizing impact of audiovisual translation on filmic representations of gender (De Marco 2012; Asimakoulas 2012; Flotow and Josephy-
Hernández 2018) as well as sociolectal and dialectal variation (Federici 2011; Yau 2012; Yau 2018).

Although the visibility of activist subtitling communities (Pérez-González 2010; Baker 2018) has risen significantly in recent years, fansubbing (Nornes 1999; Pérez-González 2006; Dwyer 2012) remains the most widely researched form of volunteer audiovisual translation. The growing importance and influence of audiovisual translations undertaken by networks of ordinary citizens, often with the goal of resisting the commercial dynamics of commercial media flows, is driving a new research strand informed by concepts and methods from digital media sociology. Drawing on the notion of self-mediation (Chouliaraki 2010), Pérez-González (2014) and Dwyer (2017) have explored the role that subtitling, as a form of immaterial collaborative work, is playing in the digital culture. Acting as a platform for the expression of subjective – both activist and playful – spectatorial experiences, participatory subtitling is paving the way for the introduction of ethnographic research methods (Li 2017). These should prove particularly useful to studying the online and offline dynamics of such discursive spaces, where non-affiliated individuals engage in complex negotiations of cultural identity and citizenship.

Further reading

Yields new insights into the role that audiovisual translation, particularly subtitling, carried out by ordinary people is already playing in the digital culture, drawing attention to potential new avenues of interdisciplinary research.

Delivers a comprehensive overview of audiovisual translation as a domain of research enquiry, surveying the main theories and methods that have been used and are being developed to advance knowledge in the field.

A collection of 32 chapters by different authors examining the professional and scholarly developments of both established and emerging modalities of audiovisual translation; includes a range of contributions that explore the social relevance of audiovisual translation in a variety of contexts.
References


