Audiovisual Translation and Multimodality: What Future?

Yves Gambier[1]*

University of Turku, Finland & Kaunas University of Technology, Lithuania

_____

Abstract

The future is always difficult to predict, but certain elements can help us anticipate what tomorrow could look like. Today, in different societies, multimodality thanks to digital technology is becoming a dominant format of communication. Therefore, the traditional Western concept of translation is challenged while technology creates new needs and expectations and opens up space for new practices. In fact, because of this rapid technological development (i.e., the digitalization of tools), certain concepts used in Translation Studies (TS) must be revisited: I will refer here in particular to text, sense, authorship, translation unit, and quality (and the variety of evaluation criteria in audiovisual translation/AVT). AVT, with its complex system of signs, is a relevant example in this new media landscape: its map shows a dynamic research field. It will also be questioned to a certain extent while the borders of TS are moving with the borders of a significant number of other disciplines.

_____

[1] iD https://orcid.org/0000-0002-1858-4281 | ✉ yves.gambier@utu.fi

* Professor Emeritus, School of Languages and Translation Studies, University of Turku, Finland

1. Introduction

In the last few decades, the translation landscape has been transformed, with new technical tools, new user behaviors, new relationships with verbal communication and new perceptions of translation – a task long-considered as subaltern, invisible (at least in Europe) and today practiced by a host of people, whether qualified as translators or not. Communication, encompassing oral and written interactions, is becoming increasingly multilingual, multimedia, multimodal, multicultural, multipurpose, and multi-authoring, with impacts on linguistics, culture, media, education, literature – and translation. Conventional key terms such as language, dialect, news, social media, culture (and its various forms: counter-, cyber-, emergent-, sub-, youth- and visual-, to name a few), ideology, community, power, etc. require updates and adjustments.

More attention in Translation Studies is now given to media. Media is here intended in its broad meaning – not only including newspapers, audiovisual devices and products (TV, DVDs, CD-ROM, Internet, mobile/smart phones, tablets, movies etc.) but also advertisements (printed and online commercials, video-clips etc.) and the different types of material in support of corporate communication, tourism information, entertainment (e.g., videogames, webseries, webtoons etc.) and scientific popularization (printed matter, brochures, leaflets, websites etc.). In addition to audiovisual translation (AVT), other fields have begun to give rise to research, for example, the translation of information, and more generally what is called journalistic translation (Davier & Conway, 2019), transcreation or translation into advertising (whose media have diversified since the 1970s and 1980s), intersemiotic translations, as in museums. But our focus here will be on AVT.

2. Multimodality

For a long time in the West, text and translation were perceived by their linguistic dimension alone. Thanks to text linguistics and semiotics, verbal communication has become more complex, to the point of taking on a multimodal dimension. Henceforth, modes of meaning or systems of signs contributing to meaning are seen as combining (and not as adding to each other), at the level of design, the production of a text as well as at the level of its distribution, its reception (Holz-Mänttäri, 1984).

While R. Barthes (1964) saw, in the study of an advertisement for pasta, a hierarchy between the verbal and the visual, with the multimodal approach, no priority is given a priori to one of the modes co-present in a text. These modes, like color, are also now recognized as having their own traditions and conventions (Kress & Leeuwen, 1996/2006; Kress, 2010).

A main mode can incorporate various sub-modes. Thus, music can include rhythm, melody, harmony, orchestration. Or a speech can be clarified with volume, intonation, voice quality, speech speed, pauses, as the static image can also be defined according to its colors, light, composition, or perspective.

Multimodal texts are not new, even in cultures with strong literary dominance. Indeed, a printed literary text can play on its typography (type and size of characters), its margins, its illustrated cover, the use of colors, and the presence of images, to name a few. Other examples include textbooks, user manuals, websites, comic books, games and video clips, blogs, operas, songs, advertisements, technical texts,

tags providing information about paintings in a museum, medical brochures, tourist brochures, lectures accompanied by slides, and audiovisual texts. All these texts, concerned with receivers, have often been called "adaptations" in Translation Studies rather than "translations" (limited to a textualist, linguistic vision). Multimodality thus covers a wide variety of texts and forms of communication and can now be associated with many types of translation and localization (Jakobson, 1959; Reiss, 1971). Modes of meaning (or semiotic codes) and materiality or medium/support of texts (which are not to be confused with the notion of book) have their own history and socio-cultural impact. A play can be printed, a performance on stage, a radio or television or cinematographic offer; a film is multimodal and can be multimedia (on a cinema screen, mobile phone, tablet, or on a DVD); a print magazine, as well a radio show, is multimodal and mono-medial (they can only be read/watched or heard). Kaindl (2013, pp. 261-262) has thus created a typology of translations according to whether they are intramodal (with a single mode), intermodal (with several modes), intra-medial (on a single medium), intermedial (based on several media) – each case can also be intra-cultural (captured in the same culture) or inter-/transcultural (between different cultures). The eight types strongly highlight the diversity of translations today, made possible by digital technology.

3. The map of audiovisual translation today

For a better understanding of the place and the challenges related to AVT, let us draw a map of the development of AVT over the last three decades (Figure 1). Reception here falls under "accessibility", assumed here in its broad meaning, i.e., aiming at all kinds of audiences and not only people who experience disabilities. Accessibility allows anybody to achieve specific goals with effectiveness, efficiency, and satisfaction in a specific context of use, and allows communication to go beyond any social, cognitive, age, gender divide and mental, sensory, or physical impairment.
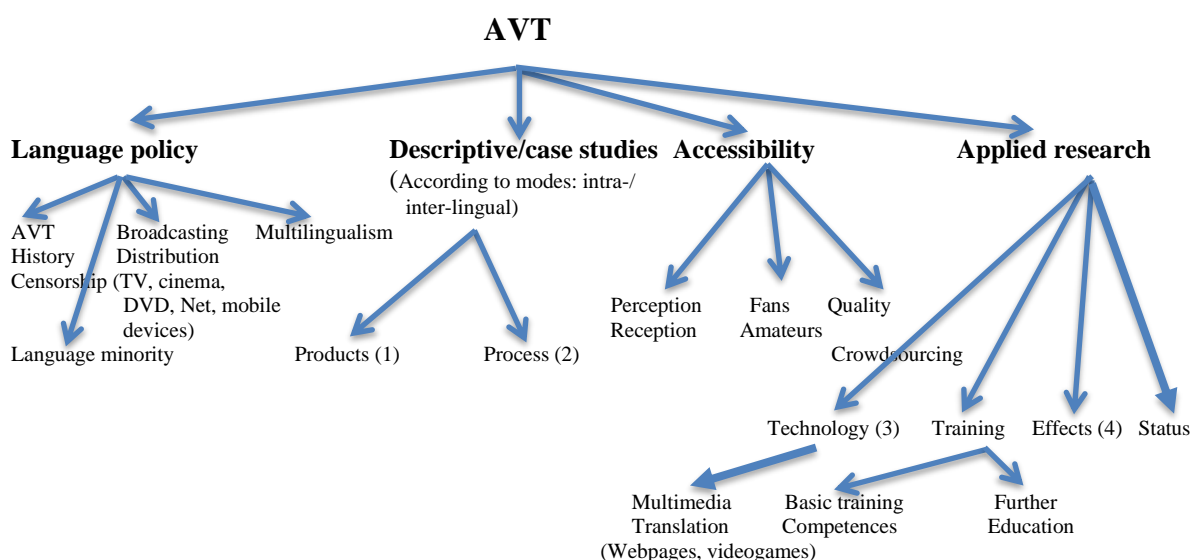

Figure 1. The map of the development of AVT over the last three decades

1) Contributions focusing on "problems" (humor, swearwords, terms of address, discourse markers, language register, cultural items, etc.) and/or "constraints" (impact of the oral code, genres, space, and time constraints, etc.).

2) Include studies on strategies, norms, conventions, the translator's voice, technical constraints, from script writing to dialogue, interplay visual/sound/verbal, etc.

3) Automation and impacts on productivity and quality, working conditions, fees, ethics, copyrights, etc. Digitization and changes in production, distribution, broadcasting, etc.

4) Effects on reading skills, reading habits, language learning, stereotypes, etc.

The four different subfields in figure 1 imply certain types of research approaches and/or tools. Thus, *Language policy* can offer historical, political, economic, and commercial perspectives while *Descriptive studies* means linguistic, pragmatic, narratological, cognitive, multimodal, imagological perspectives and case studies. With *Accessibility*, we have socio-cultural and ethnographic perspectives, among others. And *Applied Research* means technical, professional, legal, and educational perspectives.

We can view several challenges in AVT, through the holes of the map:

- How to compile, align, annotate *corpus* (parallel or not) of film subtitling? Very few exist today, such as Fortlix1, the TRACCE corpus, the Pavia corpus of film dialogue. How to set up a corpus of AD scripts, film archives? All corpora raise the issues of their representativeness, transcription, copyrights, meta-information for AV media. Corpus could impact research in automation and could help create correlations AVT strategies and film genres.

- Would observation of subtitlers be relevant for better understanding their decision-making process, taking into consideration all the different verbal and non-verbal signs on screen? A *psychosomatic approach* is still open.

- *History* of AVT. Often history is limited to repetitive arguments, such as political concerns in the studies on censorship during dictatorships in the 20th century, forgetting the weight of illiteracy then among popular audience. What about early screen translation practices, about domination of a certain mode of AVT in certain societies? Why and how does re-dubbing, re-subtitling take place?

- *English* as a source lingua franca in the international exchanges of films, TV programs and series (scenarios and pre-production scripts) and a langue-pivot for templates in subtitling.

- *Sociolinguistic role* of the translators towards minorities, in the command of native and foreign languages.

- More *Reception Studies* are required to develop a better understanding of new patterns of media participation in an ever-increasing mediated multilingual environment.

- *Processing images* (hardly ever altered when broadcasted, released) in both viewing and translating. How to define density of visual information? How to prepare the translators when digitization change images to be shown? Would they translate images as well? But also, how to identify and retrieve images in databases of film archives?

- *Automation*: To what extent could Machine Translation help translating a volume of AV products which are increasing every day? Would *open sources tools*, such as Subtitle Workshop 6.0, DivX Land Media Subtitler 2. 1.2, Avidemus 6.6.8, enhance the work and increase the number of amateurs?

- What could be the role of the translators in reproducing, adapting, manipulating *stereotypes*, clichés, and representations regarding nations, genres, races, or minorities?

- What new *theoretical references* could better describe and explain subtitling, AD, dubbing, etc.? How could Relevance Theory, Frame Theory, or Appraisal Theory be used in AVT?

4. Implications of AVT for Translation Studies

Several TS concepts should be rethought, revised, and extended when applied to AVT. We have grouped them together into three subsections.

4.1. Text, sense, and authorship

When, in a conference, a literary translator, a subtitler, a conference interpreter, and a localizer refer to "text", do they mean the same thing? Do they refer exclusively to the conventional text as a linear arrangement of sentences, or as a sequence of verbal units? Does the interpreter forget the orality, the speed of delivery, the mixture of impromptu discourse and the reading of the ready-made paper? Does the audiovisual translator forget that his "text" is maybe based on a novel (which may also be a translation), that he/she has worked from a dialogue list, watched the dialogues embodied in moving images? The screen is not only full of linguistic elements. Does the localizer forget that the source "text" was never finalized when the translation started, the source "text" was updated several times during the localization process? Does the website to be translated apply the seven criteria of textuality: cohesion, coherence, intentionality, acceptability, informativity, situationality, and intertextuality? (De Beaugrande & Dressler, 1981). The webpages are largely decontextualized, read in discontinuity by receivers with many different backgrounds and shared knowledge. In other words, the concepts of text and sense made of written lexical and morpho-syntactical devices must be revised or re-defined according to the settings and the materiality of the "text". From the historical process of delineating words, creating paragraphs, punctuation and lay-out to the interconnections between aural, oral, and visual modalities that are gaining far greater cultural prominence today, the "text" has gone through various shifts and translators have moved beyond text-to-text translation.

New formats were gradually disseminated thanks to new technology (e.g., emails, websites, SMS (short message services), chats, tweets, blogs etc.) calling for crowdsourcing and online collaborative translations. The usual concept of "text" and the interaction between text, author, reader, and translator began changing. Literature itself has cyber-formats, mixing text and media; poetry can be performed

(see slam, rap), read in public or presented in a visual exhibition. This transition from the "graphosphere" into "the video-sphere" is also becoming obvious in translation and interpreting, with the localization of videogames and the different modalities of AVT. What could be the new concept of text? (Toury, 2006).

There are differences between a text by Cicero and Virgil (to be read aloud during a public event) and a text by Proust, between a literary text, annual reports, and instructive texts even if they are all physically limited and semantically open, while hypertexts are both physically and semantically open. One does not read an electronic text without additionally referring to an interview on YouTube, a public lecture, or a map. As one does not watch a film without watching a video-clip, a trailer, or rushes beforehand. Two users can open a website with the same webpage but can end their navigation after opening different links. Three decades of the Internet and Web, including AV products, have transformed a concept which was dominant for thousands of years. From now on, "texts" are fluid, with other "texts" and other system of signs (fixed or moving images, graphs, colors, fonts, sounds, etc.). A "text" has become poly-semiotic or multimodal and exists in a permanent intertextual relationship with other "texts". While printed texts could always be dated, allocated to an author, editor or printer, digital/AV texts can be constantly updated and offered in different versions. The same uncertainty exists in news translation in which the absence of an identified source text is more frequent than not: the translation might be based, for instance, on the oral version of the speech although there is a written text. To track back to a source text can be time consuming and make collection of a corpus difficult.

Today, researchers on multimodality insist on the importance of taking into consideration the multiple modes of sense-making. Let us take the example of the multimodality of a film: 14 semiotic codes are part of the production of meaning (see Table 1).

Table 1. Semiotic codes of a film

	Audio channel	**Visual channel**
Verbal Elements **(signs)**	- *linguistic code* (dialogue, monologue, comments/voices off, reading) - *paralinguistic code* (delivery, intonation, accents) - *literary and theatre codes* (plot, narrative, sequences, drama progression, rhythm)	- *graphic code* (written forms: letters, headlines, menus, street names, intertitles, subtitles)

Non-Verbal Elements	- special sound effects/*sound arrangement code*	- *iconographic code*
(signs)	- *musical code*	- *photographic code* (lighting, perspective, colours, etc.)
	- *paralinguistic code* (voice quality, pauses, silence, volume of voice, vocal noise: crying, shouting, coughing, etc.)	- *scenographic code* (visual environment signs)
		- *film code* (shooting, framing, cutting/editing, genre conventions, etc.)
		- *kinesic code* (gestures, manners, postures, facial features, gazes, etc.)
		- *proxemic code* (movements, use of space, interpersonal distance, etc.)
		- *dress code*, including hairstyle, make up, etc.

To those 14 codes can be added "objects" (with designer imprint or branding) which participate in the funding of the movies and influence the viewers because of their emotional and symbolic connotations. Consider drinks and cars in the James Bond films, for example. All the codes are placed in a subtle relationship during the editing (montage). In the system of mixing signs, the verbal elements fulfil different functions: the translators must take them into account before deciding what to omit, condense, or make more explicit when subtitling the dialogues.

The complexity of the AV is also reflected in the concept of sense or meaning. Sense results from the interaction between signs on the screen, between those signs and the different stakeholders of the film industry, and between them and the viewers. The conventional hierarchies between producers and editors are shaken up when the "same" film can be different after the work in the cutting room, if it is planned for certain cultural audiences, certain age groups, and/or certain conditions of showing (in a cinema theatre, as a video, in a DVD format, for a flight audience or for TV channels). The final cut, with today's new technology, can depend on producers, film directors, parents, religious associations, ideological groups (who prefer politically correct projection and bowdlerized versions). A film such as *Eyes Wide Shut* (Stanley Kubrick, 1999) has not been distributed as a single version in the U.S.A, New Zealand, Australia, the U.K or the Netherlands because of the alteration of several sexually

explicit scenes during post-production or because of the verses cited from the sacred Hindu Scripture *Bhagavad Gita*. Rating systems are part of censorship. A film such as *Apocalypse Now* (Francis Ford Coppola, 1979) is another example with alternative versions, regarding the ending: the 35 mm general release version is different from the network television version, the extended version released in 2001, and the final cut shown in April 2019 – the runtime fluctuating between 2h30, 3h20 and 3 hours, respectively.

What about the concept of *authorship*? In literary studies and TS, the author is often perceived as a single individual. For a long time, the prevailing concept of authorship in literary history focused on authorial intention and originality. In AVT, the issue of authorship cannot be overlooked, since a few groups or institutions are part of the process (screen writer, producer, director, actors, sound engineers, cameraman, editors etc.). The list of credits at the beginning or end of a film is always long, encompassing intellectual and manual workers (e.g., hairdressers, carpenters, and electricians). This diversity of stakeholders manifests the diversity of semiotic signs in the meaning-making process. We may note here that in Translation Studies, different voices can be heard now in the translating process – no longer a simple face-to-face meeting between the author and the translator but a complex network of agents (literary agent, publisher, editor, copy editor, reviewer/reader, reviser, cover designer, proof-reader, typesetters, critics, librarians, booksellers etc.). In relation to authorship and voices, one can mention authorship, translation, and copyrights. From the Berne Convention (1896) to the new European Directive dated 2019, writers, artists, performers, academics, publishers, data compilers etc., are all protective of their intellectual, moral, and financial rights. Their works can be translated, adapted, arranged, broadcast, reproduced, distributed, and performed in public but only under certain conditions in order to avoid plagiarism, pirating, or copies. Within the film industry, one must realize the difficult definition of a co-production with many agents from different countries and professions (and therefore the difficulty of accessing funding) before understanding how sensitive the problem of rights really is.

4.2. Concepts of translation and translation unit

In parallel with the concept of AVT, other labels challenge the traditional Western concept of translation as a transfer and a search for equivalence. New practices, stimulated by technology and users, co-exist, and call for new denominations, such as:

- *Localization*, used first in the computing industry for software and websites and then in the videogame industry.

- *Adaptation*, for a long time in use in competition with translation, as soon as the focus is on the receiver and not on the source text, for example for comics, drama plays, children's books, advertisements, and tourist brochures. Those adaptations (or sometimes transadaptations) refer to texts with still images or music and often tackle the concept of manipulation and self-censorship.

- *Transcreation*, used in the marketing and advertising industry, mainly to underscore the creativity of international campaigns.

- *Language mediation* or how, in different multilingual and multicultural settings and between people, a change of languages occurs, a kind of code-switching that is not always controlled by a formally qualified translator or interpreter. *Translanguaging* can relate to language mediation, when multilingual speakers, using their languages in certain contexts, navigate complex social and cognitive demands by strategically employing their languages. Such a practice can take place in schools, bi- or plurilingual families, or marketplaces in a multilingual city, among others. English as a lingua franca can be one of the languages in translanguaging.

- *Transediting*, used in the print media to make clear that journalists do not "translate" but look for accuracy of information over faithfulness to the source text (if any), keep in mind the target readership and story readability, always under time pressure.

- *Multilingual technical writing* used when one writes documents in several languages from pieces of information or data (often available in English).

- *Co-drafting*, for example, of legal texts that are both legally binding.

The list is not exhaustive and does not apply everywhere at the same time. Nevertheless, it confirms that translation cannot be conceptualized by a monolithic and universal term. This diversity occurs not only within a given framework (i.e., transfer), but also with old terms in a geo-cultural area (such as mimesis, appropriation, and imitation in England, France, and Spain) and in the concepts used yesterday and today in different spaces where translation is defined as explanation, substitution, metamorphosis, or "turn around".

In interpreting, the changing practices in media settings (not only radio and TV but festivals as well, for instance) – in simultaneous or consecutive mode or with sign language – are also under scrutiny (Pöchhacker, 2018).

In addition, one should not forget media adaptation or format licensing and remakes. Game shows, reality shows, and sitcoms can be internationalized for multiple markets and produced in a new version tailored to the new audiences.

In short, within the various media, translation can be understood in different ways, because there are new types of texts, new types of relations between so-called originals and target texts and because the translation process implies the producer of the source text, the translator, and the audience (users, readers, viewers). AVT, news translation and other forms in practice contest the equivalence paradigm in TS and demonstrate why the term "translation" can be misleading.

Very often, the lay person thinks of translation in the equivalence paradigm, or the quest to convey identical meanings. The aim is to achieve a text in the target language that is "of equal value" but a word or concept may connote different meanings in another language or may be absent altogether, so the relationship between the two languages is not necessarily symmetrical. Two words may also refer to the same object, and this would not necessarily convey the intended meaning of the original text. Loyalty to the source text may result in a text that is not easily comprehensible in the target language. Some people would criticize a translation because certain words have not been replaced – thus, the famous set phrase: "Traduttore traditore". This focus on the lexical similarity of texts, however, does

not allow one to consider, describe, and explain the translation decisions and the translated output. The distinction between what is manifest (literal, direct, surface-level) and what is latent (implicit, connotative, underlying) 'misreads' the process of translation, and relegates the translator's act of interpreting the content to a task of relative obscurity.

Within TS academic studies, however, the equivalence paradigm has been criticized since the 1980s; translation theories and conceptual frameworks have shifted to include a more contextualized and socio-culturally oriented conception of the translation process. Translation has been reframed as a form of intercultural interaction. It is not languages that are translated, but rather texts that are socially and culturally situated. Within this "cultural turn" in TS, several perspectives have contributed to the critique of the long-standing equivalence paradigm – Descriptive Translation Studies (Toury, 2012); the *Skopos* theory (Reiss & Vermeer, 2013); and cultural politics (Venuti, 2018) – among others. Translation is thus viewed as a process of re-contextualization, as a purposeful action. The entire decision-making process is bound to considerations that involve the client end-receiver. Meaning is no longer considered to be a mere invariant in the source text but rather as culturally embedded, with a need to be interpreted. If scholars agree upon this broader definition of translation, which could become an umbrella term for many practices, why do we use a string of new labels?

Despite decades of academic and professional translation research, the traditional parameters configuring the equivalence paradigm persist. When scholars translate survey questionnaires, when foreign businesses discuss contracts, and viewers watch subtitled TV programs, or when language teachers use back-translation, they all rely heavily on the equivalence paradigm. Viewed from this perspective, translators are non-existent; they are passive agents, with no voice, no empathy, no subjectivity, no reflexivity, no interpreting skill, no intercultural awareness, and no qualifications, all in complete opposition to those practices described above in which translators rely on the communicative and cultural context, the objectives and functions of the translation and the target receivers. In our transition period, when digital technology is transforming texts and translation, two main concepts of translation are competing: one based on printed verbal texts (dominant since the C15th in the West), the other based on multimodal texts.

What about *translation unit*? When one subtitles or dubs, what does he/she translate? The concept of a translation unit has been defined in various ways – sometimes as the smallest verbal segment of utterance, sometimes as a unit of sense gathered from a small number of words, sometimes as a functional piece of the text combined with elements scattered in the text, sometimes as the whole text (Ballard, 2010). All those definitions are based, explicitly or otherwise, on the equivalence paradigm. Beyond this concept of a translation unit as a linguistic unit, within product-oriented research, characterizing source-text–target-text pairings with a focus on lexical items, some scholars from a process-oriented stance have defined the translation unit as a textual segment acting as cognitive input – such a source segment varies as one progresses through a text (from a single word to a word group, a clause to a sentence) depending on the translation modality (written vs oral), the text type, the translator's strategy, and possibly the ready-made segmentations provided by computer-assisted tools (e.g., translation memory software and concordancer). Whatever the theoretical framework, the translation unit always refers to verbal production. But what happens when the text is multimodal, when the meaning to be translated is dynamic, made of multiple semiotic signs? Focusing on words

and sentences can be methodologically convenient, even if it seems pointless to define a single linguistic unit. Translation and interpreting (encompassing localization, transcreation etc.) are based on a myriad of source segments, as AVT for instance pinpoints. Matters become more complex when non-linguistic signs are mixed with linguistic elements. In other words, when one subtitles, the source "text" is not the dialogue or any other linguistic item on screen but all the signs together: it is because of the interplay between verbal, visual and sound that one can decide what to omit, condense, simplify or add in the subtitles, assuming that the viewers infer from all the signs, and not only from the two lines at the bottom of the screen, the meaning of a shot and sequence. When viewers watch a film, they have a holistic approach; when subtitlers start their work, they have an analytical approach, maybe more systematic when they are novices and more routine when they are professionals. In all cases, the rhythms of the plot, the action, the dialogues, and the subtitles must coincide, the rhythm being the "binder" between all the signs. A subtitle that is too cognitively demanding at a moment when the plot is progressing rapidly disturbs the viewers and their reception of the film.
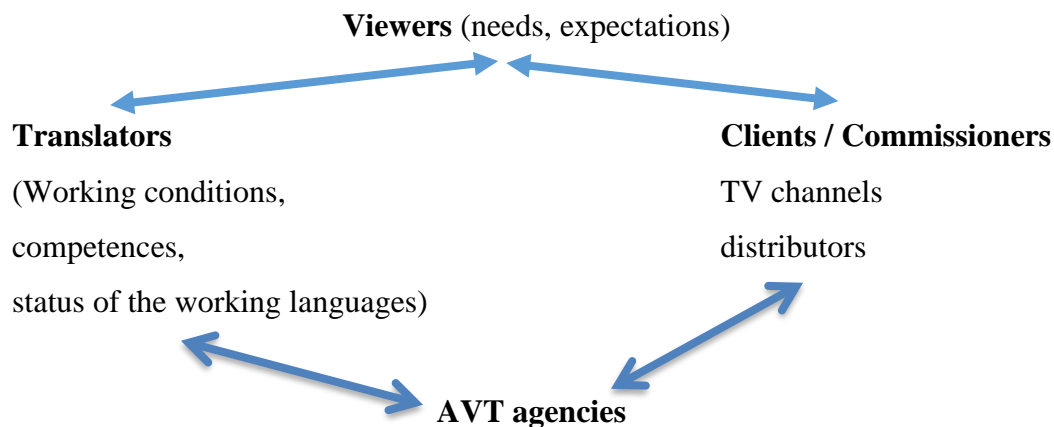
4.3. The concept of quality

There are different views and models of translation quality assessment which presuppose a certain conceptualization of translation. Most of them, based on a relationship between source and target text or a relationship between textual features and their perception by the translator, are framed within the equivalence paradigm. Many approaches focus on text (as a verbal entity), neglecting the purpose and context of the translation, as well as the different agents involved in the process, as if quality were an intrinsic characteristic and not at stake in the relationship between the commissioner, the client, the translator, the receivers, and the text to be delivered. See Gouadec (2010) who makes a distinction between the quality of a product and the quality of the transaction (the service provided).

Again, as with the concepts of text and authorship, quality assessment and control are challenged by localization of software, Machine Translation, AVT, and crowdsourced translation workflows, among others. The productivity of the translator and the type of material to be translated have changed and are still changing. A long life-span text, a highly perishable document, the raw output of a Machine Translation and a finalized and published text cannot be assessed in the same way, with the same set of evaluation criteria. In addition, quality can be assessed throughout the translation process: before signing a contract (how would the translation be able to meet the requirements of the commissioner?), during the work (to check that certain decisions are adequate for the purpose of the translation), and after the project (is it necessary to revise the translated text?). Quality (because it correlates with the time available to the translator and to the fees) takes place when planning, allocating tools and resources, recruiting freelancers, ensuring a return on investment, and measuring the impact of change.

In other words, and clearly true of all the media sectors (e.g., news translation and AVT), quality depends on a collective organization, including the client (buyers: news agencies, TV channels, film importers and distributors, and end users: readers and viewers), the language service providers (AVT companies) and the translators (staff, freelancers, outsourced agencies, subcontracted translators, and amateurs working online). Quality has many different parameters: external (needs and expectations of the stakeholders) and internal (related to organization, management, and competences). This complexity is summarized in Figures 2 and 3:

Expected Quality → Quality Sought

Client +/-   Translators'
Satisfied   Performance

Perceived Quality ← Quality Obtained

Figure 2. Types of qualities in AVT

Viewers (needs, expectations)

Translators   **Clients / Commissioners**

(Working conditions,   TV channels

competences,   distributors

status of the working languages)

AVT agencies

Figure 3. Quality and stakeholders in AVT

Beyond the set of linguistic criteria as exclusive evaluation criteria, AVT follows new challenging criteria. Among them, accessibility is a key word in AVT, not only as a legal and technical issue but also as a concept that shakes up the dominant way of assessing the quality of a translation, the aim being to optimize the user-friendliness of AVT, software, web sites and other applications. It covers a variety of features in relation to subtitles including:

- *Acceptability*, related to language norms, stylistic choice, rhetorical patterns, terminology.
- *Legibility*, defined in terms of font, position of the subtitles, subtitle rate.
- *Readability*, also defined for subtitling in terms of reading rates, reading habits, text complexity, semantic load, shot changes and speech rates.

In relation to dubbing, voice-over and free commentary:

- *Synchronicity*, including lip-synchrony (appropriateness of the speech to lip movements), syllable articulation synchrony, kinetic synchrony (utterance in relation to the gestural and facial expressions), iso-synchrony (synchronization of duration of what is said to what is shown in the pictures), and audio- or voice-synchrony (a voice, understood as vocal type, tone, timbre, pitch of the voice, prosody, accent, that matches the personality of the visible actor).

Common to both modes:

- *Relevance*, in terms of what information is to be conveyed, deleted, added, or clarified to not increase the cognitive effort involved in listening or reading.

With such predominant criteria, the usual ones (accuracy, appropriateness, and coherence) become secondary and even minor if the time and space constraints have priority. This is confirmed in reception studies dealing with different modes of AVT and using different research methods. With other multimodal texts, other criteria can be applied: for instance, *singability* for songs, operas, and musicals, and *speakability* and actability for theatre. Multi-semiotic "texts" move away from the equivalence paradigm and a purely linguistic view of translation.

4.4. Research methods in multimodality

What are the research methods on multimodality? A sketch of these methods is useful to complete the panorama that we wanted to draw of this booming field. Three types can be distinguished: methods based on Halliday's systemic functional linguistics (1978), those based on the exploitation of a corpus and those reception oriented (Tuominen et al., 2018). Obviously, audiovisual translation cannot be reduced to multimodality alone. Other approaches to AVT exist, such as the descriptive, pragmatic, empirical (observation, experimentation, testing) or even historical one.

The three meta-functions distinguished by Halliday (ideational, interpersonal, textual) have been used in many works. However, such an analysis often risks becoming monomodal, as verbal-non-verbal interaction (visual, acoustic) is a real challenge if the modes are not taken together from the outset. About corpora, we must recognize that these in AVT are few and often little provided in titles (see section 3). Without even insisting on copyright problems, these corpora then raise many questions, such as representativeness, compilation, alignment, metadata associated with a plan or sequence and their non-verbal elements, necessary annotations, to make the analysis feasible and understandable. In addition, the problem remains on how to facilitate the segmentation of AV text from 24 frames per second into (multimodal) analysis units. As for research on reception, which has recently become more popular, it is often "cumbersome" to conduct: how do we treat translated multimodal texts or, in other words, how do we read, interpret, and memorize them? What are the stimuli to consider? What is the role of redundancy in our understanding of talking images that scroll? What types of inference are used from certain visual, sound, verbal data to make sense? We must not specify only the types of reception (Gambier, 2018, pp. 55-57) but also the variables we are trying to measure, knowing that all the variables interact with each other, making it dangerous to isolate only one of them (e.g. the speed of reading subtitles at the expense of reading habits, age and education level of the viewer, characteristics of subtitling (scrolling speed, font size, readability on the screen). In any case, there are various offline (questionnaire, interview, focus group) and online (such as the use of eye tracking) ways to measure various aspects of reception.

5. What future for AVTS?

Moving beyond text-to-text translation, scholars have begun to explore different kinds of translation – text-to-image, text-to-music, image-to-text, and text-to-dance, to name a few. For them, the very terms

text, translation and other similar forms have become inadequate for describing the full range of interactions involving the production and transfer of meaning in fluid genres. If we consider texts embedded in other systems of signs, the future of audiovisual translation studies (AVTS) could be integrated as depicted in Figure 4:
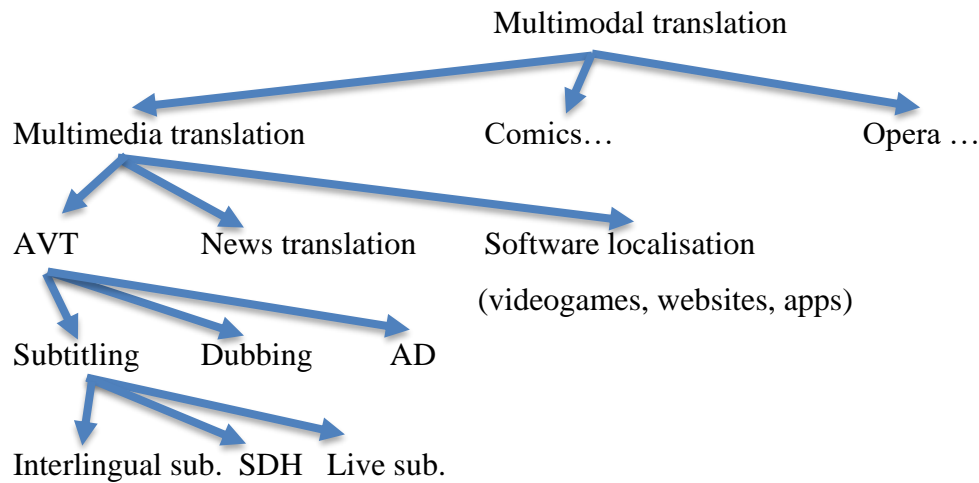
Figure 4. Future position of AVTS in multimodal translation

TS are twisted regarding three sets of questions:

1) Those coping with its scope and object of investigation: "translation". But what does it mean when many labels are supposed to describe the wide range of practices today that are found in different cultural areas, and there is an on-going debate concerning whether intralingual translation is part of TS? The proliferation of terms designating the linguistic-cultural transformation for which "translation" would once have sufficed is indicative of a conceptual disruption, of the communication value being added to the nodes of a burgeoning global-wide network and of the expansion of competences required from "translators".

2) Transformations today demand an increasingly transversal approach. Notable examples are climate change, sustainable development, cognitive sciences, artificial intelligence, public health care, bioethics, gerontology, and human-machine interactions. Obviously, the rapid changes in technology open new priorities, specializations, and communications which blur several borders. In such a perspective, do we have to maintain separated *Translation Studies, Adaptation Studies, Intercultural Studies, Transfer Studies, Media Studies, Knowledge Management, Internet Studies, Web Science,* and *Globalization Studies*, just to name a few research domains which also handle communication in its broad sense[1]? This does not mean that TS should or would swallow up neighboring disciplines. All of them acknowledge the complexity of communications and behaviors and deal with them to some extent, but they still

---

[1]. These disciplines are, in principle, partially interrelated but separated as academic disciplines, if not as university departments.

tend to ignore translation, as if information, data, knowledge, and documents could be delivered in seconds around the globe without translations!

3) In Translation Studies today, two paradigms are evolving. On the one hand, there is the more conventional conceptualization of translation which has endured for centuries through the equivalence paradigm and has evolved into one that is more oriented towards the public – audience targeted. This is the 'cultural turn' paradigm. It exists concurrently with another changing paradigm, one which reflects the platforms and mediums through which the activity of translation is now being carried out. In this sense, the paradigm of the book (upon which the paradigm of equivalence is based) transforms into one that is digital and of the Web (where the text to translate becomes multimodal). This double clash of paradigms is happening now and explains why Translation Studies has doubts about its identity, hesitating between a fragmentation into "turns" and a consilience of concepts and methods.

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

Ballard, M. (2010). Unit of translation. In Y. Gambier & L. van Doorslaer (Eds.), *Handbook of translation studies* (Vol. 1, pp. 437-440). John Benjamins.

Barthes, R. (1964). Rhétorique de l'image. *Communications, 4*, 40-51.

Davier, L., & Conway, K. (2019). Introduction. In L. Davier & K. Conway (Eds.), *Journalism and translation in the era of convergence* (pp. 1-11). John Benjamins.

De Beaugrande, R., & Dressler, W. (1981). *Introduction to text linguistics*. Longman.

Di Giovanni, E., & Gambier Y. (Eds.) (2018). *Reception studies and audiovisual translation*. John Benjamins.

Gambier, Y. (2018). Translation Studies, audiovisual translation and reception. In E. Di Giovanni & Y. Gambier (Eds.), *Reception studies and audiovisual translation* (pp. 43-66). John Benjamins.

Gambier, Y., & van Doorslaer, L. (Eds.). (2010). *Handbook of Translation Studies* (Vol. 1). John Benjamins.

Gambier, Y., & van Doorslaer, L. (Eds.). (2011). *Handbook of Translation Studies* (Vol. 2). John Benjamins.

Gambier, Y., & van Doorslaer, L. (Eds.). (2012). *Handbook of Translation Studies* (Vol. 3). John Benjamins.

Gambier, Y., & van Doorslaer, L. (Eds.). (2013). *Handbook of Translation Studies* (Vol. 4). John Benjamins.

Gambier, Y., & van Doorslaer, L. (Eds.). (2021). *Handbook of Translation Studies* (Vol. 5). John Benjamins.

Gouadec, D. (2010). Quality in translation. In Y. Gambier & L. van Doorslaer (Eds.), *Handbook of Translation Studies* (Vol. 1, pp. 270-275). John Benjamins.

Halliday, M. A. K. (1978). *Language as social semiotic: The social interpretation of language and meaning.* Edward Arnold.

Holz-Mänttäri, J. (1984). *Translatorisches handeln: Theorie und methode.* Suomalainen Tiedeakademi.

Jakobson, R. (1959). On linguistic aspects of translation. In R. A. Brower (Ed.), *On translation* (pp. 232-239). Harvard Uiversity Press.

Kaindl, K. (2013). Multimodality and translation. In C. Millán & F. Bartrina (eds.), *The Routledge handbook of translation studies* (pp. 257-269). Routledge.

Kress, G. (2010). *Multimodality: A social semiotic approach to contemporary communication.* Routledge.

Kress, G., & van Leeuwen, T. (1996/2006). *Reading images: The grammar of visual design.* Routledge.

Kruger, J. L. (2020). Translation, multimodality and cognition. In A. Jakobsen & F. Alves (Eds.), *The Routledge handbook on translation and cognition* (pp. 433-444). Routledge.

Pöchhacker, F. (2018). Media interpreting; From user expectations to audience comprehension. In E. Di Giovanni & Y. Gambier (Eds.), *Reception studies and audiovisual translation* (pp. 253-276). Hohn Benjamins.

Reiss, K. (1971). *Möglichkeiten und grenzen der übersetzungskritik: Kategorien und kriteren für eine sachgerechte beurteilung von übersetzungen* [Translation criticism - the potentials and limitations: Categories and criteria for translation quality assessment]. St Jerome.

Reiss, K., & Vermeer, H. (2013). *Towards a general theory of translational action: Skopos theory explained* (C. Nord, Trans.). St Jerome. [Original work in German, 1984]

Toury, G. (2006). Conducting research on a "wish-to-understand" basis. In J. Duarte, A. Assis Rosa, & T. Seruya (Eds.), *Translation studies at the interface of disciplines* (pp. 55-66). John Benjamins.

Toury, G. (1995/2012). *Descriptive translation studies – and beyond.* John Benjamins.

Tuominen, T., Jiménez Hurtado, C., & Ketola, A. (Eds.). (2018). *Methods for the study of multimodality in translation.* https://doi.org/10.52034/lanstts.v17i0

Venuti, L. (1995/2018). *The translator's invisibility: A history of translation.* Routledge.