

Audiovisual Translation under the Influence of the High-Tech: An Interview with Yves Gambier¹

◎ WANG Yili, Yves GAMBIER

Ningbo University, University of Turku

Abstract: This interview was conducted between Yves Gambier and Wang Yili in February 2024 in Turku, Finland, in order to give an overview of audiovisual translation (AVT) in the context of modern technology. In this interview, the current status of AVT in translation studies was discussed first, followed by the impact of technology on emerging issues such as cloning of voice, manipulation of images, and the changes of roles of audiences in AVT. The subsequent discussion focused on the difficulty in investigating audience receptions and possible solutions. The last part was an analysis of the future development of AVT.

Keywords: audiovisual translation, reception study, influence of technology

CLC: H059 **Document Code:** A **Article ID:** 2096-4374(2025)02-0118-07

DOI: 10.53397/hunnu.jflc.202502010

Yves Gambier is Professor Emeritus of the School of Language and Translation Studies at the University of Turku, Finland. He is a pioneer in audiovisual translation (AVT) research and has published significantly on the topic. His most recent contributions are coediting a special issue of *Babel* regarding multimodality and concepts of text/context and coediting with Haina Jin the newly published book *Chinese Films Abroad: Distribution and Translation*.

Wang: Professor Gambier, recently, audiovisual translation (AVT) has received a lot of attention. In the book *Audiovisual Translation: Theoretical and Methodological Challenges*, you have mentioned that AVT as an area of study within Translation Studies has been fully acknowledged and no longer causes eyebrows to be raised. What do you think is the current role of audiovisual translation and the role of translators? Have they been changed?

Gambier: About twenty years ago audiovisual translation was not acknowledged as an independent branch of translation. AVT primarily dealt with movies, and its purpose was to spread

the culture value embedded in movies, and as its translation was largely in the literary context, AVT was categorized under literary translation. Fortunately, this has changed and a considerable amount of work has been accomplished. The modalities of AVT have been significantly broadened and enhanced by a considerable number of publications, conferences, and associations solely focused on AVT. It has now risen to increased recognition within translation studies and become a dynamic and multifaceted field.

AVT research nowadays is increasingly interdisciplinary, with methods and concepts borrowed from literary studies, sociology, experimental psychology, film studies, reception studies, etc. In addition to interdisciplinary research, collaboration is also happening between the field of research and industries. The industries have felt the need for scientific supports to test the impact and outcome of certain tools or formats. Similarly, collaboration with conferences and professional associations have also been expanded from national to international levels.

Wang: As you have pointed out, the AVT landscape has been transformed from a dominant duality of dubbing and subtitling to a much more diversified one. Today, an AV product can be subtitled for the deaf and hard-of-hearing, live subtitled, surtitled, audio-described, interpreted or translated online by fans. Because of the rapid technological development—for example, the digitization of tools and artificial intelligence—AVT has been deeply impacted. In 2012, in your article “The Position of Audiovisual Translation Studies,” you mentioned five impacts of technology, including the production of AVT, new demands and formats of audiences such as short films and mobisodes, fan-subbers and amateurs subtitling, working process, and so on. How have the advancements in technology, such as machine translation and AI, influenced the practice of AVT?

Gambier: First and foremost, it is crucial to acknowledge that the advent of new technologies invariably gives rise to new challenges, which in turn call for new technologies. Presently, artificial intelligence has become a ubiquitous buzzword, saturating media outlets with articles that evoke threat by proclaiming imminent transformations in all fields. Nonetheless, it is worth noting that technology has been an integral part of our endeavors throughout the preceding. Artificial intelligence will not fundamentally alter this paradigm.

Digitization and technology could improve sound quality and allow analysis and re-synthesis of the actor’s voices but raises new problems. Lately, cloned voice has been used to assimilate the actor’s voice, in all fields of dubbing. Thus, this sort of cloned voice raises an important and new issue: voice right. The advent of new technology, including AI, has also raised another debatable question as it introduces tools allowing for the manipulation and creation of images. This technology allows for the clone of an actor’s facial expressions and well-synchronized lip movements, and could be used to serve different purposes. The safeguarding of voice right and the manipulation of virtual characters should be international concerns and a regulatory framework should be established.

The second remark pertains to the clarification of the process of AVT. In the process of subtitling, for instance, technological and AI-based automatic translation predominantly deals with linguistic elements, whereas, this does not work in AVT, where the translation process extends beyond the linguistic level to interpret visual and auditory cues. This discrepancy explains the slow adoption of automation in AVT and the not-always-high evaluation of the outcomes.

Thirdly, technology brings new modalities such as video-streaming, podcasting, and portable

devices including mobile phone and tablets. These new forms generate new demands and new needs. Instead of addressing existing problems, they inadvertently create new challenges and problems, which in turn requires the advent of new technologies. Thus, while technology and artificial intelligence undeniably provide solutions to some problems, they also introduce fresh complications. In the contemporary era of artificial intelligence, a predominant issue is the management and control of the data collected and used. This concern is particularly salient given the pivotal role data plays in shaping the performance and outcomes of AI systems.

Wang: Now that you have mentioned the role of data, how do we use these data to evaluate the quality of translation and the reception of AVT? How do you see the key dimensions or aspects of reception study when assessing the effectiveness of AVT?

Gambier: Reception is an idea which hasn't reached consensus and needs to be distinguished from perception. The concepts of "perception" and "reception" in AVT differ in their focus: perception involves the immediate sensory impressions and personal interpretations formed by viewers during the watching experience, while reception studies how AV products are processed, consumed, and remembered within specific sociocultural contexts. Reception studies aim to explain the interactions between viewers and AV products, examining the influence of AVT on the consumption patterns and preferences regarding foreign-language films and cinema attendance.

However, the reception in AVT context diverges significantly from commercial considerations. It refers to how viewers receive or interpret audiovisual content. The concept of reception must be specified when conducting research to avoid the conflation of broader influence of media studies, such as the influence of the depicted alcohol or violence on young people. In AVT, the focus is to explore the cognitive processes involved in how translations are received, evaluated and appreciated by reviewers. Studies can also be carried out on their satisfaction and their evaluation. From this perspective, reception is more an individualized experience, defined by viewers' various temporal, geographical, social, and technical settings. Because there are so many variables and factors to be considered, it is ambitious to discuss reception within an AVT context. In practice, it is very difficult to master all variables, thus you have to prioritize certain variables over others. Therefore, determination of the research focus is the critical first step in conducting a study.

Wang: Then what about the key dimensions or aspects of reception that are considered when assessing the effectiveness of AVT?

Gambier: Reception studies are complicated by numerous variables that can influence outcomes, including sociology variables such as the age of the viewers, their educational backgrounds, personal expectations, individual needs, reading habits, and so on, and AV variables like broadcasting time, types of TV channels (public/commercial), film genre, and interplay of images/dialogue. These parameters underscore the complexity of reception studies and the necessity for meticulous research approaches to accurately reflect the diverse experiences and perspectives of viewers.

The first sociocultural parameter influencing reception study is the age of the viewers. Subtitling, for instance, exemplifies how the target groups of young and old audiences require different translation strategies for the variations in their reading speed and lexical preferences.

Translators must therefore consider both textual and para-textual elements within subtitles (types of selected words, use of punctuation, etc.), which are crucial in tailoring the subtitles to the needs of target groups of different ages. Understanding these nuances is essential for optimizing AVT practices and audience reception.

The second parameter influencing reception study is the individual viewers' expectations and needs, which can significantly affect their media-based experience. For example, correct spelling in subtitles might be relevant in a given society but not universally important. Viewers of the same program would provide differing responses, making it challenging for researchers to draw universal conclusions without considering all these variables. Therefore, acceptance and tolerance of media services could vary among cultures, highlighting the caution that needs to be exercised in the generalization of any results.

Another crucial aspect to consider is the difficulty of altering cultural habits compared to consumer behaviors. While consumer preferences can adapt quickly to changes in circumstances, cultural habits are more resistant to change. For example, in Nordic countries like Finland and Norway, attempts to replace subtitling with dubbing met with difficulty, leading to a swift reversion to the traditional subtitling practice. Such instances demonstrate that cultural habits are resistant to rapid change. This insight illustrates the importance of considering cultural norms and practices in understanding viewer reception and preferences.

For each set of above-mentioned parameters, there are maybe four or five different sub-features, which indicates more than twenty parameters to consider. These parameters make it challenging to empirically assess the influence of specific features. The validity of findings from certain features could always be compromised by the rest of the parameters. Consequently, a direct correlation between controlled parameters and the other ones cannot be ensured. This again adds to the complexity of reception study within the AVT context.

In addition to the previously discussed parameters, researchers must also consider the diversity of cognitive processes among individual viewers. An individual's attention, a pivotal cognitive process, manifests in various forms. These include iconic attention, verbal attention, and narrative attention. This differentiation in attention types is crucial as it challenges the assumption that all viewers engage with audiovisual content in a universal manner. Thus, researchers must account for these varied focuses to accurately interpret data from studies on media consumption. This highlights the imperative to integrate cognitive diversity into the theoretical frameworks and methodologies of audiovisual translation studies, thereby enhancing the validity and reliability of research findings in this field.

This kind of parameter analysis limits the scope of study to relatively small populations, hindering the possibility of generalizing findings within reception studies. Thus, researchers should be cautious in data collection and in making generalizations from the findings. As Di Giovanni points out, there are hardly any large-scale studies on the reception of subtitles, which means that cooperative approaches, in which researchers would work together to expand the scope of audience, are much needed, and the potential approaches that could be used are worth exploring (58–78).

Wang: Based on Chesterman's research work, you have proposed three meanings of reception applied to AVT—response, reaction and repercussion (57). What criteria or metrics are used to evaluate audience reception of AVT? Are there any standardized evaluation frameworks or guidelines that inform this process?

Gambier: AVT reception can be understood through three primary dimensions—response, reaction, and repercussion. Each of these dimensions offers a distinct perspective on how subtitles and other forms of AVT are processed and perceived by audiences.

The first dimension, response, relates to the immediate sensory processing of visual information, such as subtitles. This aspect has been extensively studied by experimental psychologists who investigate how viewers allocate attention between images and subtitles, whether subtitles are read word-for-word, and the conditions under which viewers may re-read subtitles or avoid reading them altogether (Nikolic 179–198). Legibility concerns, which fall under this category, include the speed at which subtitles appear, their contrast against the background, and the size of the font, all of which affect the viewer’s ability to decode visual stimuli physically.

The second dimension, reaction, addresses the psycho-cognitive engagement with the subtitles. This involves the viewer’s mental processing and interpretation based on their prior knowledge, educational background, and personal experiences. These factors significantly influence translation strategies, as the greater the viewers’ processing capability, the lower the efforts of the translation becomes.

Finally, the dimension of repercussion considers the broader sociocultural and attitudinal impacts on reception. This includes examining viewer preferences and habits regarding AVT modes, as well as the values and ideologies transmitted through AV programs.

In the exploration of reception within the realm of AVT, it is also crucial to understand the concept of accessibility. Accessibility in AVT fundamentally challenges traditional metrics for evaluating translation quality by prioritizing user-friendliness in AVT applications, software, and websites.

By integrating these dimensions into a coherent framework, we can better understand the complex interplay of physiological, psychological, and contextual factors that influence how audiovisual content is received.

Wang: What could be the methodological approach in evaluating the reception in AVT? Do you primarily rely on quantitative measures or qualitative methods for evaluating audience reception of AVT?

Gambier: The investigation of reception within audiovisual translation necessitates distinct methodological approaches tailored to its multifaceted aspects. For the initial dimension—tracking individual responses—techniques such as eye-tracking and heart rate monitoring have been prevalent. However, these methodologies inherently focus on single participants and are time-intensive, limiting their utility for large-scale data collection. Moreover, the resultant data, characterized by dots and lines, presents interpretative challenges.

Similarly, the assessment of legibility confronts comparable hurdles. Given the myriad of influencing factors, relying solely on one methodological tool for a small sample size may yield results of limited relevance. Even employing triangulation—integrating eye-tracking data with interviews or questionnaires—only partially addresses these challenges. Such approaches can reveal discrepancies between perceived and actual viewer behaviors; participants might report not following subtitles, yet eye-tracking data could indicate otherwise.

The use of focus groups represents a less frequently employed but potentially fruitful method, particularly for gauging collective perceptions and experiences. This approach involves

assembling a group of non-translator participants to discuss their views under the guidance of a skilled moderator. This setting often facilitates a more natural exchange of ideas among peers compared to one-on-one interviews conducted by researchers. However, the efficacy of focus groups can vary; participants may not provide feedback as anticipated, particularly when they lack familiarity with the technical aspects of AVT. This method, while valuable, does not guarantee reliable insights into specific preferences for dubbing versus subtitling, as the discussion outcomes can be unpredictable due to the participants' general lack of expertise in translation studies.

Another major challenge in AVT research is understanding the relationships between verbal and non-verbal signs. Many researchers approach their analyses as if these signs operate along parallel lines. They assert that films are multi-semiotic entities but then treat linguistic data in isolation, overlooking the intricate and dynamic nature of the meaning-making process. This raises significant methodological problems in addressing the multiplicity of signs, with the multimodal approach being one potential solution (Taylor 192–205).

Wang: What then are the areas for future research and development in the evaluation of audience reception of AVT? How do you ensure continuous improvement in AVT practices based on insights gained from research in the 3Rs?

Gambier: In the realm of literary studies, numerous corpora are readily available, yet the field of audiovisual translation research is markedly deficient in this respect. Currently, only a limited number of AVT-specific corpora exist, such as Fortlix 1, the TRACE corpus, and the Pavia corpus for films. These corpora present several technical challenges including issues of transcription, representativeness, and meta-information, alongside legal concerns such as copyright restrictions. Predominantly, these corpora are confined to dialogues, neglecting other crucial audiovisual elements like imagery, lighting, and color schemes. This limitation constrains the scope of AVT studies, restricting the analysis to verbal content and overlooking the integral role of visual components in multimedia contexts.

Another significant limitation in current AVT research is the homogeneity of sample populations used in studies, which are often comprised predominantly of students. This demographic is not representative of the broader society, thereby skewing research outcomes and potentially limiting the generalizability of findings to wider audiences. This underscores the need for a more diverse selection of viewer samples in AVT research to ensure findings are reflective of varied audience demographics and viewing experiences.

The third and less-explored dimension of audiovisual translation (AVT) research pertains to the psychosomatic behavior of translators, particularly subtitlers. Interviews with subtitlers reveal that they perceive a necessity to physically move in sync with the rhythm of the film, including its plot, dialogues, and their own subtitles. A mismatch in these rhythms can create cognitive dissonance in viewers, potentially leading to a negative reception of the film. Subtitlers, therefore, engage in free body movements as a method to intuitively capture and reflect the film's rhythm in their subtitles. Conversely, dubbing actors—who translate not through text but through voice—often restrict their physical movements to better control their vocal delivery and intonation. This contrast in the psychosomatic approaches between subtitlers and dubbers highlights a significant yet under-investigated area in AVT research. Such research would enhance our understanding of the complex interplay between the physical engagement of the translator and the cognitive

reception of the translated media by audiences.

The utilization of a pivot language in indirect translation represents a critical aspect of contemporary AVT practices. English is usually positioned as a source lingua franca and a pivotal template for subtitling across diverse linguistic landscapes. However, the reliance on English as a pivot language introduces complexities. Industries may employ templates based on English syntax and lexical structure, operating under the assumption that all languages exhibit similar structural characteristics. While such templates might accelerate the subtitling process, the linguistic diversity and structural variations across languages can be overlooked. This can lead to increased time and financial costs, as the initial templates may require significant adjustments to accommodate languages. Thus, there is a pressing need for further research into the implications of using English as a pivot language in AVT.

The landscape of AVT is evolving beyond traditional subtitling and dubbing, encompassing over a dozen modalities including voice-over, audio description, and interpreting. The proliferation of devices and the diversification of media consumption contexts, such as mobile viewing prevalent in places like Korea, have given rise to new content formats like mobisodes. These short episodes, tailored for brief, intermittent viewing on small screens, necessitate adaptations in translation practices. While subtitles have been a cost-effective and popular choice, the small screens of mobile devices may enhance the appeal of dubbing due to legibility concerns, potentially increasing its demand and reevaluating its role in AVT.

As media consumption becomes increasingly multimodal, AVT is poised to become a critical sub-branch of multimodal translation, paralleling other fields such as literary and technical translation. This shift underscores the need for comprehensive research into how visual and auditory information is processed and translated across various media, ensuring that translation practices can adapt to the changing technological and cultural landscapes of global media consumption.

Note

1. This article is supported by Zhejiang Provincial Philosophy and Social Science Planning Project (Project Number: 25NDJC044YBMS).

Works Cited

- Di Giovanni, E. "Reception Studies in Audiovisual Translation Research: The Case of Subtitling at Film Festivals." *trans-kom*, vol. 9, no. 1, 2016, pp. 58–78.
- Gambier, Yves. "Translation Studies, Audiovisual Translation and Reception." *Reception Studies and Audiovisual Translation*, edited by Elena Di Giovanni and Yves Gambier, John Benjamins, 2018, pp. 43–68.
- Gambier, Yves, and R. S. Pinto, editors. *Audiovisual Translation: Theoretical and Methodological Challenges*. John Benjamins, 2018.
- Nikolic, K. "Reception Studies in Audiovisual Translation-interlingual Subtitling." *Reception Studies and Audiovisual Translation*, edited by Elena Di Giovanni and Yves Gambier, John Benjamins, 2018, pp. 179–198.
- Taylor, C. "Multimodal Transcription in the Analysis: Translation and Subtitling of Italian Films." *The Translator*, vol. 9, no. 2, 2003, pp. 192–205.