

# BOOK REVIEW: *INTRODUCING AUDIOVISUAL TRANSLATION*

**Mahmoud Afrouz**

Associate Professor University of  
Isfahan, Isfahan, Iran.

[m.afrouz@fgn.ui.ac.ir](mailto:m.afrouz@fgn.ui.ac.ir)

[http://orcid.](http://orcid.org/0000-0003-3051-4769)

[org/0000-0003-3051-4769](http://orcid.org/0000-0003-3051-4769)

**Editors:** Agnieszka Szarkowska & Anna  
Jankowska

**Title:** *Introducing Audiovisual Translation*

**Year of publication:** 2025

**City:** London and New York

**Publisher:** Routledge

**ISBN:** 978- 1-003- 03812- 2 (ebk)

**Number of pages:** 225

Agnieszka Szarkowska and Anna Jankowska bring in *Introducing Audiovisual Translation* comprehensive insights into the major modalities of audiovisual translation (AVT) and media localization. Both authors conduct research on AVT and serve as media localisation trainers at the university level. Szarkowska serves as faculty at the University of Warsaw, while Jankowska teaches at the University of Antwerp and serves as Editor-In-Chief of the *Journal of Audiovisual Translation*.

While noting that previous volumes have brought specific modalities of AVT to the spotlight before, the authors claim that this work stands out in providing a “comprehensive” source to introduce “all AVT modalities” (p. 1). They blend theoretical concepts illustrating them with practical examples with the aim of breaking down key principles of AVT.

This volume is divided into seven chapters with the first introducing the multifaceted world of AVT, and the ensuing chapters covering specific AVT and media accessibility (MA) modalities. Major AVT and MA types are explored by the authors, with a caveat, “that AVT encompasses more types, including, for instance, audio subtitling, sign language interpreting, or audio introductions”, and they encourage “curious readers” to study other resources, such as (2003), and Greco and Jankowska (2020) “to delve deeper” into other AVT types (p. 10).

As said above, the first chapter, titled “Audiovisual Translation,” opens the volume in defining AVT and providing as a solid entry point for students. It frames AVT as an evolving area of translation studies and provides a robust theoretical foundation, a useful typology of AVT modalities, and a compelling historical narrative.

The author acknowledges the terminological variations in the field, such as “language transfer,” “screen translation,” “film translation,” or “multimedia translation,” and considers AVT as the most academically established term (p. 5). This terminology shift may reflect broader technological and commercial trends. Meanwhile, the discussion on



Received: 2025-07-22 / Accepted: 2026-01-23 / Publicado: 2026-02-13

<https://doi.org/10.17533/udea.ikala.361595>

Editor: Luanda Sito, Universidad de Antioquia, Medellín, Colombia.

Patrimonial rights, Universidad de Antioquia, 2025. This is an open access article, distributed in compliance with the terms of the Creative Commons license BY-NC-SA 4.0 International.



*Íkala, Revista de Lenguaje y Cultura*

MEDELLÍN, COLOMBIA, VOL. 31 ISS. 1 (JANUARY-APRIL, 2026), PP. 1-6, ISSN 0123-3432

[www.udea.edu.co/ikala](http://www.udea.edu.co/ikala)

“diasemiotic” vs. “isosemiotic translation” (p. 6) could have benefited from more contemporary instances. For example, how streaming platforms manipulate these channels in interactive media. The use of Jakobson’s (2012) tripartite model to AVT is noteworthy, although we would have expected that the authors further explored the limitations of this framework and enriched their discussion by referring to critiques from scholars who might question rigid classifications in digital media environments. Moreover, a forward-looking perspective on together with a brief introduction to the emerging trends would have strengthened the section “AVT Modalities” (p. 9).

Recent innovations like fansubbing, game localisation, and AI-generated subtitling could also accompany traditional modalities presented in this section (i.e., voice-over, dubbing, subtitling, and audio description). However, it should be noted that the authors have presented very informative guidance concerning these subjects in subsequent chapters (see p. 56). As the reader engages with the content of the book, s/he will observe the authors frequent references to diverse audiences (i.e., primary and secondary audiences). For example, “intralingual subtitling is primarily designed for individuals with hearing impairment, but it is also recognized as beneficial for foreign-language learners” (p. 10).

Concerning intralingual subtitling, the issue of multiple-language versions (MLVs) is interesting, even though it might have investigated how contemporary streaming platforms, like Netflix, which produces parallel multilingual versions, are reviving this practice. The subsection entitled “Sub/dub wars” (p. 17) is engaging but it seems somehow essentialising. In fact, a deeper delve into economic or even ideological factors would have added nuance; for instance, considering dubbing as a tool of cultural hegemony.

In the section “AVT around the World”, while the authors note Netflix’s role in diversifying AVT options, there is little discussion of how

algorithmic recommendations and market forces influence translation choices. A more critical standpoint regarding the homogenising effects of globalisation could be beneficial. The concluding section “The Future” (p. 22) rightly highlights AI and machine learning as transformative forces: however, it does not sufficiently address ethical concerns, such as job displacement, and quality erosion in subtitles generated by AI or machines. Future editions could benefit from incorporating deeper critical analysis of the socio-cultural aspects of AVT practices.

In the second chapter “Pre-recorded subtitling”, translation strategies for subtitling, its tools, templates, synchronisation principles and automation are discussed. Even though the discussion could benefit from a more critical examination of how streaming platforms, like Netflix, Disney+, etc., are reshaping subtitling norms, particularly in terms of creative subtitling, one of the chapter’s strengths is its systematic discussion of the spatial and temporal constraints inherent in subtitling.

The classification of subtitling (professional vs. fansubs, open vs. closed, interlingual vs. intralingual) is pedagogically useful. The distinction between pre-recorded and live subtitling is clear, though the section the rise of AI-generated subtitles like YouTube’s auto-captions, and their implications for quality and employment in the industry might have further added to the discussion. Furthermore, a critique of Netflix’s 20 cps standard—often criticized for compromising readability—would strengthen the section “Temporal Constraints.”

Likewise, the distinction between omission and condensation is finely illustrated with examples from “Gilmore Girls” and “Sherlock” (pp. 36–38). However, readers may miss some delving deeper into ethical dilemmas like when does condensation distort characterization? Additionally, in the section “Cultural References”, although the Netflix guideline against substitution is noted, the tension between corporate mandates and creativity of the subtitler could be explored further.

As “the most target-oriented strategy” by which “a foreign cultural item is replaced with a target-culture term on the assumption that it will be easier to understand or identify with”, substitution is introduced by the authors as a strategy which can mainly lead to “cultural incongruity and mislead viewers” (pp. 43–47). When viewers, who are deeply “immersed in a foreign reality”, encounter substitution, they are likely to “experience disruption to their suspension of disbelief” (Chandler & Munday, 2016) and “feel that the fictional contract of illusion has been broken by the cultural clash” (Pedersen, 2011, as cited in Szarkowska & Jankowska, 2025, p. 43).

The authors provided three examples from different movies (i.e., *Mrs. Doubtfire*, *FNAC*, *Parasite*) just to show the negative aspects of employing substitution as a strategy for subtitling. In order to balance the discussion, however, they could at least refer to some instances where the creative strategy of substitution could possibly be used effectively—where, for example, retention of a proper name, a symbol, or a number merely results in the viewers not understanding what these culture-specific terms actually refer to. Moreover, in this section, the ethical implications of the use of some strategies like substitution deserve sharper critique. In the final section of the chapter, “The Future” of subtitling, the predictions about automation, including translation management systems (TMS), computer-assisted translation (CAT), large language models (LLMs) and MT, are pragmatic and the authors claim that “[s]ubtitling workflows will soon become increasingly automated on many different levels” is quite tangible these days (p. 56). Overall, the inclusion of real examples and international comparisons regarding subtitling enhances the understanding of concepts and increases the applicability of the information presented. Also, the chapter mostly excels as a pedagogical primer providing a thorough and well-organized introduction to pre-recorded subtitling, I would have liked a more critical treatment of issues on ethics, labor, and cultural power

dynamics, such as dealing with subtitling and power by examining fansubs vs. corporate subtitling, or address AI’s labor implications, or maybe include case studies from non-Western contexts such as Bollywood and Nollywood.

The third chapter “Dubbing” offers a thorough and well-structured exploration of “the art of dubbing” (p. 2). The distinction made in this chapter between dubbing and revoicing is useful, but the conflation of “voice-over” (VO) with “UN-style dubbing” (p. 63) risks confusion. Also, the claim that lip sync “precedes semantic equivalence” (p. 64) warrants deeper ethical scrutiny: when does prioritizing lip movements over meaning distort character intent? On the other hand, the inclusion of “dubbese” as a concept is particularly insightful, shedding light on the artificiality of dubbed dialogue and its impact on target-language norms (p. 69). In “Translation Strategies”, the ethical implications of substitution are glossed over.

In the two succeeding sections, “Dubbing Process” and “Technologies in Dubbing” are welcome additions in which the coverage of AI (e.g. deepfake lip-sync) is forward-thinking but overly optimistic—issues concerning ethical concerns (voice cloning, job displacement) are missed. In the last two sections, Chaume’s (2020) quality criteria are somewhat comprehensive, but the lack of viewer-centric metrics (e.g., audience reception studies) is noteworthy. The section on “multilingualism” (p. 81) provides a timely overview of third language (L3) and adeptly tackles L3 challenges, even though it could critique platforms like Disney+ for often subtitling L3s even in dubbed versions. Although the chapter serves as a useful resource for further familiarity with dubbing, audience reception data would have enriched the volume. Finally, it is noted AI’s potential to revolutionize dubbing and lip-syncing, even though the limitations of such technologies in capturing the nuance of human performance are underexplored.

In the fourth chapter “Voice-over”, diverse voice-over traditions and common audiovisual content types that depend on VO are meticulously reviewed, and techniques of synchronization and tips for script preparation are presented. One of the chapter’s strengths lies in its detailed discussion of “synchronization” types—primarily “isochrony,” “action synchrony,” “kinetic synchrony,” and “literal synchrony” (pp. 92–95). The critique of literal synchrony (p. 95) is clever, highlighting tensions between fidelity and naturalness, but again audience reception data are missed (e.g. how Polish viewers react to mismatched proper nouns). The note on Netflix’s (2019) style guide recommendations, such as allowing the original dialogue to be heard for 1–3 seconds before the VO begins, adds practical relevance to the theoretical discussion (p. 93).

4

In section “Translation Strategies”, the emphasis on faithfulness in non-fiction VO (p. 99) contrasts sharply with fiction’s condensation tactics (p. 100), but the chapter lacks concrete examples of how translators navigate this. A deeper dive into this issue would elevate this section. The emphasis on “condensation,” omission of hesitations, and standardization of language is well-supported by references to empirical studies (pp. 99–101). All in all, the chapter may have benefited from a more critical reflection on the ideological consequences of such strategies, particularly in non-fiction VO, where the imperative of faithfulness may conflict with the need for cultural adaptation or inclusivity. The coverage of synthetic voices, MT and text-to-speech tools (p. 102) is forward-thinking but uncritical, underplaying the limitations of current technologies in capturing the subtleties of spoken discourse and cultural nuance. While synthetic voices may suffice for corporate training videos, how do they handle emotion-laden documentaries or multilingual interviews?

The fifth chapter opens with a highly informative and practical introduction to audio description (AD), a rapidly evolving field and a media access service that renders visual content

into verbal descriptions for blind audiences and visually-impaired viewers. The chapter shines in its breadth of coverage, offering a well-rounded introduction appropriate for both students and professionals. However, the reader may notice that depth varies across sections, with some areas (e.g., historical development) receiving more attention than others (e.g., AD for non-film contexts). While film is a primary AD domain, the chapter’s limited coverage of non-film contexts (e.g., video games, museums) feels like a missed opportunity. For instance, live AD for sports or interactive AD in gaming warrants deeper exploration. Integrating disability theory more explicitly to frame AD as a social, not just technical, practice would enrich this topic. Although the chapter’s exploration of technological developments, especially AI-driven AD and text-to-speech systems, is timely and forward-looking, I would have liked to see AI’s role in AD examined, weighing automation against human nuance. Overall, while the chapter could benefit from a deeper engagement with emerging technologies, it remains an essential read for anyone interested in accessibility and audiovisual translation.

Chapter 6 “Subtitling for the deaf and hard of hearing’ (SDH) provides an in-depth exploration of SDH, a crucial accessibility service that ensures audiovisual media is inclusive for hearing-impaired viewers. The chapter is rich in practical detail, providing actionable insights into “speaker identification” and sound descriptions like door creaks (p. 169). It clarifies the often-confused terms SDH and closed captions. The authors’ decision to treat SDH and CC as essentially the same service—albeit under different names depending on the geographical and institutional context—reflects a pragmatic and inclusive approach. One of the chapter’s strengths lies in its detailed discussion of the target audience. However, it could have benefited from a more in-depth discussion of user reception studies, particularly regarding preferences for different SDH features (e.g., speaker tags, sound descriptions, font styles, and color coding).

The chapter primarily focuses on “intralingual SDH” (same-language transcription), and somehow sidelines “interlingual SDH” (translated subtitles for deaf and hard-of-hearing (DHH) audiences). Given the rise of global streaming platforms, more attention to multilingual accessibility (e.g., Netflix’s SDH in non-English languages) would be beneficial. The chapter also provides a useful overview of the technical and stylistic components of SDH, including timing, reading speed, and formatting conventions. Nevertheless, a more critical examination of current industry standards—especially the lack of consistency across platforms—would have added depth to the analysis. While the chapter mentions AI-driven captioning, the analysis may have benefited from examining limitations, such as ethical concerns about replacing human subtitlers, accuracy issues with live SDH (e.g. misheard dialogue), and the need for hybrid models (i.e., AI + human editing). The chapter also briefly notes that “some deaf viewers may prefer sign language interpreting” (p. 171) over SDH. At this point, I would like to see some insights on hybrid accessibility models (e.g., SDH + sign language options). All in all, the chapter’s strengths lie in terminological clarity, audience sensitivity, and real-world examples.

In chapter 7, “Live Subtitling,” the authors have explored creation methods like “stenography” and “stenocaptioning” (Robson, 2004), “respeaking” (Romero-Fresco & Eugeni, 2020), and automatic speech recognition (ASR), as well as addressing challenges and methods of assessing the accuracy of speech recognition systems like the NER (“Number of words”, “Editing errors”, “Recognition errors”) model (Romero-Fresco & Martinez, 2015). The discussion of quality metrics, particularly the NER model, is detailed. For further editions, a more critical examination of how such metrics are operationalized in different national contexts and how they might privilege certain approaches over others would be most welcome, as well as a more critical stance, because the issue of automation in translation raises not

only technical but also ethical and quality assurance concerns—particularly when applied to live settings where timing and accuracy are nonnegotiable, as the authors noted when referring to the current limitations in ASR and machine translation and claimed that their improvement is inevitable. For further editions, also, user-centered ethnographic findings—particularly from d/Deaf communities—would provide the opportunity to question whether current “best practices” (p. 208) truly align with user needs beyond accuracy and latency and to delve deeper on multilingual live subtitling, challenges and viewer preferences, as well as a critical assessment of AI’s ethical implications, including bias and labor impacts.

Overall, the book is a very valuable resource for students and scholars, offering a clear, well-organized introduction to AVT. With a slightly more critical lens and broader linguistic scope, it could serve as an even more definitive reference in the years to come. Additionally, a more global perspective and deeper interrogation of industry practices would have made it even stronger. However, it is undoubtedly a timely and largely successful contribution to the field. Its effort in balancing theoretical foundations with practical considerations has made it a valuable resource for students, practitioners and emerging professionals.

## References

- Chandler, D., & Munday, R. (2016). *Suspension of disbelief*. In *A dictionary of media and communication* (2nd ed.). Oxford University Press.
- Gambier, Y. (2003). Introduction. Screen transadaptation: Perception and reception. *The Translator*, 9(2), 171–189. <https://doi.org/10.1080/13556509.2003.10799152>
- Greco, G. M., & Jankowska, A. (2020). Media accessibility within and beyond audiovisual translation. In L. Bogucki & M. Deckert. (Eds.), *The Palgrave handbook of audiovisual translation and media accessibility* (pp. 57–81). Palgrave Macmillan.
- Jakobson, R. (2012). On linguistic aspects of translation. In L. Venuti (Ed.), *Translation studies reader* (3rd Ed.) (pp. 126–131). Routledge (Original work published 1959).

- Netflix. (2019). Dubbed audio style guide: vo-style dubbing. Netflix. <https://partnerhelp.netflixstudios.com/hc/en-us/articles/115016062708-Dubbed-Audio-Style-Guide-VO-Style-Dubbing>
- Pedersen, J. (2011). *Subtitling norms for television: An exploration focusing on extralinguistic cultural references*. John Benjamins.
- Pérez-González, L. (Ed.). (2018). *The Routledge handbook of audiovisual translation*. Routledge. <https://doi.org/10.4324/9781315717166>
- Robson, G. D. (2004). *The closed captioning handbook*. Focal Press.
- Romero- Fresco, P., & Eugeni, C. (2020). Live subtitling through respeaking. In Ł. Bogucki & M. Deckert (Eds.), *The Palgrave handbook of audiovisual translation and media accessibility* (pp. 269–295). Palgrave.
- Romero- Fresco, P., & Martínez, J. (2015). Accuracy rate in live subtitling: The NER model. In J. Díaz Cintas & R. Baños-Piñero (Eds.), *Audiovisual translation in a global context: Mapping an ever-changing landscape* (pp. 28–50). Palgrave.

**How to cite this review:** Afrouz, M. (2026). Book review: *Introducing audiovisual translation*. *Íkala, Revista de Lenguaje y Cultura*, 31(1), e361595. <https://doi.org/10.17533/udea.ikala.361595>