

Małgorzata Tryuk 
Uniwersytet Warszawski
m.tryuk@uw.edu.pl

**Recenzja książki Agnieszki Szarkowskiej
i Anny Jankowskiej
pt. *Introducing Audiovisual Translation*,
Abingdon and New York: Routledge, 2024, 218 s.**

Tłumaczenie audiowizualne (znane pod angielskim akronimem AVT, tj. *Audiovisual Translation*) służące przekazywaniu treści audiowizualnych ponad barierami językowymi i kulturowymi jest niezwykle prężnie rozwijającą się odmianą przekładu, która stanowi także autonomiczny przedmiot badań translatorycznych, noszących nazwę *Audiovisual Translation Studies* [Tomaszkiewicz, 1993: 2006].

O zainteresowaniu naukowców tłumaczeniem audiowizualnym świadczą podejmowane projekty naukowe (np. zakończone już europejskie projekty, takie jak: ADLAB PRO, DTV4ALL, SURE, czy też prowadzone w ostatnich latach projekty: ILSA, SMART, MATRIC), artykuły, monografie, także czasopisma, np. *JAT Journal of Audiovisual Translation* – pierwsze międzynarodowe czasopismo poświęcone w całości badaniom nad tłumaczeniem audiowizualnym, którego redaktorką naczelną jest Anna Jankowska, współautorka recenzowanej książki. Na ten temat publikowane są podręczniki akademickie, organizowane konferencje poświęcone różnym aspektom AVT (np. seria konferencji *Media4ALL*)

czy w końcu niezwykle energicznie rozwijana jest dydaktyka przekładu audiowizualnego (np. projekty *ClipFlair*, *Tradilex*). Ponadto zawiązują się organizacje naukowe i zawodowe skupiające ekspertów tego typu przekładu, np. ESIST, czyli European Association for Studies in Screen Translation, w Polsce – STAW, tj. Stowarzyszenie Tłumaczy Audiowizualnych. Badania nad specyfiką tłumaczenia audiowizualnego podejmowane są w ramach różnych paradygmatów, które ukazują nowe terytoria i problemy związane z AVT, jak np. historia AVT, aspekty psychologiczne, nowe technologie i w końcu dostępność mediów. W ostatnim pięcioleciu ukazały się w samym tylko wydawnictwie Routledge znaczące publikacje: w 2018 r. – *The Routledge Handbook of Audiovisual Translation* (pod redakcją Luisa Péreza-Gonzáleza) oraz w 2022 r. – *The Routledge Handbook of Audio Description* (pod redakcją Christophera Taylora i Elisy Perego). Na 2025 r. zapowiedziana jest publikacja Patricka Zabalbeascoa pt. *Audiovisual Translation* (w serii Routledge Introductions to Translation and Interpreting). Wydawnictwo Routledge zapoczątkowało także osobną serię monografii Routledge Research in Audiovisual Translation, w której ukazują się prace na temat nowatorskich i oryginalnych kierunków badań w ramach AVT, takich jak: wykorzystanie tłumaczenia audiowizualnego w dydaktyce języków obcych, tłumaczenie w postaci nadpisów w teatrze i podczas performansów artystycznych, tłumaczenie gier komputerowych, adaptacja filmowa i tłumaczenie klasyków literatury brytyjskiej. Na rok 2025 zapowiedziana jest publikacja pracy poświęconej zjawisku *danmu*, tj. uczestnictwa na żywo w wydarzeniach tłumaczonych na platformach streamingowych (Lu, Chen, 2024). Publikacje te ukazują obecny stan badań oraz wpływ nowych zjawisk na kształtowanie przyszłości AVT i komunikacji za pomocą technologii. Badania nad tłumaczeniem audiowizualnym wpisują się także w rozleglejszy obszar zainteresowań naukowych, które dotyczą dostępności do mediów (ang. *Media Accessibility Studies*) bądź badań nad szeroko pojmowaną dostępnością (ang. *Accessibility Studies*), a które podejmują kwestie walki z wykluczeniem dotyczącym możliwości korzystania z mediów przez liczną rzeszę osób niepełnosprawnych z powodu problemów ze wzrokiem lub słuchem. To właśnie tłumaczenie audiowizualne proponuje konkretne rozwiązania pozwalające na przezwyciężenie tych ograniczeń.

Tak szerokie zainteresowanie badaczy tłumaczeniem audiowizualnym musiało doprowadzić do ukonstytuowania się odrębnej, autonomicznej subdziedziny przekładoznawstwa, komplementarnej wobec badań nad

przekładem pisemnym i badań nad przekładem ustnym. *Introducing Audiovisual Translation* autorstwa Agnieszki Szarkowskiej i Anny Jankowskiej stanowi nie tylko kompleksowy wgląd we wspomniane aspekty tłumaczenia audiowizualnego, lecz jednocześnie uzupełnia aktualny obraz przekładoznawstwa przekraczający podział na tłumaczenie pisemne i tłumaczenie ustne. Recenzowana praca stoi dziś obok klasycznych, wielokrotnie wznawianych monografii Jeremy'ego Mundaya pt. *Introducing Translation Studies* (pięć wydań od 2001 r.) oraz *Introducing Interpreting Studies* Franza Pöchhacker'a (trzy wydania od 2004 r.) opublikowanych w wydawnictwie Routledge. Autorki monografii, Agnieszka Szarkowska z Uniwersytetu Warszawskiego i Anna Jankowska z Uniwersytetu w Antwerpii i Uniwersytetu Jagiellońskiego w Krakowie, są uznanymi w świecie badaczkami tłumaczenia audiowizualnego oraz problematyki dostępności mediów. Obok zainicjowanych przez autorki licznych badań naukowych i publikacji na podkreślenie zasługuje ich szeroka i międzynarodowa działalność dydaktyczna i popularyzatorska w dziedzinie tłumaczenia audiowizualnego i dostępności do mediów.

Recenzowana praca stanowi przegląd stanu badań nad tłumaczeniem audiowizualnym, poczynwszy od pierwszych refleksji i prób opisu AVT do współczesnych projektów związanych z zastosowaniem tłumaczenia rozszerzonego (ang. *augmented translation*), tłumaczenia maszynowego i sztucznej inteligencji. Książka adresowana jest do szerokiego grona odbiorców: badaczy, studentów kierunków tłumaczeniowych, także do odbiorców treści audiowizualnych zainteresowanych teoretycznymi i praktycznymi aspektami oraz odmianami AVT. Z tego też powodu każdy rozdział zakończony jest serią pytań pogłębiających omawianą problematykę, nad którymi czytelnik może się zastanowić, oraz listą lektur uzupełniających poszerzających teoretyczną wiedzę w zakresie AVT. Celem książki jest wypełnienie luki między teoretyczną refleksją nad tym, czym jest tłumaczenie audiowizualne, a praktycznymi rozwiązaniami, poczynwszy od klasycznych typów (tj. tłumaczenie w postaci dubbingu, napisów i wersji lektorskiej) do najnowszych jego odmian (np. audiodeskrypcja i respeaking). Główną zaletą książki jest ukazanie rosnącej roli technologii we wszystkich odmianach tłumaczenia audiowizualnego.

Introducing Audiovisual Translation składa się z 7 rozdziałów, w których omówione jest 6 typów AVT. Pracę rozpoczyna rozdział 1, który przedstawia rozwój AVT i jego kluczowe momenty, poczynwszy od ery kina niemego, ukazuje kolejne etapy tłumaczenia towarzyszące rozwojowi

sztuki filmowej. Omówione są narodowe preferencje oraz wpływ technologii na nawyki oglądania i odbierania treści w innym języku niż w języku oryginalnej ścieżki dźwiękowej. W tym wprowadzającym rozdziale Autorki dokonują klasyfikacji różnych modalności AVT w oparciu o typ tłumaczenia (tłumaczenie wewnątrz- i międzyjęzykowe) i charakter tłumaczonych treści (treści audio-werbalne, audio-niewerbalne, wizualno-werbalne i wizualno-niewerbalne). Choć według rozmaitych podejść badawczych istnieją także inne odmiany tłumaczenia audiowizualnego (s. 8-9), Autorki decydują się na analizę w kolejnych rozdziałach następujących sześciu odmian AVT. Są to: napisy międzyjęzykowe, dubbing, tłumaczenie w wersji lektorskiej, napisy dla niesłyszących, audiodeskrypcja i tworzenie napisów na żywo (*respeaking*). Trzy pierwsze odmiany AVT to najczęściej stosowane techniki tłumaczenia, a więc najbardziej popularne wśród odbiorców. Trzy kolejne są bardziej skoncentrowane na zapewnieniu dostępności do treści audiowizualnych dla odbiorców z niepełnosprawnością związaną ze wzrokiem lub słuchem. W podsumowaniu Autorki stwierdzają, iż biorąc pod uwagę dynamiczne zmiany technologiczne, w tym wykorzystanie tłumaczenia maszynowego i sztucznej inteligencji, przyszłość tłumaczenia audiowizualnego jest bardzo obiecująca. I tak np. rosnąca popularność platform streamingowych oferuje rozmaite opcje dostępu do treści audiowizualnych i umożliwia odbiorcom aktywny wpływ na jakość tłumaczenia.

Rozdział 2 zatytułowany „Napisy międzyjęzykowe” jest przeglądem najważniejszych problemów związanych z tworzeniem napisów międzyjęzykowych, a mianowicie dotyczy takich zagadnień, jak: synchronizacja obrazu i napisów, ograniczenia czasowe i przestrzenne (czas wyświetlania napisów i ich wymiar), operacje językowe dokonywane na tekście oryginalnych dialogów (opuszczenia, kondensacja oraz zwięzłość), specyfika tłumaczenia treści werbalnych innych niż dialog (teksty piosenek, występowanie języka trzeciego) oraz techniki tłumaczenia elementów kulturowych. Kwestie techniczne tworzenia napisów i posiłkowanie się szablonami z kodami czasowymi (ang. *template*), które ułatwiają tłumaczenie, w tym jednoczesne tłumaczenie na kilka języków, są bardzo szczegółowo opisane w rozdziale. Doskonale dobrane wielojęzyczne przykłady ilustrują wszystkie sygnalizowane problemy. Rozdział kończą refleksje na temat przyszłości tej techniki, a w szczególności automatyzacja i wykorzystanie narzędzi do rozpoznania mowy w tworzeniu napisów, co, jak twierdzą Autorki, wciąż wymagać będzie ludzkiej kontroli i post-edycji.

Rozdział 3 dotyczy dubbingu. Autorki charakteryzują jego specyfikę, standardy, cechy wyróżniające, strategie tłumaczeniowe oraz kryteria jakości tłumaczenia. Istotne miejsce w rozdziale zajmuje kwestia najważniejsza w tej technice, tj. synchronizacja: ruchu warg, kinetyki (gesty i mimika ekranowych bohaterów), charakteru postaci (płeć, wiek), treści i czasu trwania wypowiedzi. Synchronizacja jest podstawowym kryterium oceny jakości dubbingu w podobnym stopniu co naturalność, wierność i oczywiście techniczna dokładność. W tym rozdziale Autorki sygnalizują istnienie nieprofesjonalnego tłumaczenia i zjawisko *fandubbingu*, a także, co jest interesujące dla polskiego czytelnika, cytują stanowisko znanego polskiego tłumacza audiowizualnego Bartosza Wierzbieży, dla którego najważniejsze elementy w dubbingu to: synchronizacja, atrakcyjność formy i zrozumienie treści przez odbiorcę (s. 72). Osobne miejsce w rozdziale zajmuje kwestia dubbingu w wypadku obcych akcentów, dialektów, tłumaczenia na język angielski i w końcu jakość tłumaczenia. Przedstawione rozwiązania ilustrowane są licznymi wielojęzycznymi przykładami. I tak jak w poprzednim rozdziale Autorki podkreślają, iż przyszłość dubbingu będzie kształtowana przez technologię, w tym przez możliwości oferowane przez sztuczną inteligencję, np. automatyczny dubbing, wykorzystujący automatyczne rozpoznanie i syntezę mowy, a przede wszystkim przez technologię pozwalającą na manipulowanie ruchami warg w celu dopasowania ich do przetłumaczonego tekstu.

Tłumaczenie w wersji lektorskiej (ang. *voice over*), które ma miejsce, gdy lektor lub aktor odczytuje listę dialogową w języku przekładu, podczas gdy głos oryginalny ścieżki dźwiękowej jest ściśniony, zostało przedstawione w rozdziale 4 książki. Autorki dokonują przeglądu rozmaitych narodowych tradycji lektorskich i dominujących typów treści audiowizualnych wykorzystywanych w tej odmianie AVT, np. w zależności od gatunku filmowego (film fabularny i film dokumentalny). Opisują stosowane strategie tłumaczeniowe i tak jak w poprzednich rozdziałach zwracają uwagę, że w przyszłości ta odmiana AVT znajdować się będzie pod silnym wpływem nowych technologii (np. automatyzacja czy wykorzystanie syntezy mowy). Jednak, jak piszą, w miarę rozwoju nowych technologii ważne będzie, by tłumacze zapoznawali się na bieżąco z najnowszymi trendami i technikami w celu zapewnienia wysokiej jakości tłumaczeń lektorskich, które mają spełniać różne potrzeby odbiorców.

Trzy ostatnie rozdziały dotyczą odmian AVT związanych z problematyką dostępności mediów dla odbiorców z deficytami słuchu lub wzroku.

W tym wypadku tłumaczenie, tak audiodeskrypcja, jak i tworzenie napisów, może być wewnątrzjęzykowe (w języku źródłowym) lub międzyjęzykowe (w języku innym niż źródłowy).

Rozdział 5 natomiast został poświęcony audiodeskrypcji (AD), która stanowi odmianę przekładu intersemiotycznego polegającą na werbalnym opisie tego, co widać na ekranie. Narracja jest wstawiana w przerwy w dialogach i przeznaczona głównie dla widzów niewidomych lub niedowidzących. Audiodeskrypcja to bardzo skomplikowane działanie translatorskie, gdyż wymaga opracowania specyficznych strategii i scenariuszy niezbędnych do syntetycznego opisu tego, co jest pokazywane na ekranie: obrazów (a więc opisu, co się dzieje, gdzie, jak, kiedy, z czym udziałem i w końcu z jakimi emocjami się wiąże), dźwięków (innych niż werbalne i trudne do rozpoznania), elementów werbalnych napisanych oraz elementów kulturowych. Autorki przedstawiają historię audiodeskrypcji, omawiają przyjęte narodowe wytyczne, strategie tłumaczenia oraz szablony umożliwiające wielojęzyczną audiodeskrypcję. Do najtrudniejszych zadań dla twórców audiodeskrypcji należy zachowanie równowagi między obiektywnym opisem a subiektywnością audiodeskrybera. W podsumowaniu tego rozdziału autorki zwracają uwagę na rosnący ruch na rzecz dostępności i liczne narodowe inicjatywy legislacyjne, które zwiększają zapotrzebowanie na tę odmianę (lub usługę) tłumaczeniową. Już obecnie projekty automatyzacji audiodeskrypcji przynoszą wymierne i praktyczne korzyści. Jednak, jak podnoszą Autorki, opisywanie postaci i obiektów oraz przedstawianie fabuły za pomocą spójnej narracji w dużej mierze pozostaje poza obecnymi możliwościami tłumaczenia maszynowego.

Tematem kolejnego rozdziału 6 jest tworzenie napisów dla osób z dysfunkcją słuchu (ang. *Subtitling for Deaf and Hard of Hearing, SDH*), które obejmuje transkrypcję lub tłumaczenie dialogów mówionych, dźwięku i identyfikatorów mówcy. W rozdziale tym ukazana jest specyfika tworzenia tej odmiany AVT dla konkretnej grupy odbiorców, którzy mają inne potrzeby niż słyszący odbiorcy napisów. Tak np. w SDH napisy powinny zawierać informacje dotyczące identyfikacji mówcy/mówców, dźwięków diegetycznych i niediegetycznych, charakteru muzyki, akcentów, obcej wymowy lub innych języków. Autorki zwracają też uwagę na specyfikę napisów tworzonych z myślą o dzieciach. Postęp technologiczny pozwala na większą personalizację SDH, np. pod względem czcionek lub kolorów używanych w napisach. Jeśli chodzi o przyszłość tej odmiany tłumaczenia audiowizualnego Autorki są zdania, że podobnie jak w przypadku innych

rodzajów AVT SDH najprawdopodobniej będzie coraz bardziej automatyzowane, przede wszystkim dzięki potencjałowi automatycznego rozpoznawania mowy. Jednak, jak słusznie zwracają uwagę, obecnie automatycznie tworzone napisy są wciąż zbyt dosłowne, co niekorzystnie wpływa na rozumienie tłumaczonej treści.

Zamykający książkę rozdział 7 został poświęcony najnowszej odmianie AVT, tj. tworzeniu napisów na żywo – *respeakingowi*. Jest to usługa także przeznaczona dla widzów niesłyszących lub niedosłyszących podczas projekcji filmowych, w przekazie telewizyjnym, podczas imprez otwartych (np. mitingi lub konferencje). W *respeakingu* tłumacz słucha tekstu źródłowego, powtarza go dokonując niezbędnej edycji (np. dodając informacje o interpunkcji). Dzięki systemowi rozpoznania mowy jest ona przetwarzana na wyświetlany na ekranie tekst. Napisy na żywo można tworzyć na różne sposoby – od szybkiego pisania na standardowej klawiaturze QWERTY po w pełni zautomatyzowane rozwiązania oparte na sztucznej inteligencji. W *respeakingu* wewnątrzjęzykowym tekst jest tworzony w języku źródłowym przez jedną osobę, w *respeakingu* międzyjęzykowym zazwyczaj pracują dwie osoby: tłumacz ustny, który symultanicznie tłumaczy tekst, i *respeaker*, który dokonuje przekształcenia mowy na tekst pisany. W tym rozdziale Autorki poświęcają szczególne miejsce metodologii oceny jakości *respeakingu*, w tym modelom NER i NTR opracowanym przez Pabla Romera Fresca (2015, 2017) oraz dobrym praktykom w redagowaniu napisów. Także w tym wypadku przyszłość *respeakingu* leży w automatyzacji i wykorzystaniu rozpoznawania mowy oraz sztucznej inteligencji z korzyścią dla tych, którzy potrzebują dostępu do treści audiowizualnych podczas wydarzeń na żywo. Popularność napisów powstających na żywo w połączeniu z tłumaczeniem maszynowym wciąż rośnie, podobnie jak liczba potencjalnych odbiorców, którzy mogą skorzystać z tej usługi (np. podczas nauki języków obcych). Jednak kwestią zasadniczą pozostaje wciąż jakość napisów tworzonych w czasie rzeczywistym.

Recenzowana książka stanowi ważny wkład do badań nad przekładem audiowizualnym. Jest solidnym i wyczerpującym przeglądem literatury na temat rozwoju i przyszłości tej odmiany tłumaczenia. Istotną część pracy stanowi przedstawienie możliwości rozwoju AVT związanego z nowymi technologiami. Omawiane problemy są ilustrowane wielojęzycznymi przykładami, a podstawowe pojęcia – opatrzone jasnymi definicjami. Każdy rozdział kończy się omówieniem możliwości, jakie oferują systemy automatycznego rozpoznania i syntezy mowy, tłumaczenia maszynowego

i sztucznej inteligencji. Wartością dodaną są proponowane na zakończenie każdego rozdziału pytania sygnalizujące nowe tematy i zagadnienia zachęcające czytelników do znalezienia własnych odpowiedzi i opinii. Każdy rozdział opatrzony jest dodatkowo odpowiednią i aktualną bibliografią.

Do zadań recenzenta należy też zasygnalizowanie pewnych braków, które jednak nie umniejszają wartości i znaczenia omawianej książki. Należy do nich m.in. niedostateczne opisanie zjawiska nieprofesjonalnego AVT, tj. fansubbingu i fandubbingu, czyli fanowskiego tłumaczenia ścieżki werbalnej filmów, seriali, gier komputerowych, które jest możliwe dzięki nieograniczonemu dostępowi do specjalistycznych oprogramowań. Tłumaczenie fanowskie jest obecnie szeroko komentowane zarówno z punktu widzenia translatorycznego, gdyż ukazuje nową praktykę tłumaczeniową i jej liczne innowacje, jak i prawnego, gdyż ma wpływ na stosowanie lub nie prawa autorskiego do tłumaczonych treści (Tryuk, 2018).

Historia tłumaczenia audiowizualnego ukazuje także, że operacje tłumaczenia treści audiowizualnych z łatwością mogą podlegać manipulacji lub cenzurze, niezależnie od tego, czy będzie to cenzura instytucjonalna czy autocenzura tłumacza (Merkle, Baer, 2024). Liczne przykłady cenzurowania (a nawet całkowitej zmiany) treści audiowizualnych w państwach autorytarnych i nie tylko ukazują nieograniczone możliwości i inwencję cenzorów w tym zakresie. Tłumaczenie audiowizualne, jak żadne inne, z łatwością poddaje się operacjom manipulacji treści. Może warto wspomnieć ten aspekt AVT.

Na koniec należy zauważyć, że tłumaczenie audiowizualne w postaci napisów wewnątrz- lub międzyjęzykowych może być z powodzeniem stosowane przy tworzeniu treści dostępnych w języku prostym (ang. *easy language*) dla odbiorców z deficytami kognitywnymi, a także napisów w językach mniejszościowych lub mniej rozpowszechnionych (ang. *LLD – languages of lesser diffusion*). Perspektywa wykorzystania tłumaczenia maszynowego i sztucznej inteligencji w tworzeniu napisów w języku prostym lub w języku mniejszościowym jest także ciekawa z punktu widzenia rozwoju społeczeństwa inkluzyjnego.

Jak sygnalizuje tytuł książki, *Introducing Audiovisual Translation* stanowi doskonałe wprowadzenie do problematyki tłumaczenia audiowizualnego, tej trzeciej odmiany przekładu, różniącej się od przekładu pisemnego i ustnego odrębnym przedmiotem badań, metodologią i przede wszystkim wykorzystaniem potencjału technologii zarówno w pracach teoretycznych, jak i w praktyce AVT. Ponadto wpisanie AVT w obszar

badan nad dostępnością do mediów ukazuje jego społeczne znaczenie, co także zostało mocno podkreślone w książce. W podsumowaniu należy zaznaczyć, iż jest ona nie tylko aktualnym przewodnikiem po problematyce AVT, ale także potrzebnym podręcznikiem, który może i powinien być wykorzystany podczas kształcenia tłumaczy audiowizualnych na poziomie akademickim.

Bibliografia

- Lu, S., Chen, X. (2024), „An Experimental Study on Viewing Perception and Gratification of Danmu Subtitled Online Video Streaming”, *Journal of Specialised Translation*, No. 42, <https://doi.org/10.26034/cm.jostrans.2024.5990>.
- Merkle, D., Baer, B.J. (eds) (2024), *The Routledge Handbook of Translation and Censorship*, Routledge, Abingdon.
- Munday, J. (2022), *Introducing Translation Studies*, Routledge, Abingdon.
- Pérez-González, L. (ed.) (2018), *The Routledge Handbook of Audiovisual Translation*, Routledge, Abingdon.
- Pöchhacker, F., 2022, *Introducing Interpreting Studies*, Routledge, Abingdon.
- Romero Fresco, P., Martinez, J. (2015), „Accuracy Rate in Live Subtitling: the NER Model”, w: Diaz Cintas, J., Baños Piñero, R. (eds), *Audiovisual Translation in a Global Context: Mapping an Eever-changing Landscape*, Houndmills–Basingstoke–Hampshire–New York, 28-50.
- Romero Fresco, P., Pöchhacker, F. (2017), „Quality Assessment in Interlingual Live Subtitling: the NTR Model”, *Linguistica Antverpiensia. New Series: Themes in Translation Studies*, 16, s. 149-167.
- Taylor, Ch., Perego, E. (eds) (2022), *The Routledge Handbook of Audio Description*, Routledge, Abingdon.
- Tomaszkiewicz, T. (1993), *Les opérations linguistiques qui sous-tendent le processus de sous-titrage des films*, Wydawnictwo Naukowe UAM, Poznań.
- Tomaszkiewicz, T. (2006), *Przekład audiowizualny*, Wydawnictwo Naukowe PWN, Warszawa.
- Tryuk, M. (2018), „Le fansub et le fansuber dans la réalité polonaise”, w: Monti, E., Schnyder, P. (réd.), *Traduire à plusieurs/Collaborative translations*, Orizons, Paris, 413-427.