

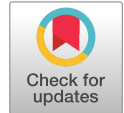
İstanbul Üniversitesi Çeviribilim Dergisi Istanbul University Journal of Translation Studies

Research Article | Araştırma makalesi

Open Access | Açık Erişim

Global Voices, Local Losses: The Quality Assessment of AI-generated Dubbing on the YouTube channel ‘Vlad’

Küresel Sesler, Yerel Kayıplar: YouTube Kanalı “Vlad”da Yapay Zekâ Tarafından Üretilen Dublajın Kalite Değerlendirmesi



Ebru Çavuşoğlu¹ & Erdem Celal Gören²

¹ Samsun University, Department of English Translation and Interpreting, Samsun, Türkiye

² Samsun University, Department of English Translation and Interpreting, Samsun, Türkiye

Abstract

AI-generated dubbing on YouTube has enabled fast, scalable multilingual content distribution, yet raises serious concerns about cultural manipulation, ideological reframing, and narrative control. With an emphasis on changes in tone, pragmatics, comedy, and cultural references, this study explores how AI-generated dubbing on YouTube shapes meaning. AI-generated dubbing often prioritizes just lip-sync accuracy and serving multiple language options to viewers over cultural integrity, leading to errors in meaning while supporting commercial goals, such as reaching more followers and encouraging them to subscribe to the channel. In line with this fact, the current study aims to examine how AI-generated dubbing affects cultural representation by conducting a qualitative review of data derived from purposive sampling of Turkish dubbing translations from the children’s content channel ‘Vlad’. The data derived from the study is classified according to recurring patterns under Larson’s Meaning-based Translation Quality Assessment (TQA) model. Meaning, purposefully selected samples are classified in the frame of five error types: (1) semantic inaccuracies, (2) pragmatic mismatches, (3) omissions and additions, (4) audiovisual misalignment, and (5) speech synthesis-related problems, which align with Larson’s criteria for evaluating translation quality, including accuracy, naturalness, and clarity. This study contributes to the growing conversations in audiovisual translation and digital media ethics by situating AI dubbing within broader discourses on cultural manipulation and emphasizing the need for critical frameworks and human intervention in AI-mediated cultural transfer processes.

Öz

YouTube’da yapay zekâ tarafından üretilen dublaj, hızlı ve ölçeklenebilir çok dilli içerik dağıtımını mümkün kılmıştır; ancak kültürel manipülasyon, ideolojik yeniden çerçeveleme ve anlatı kontrolü bağlamında ciddi endişeler doğurmaktadır. Bu çalışma, üslup, pragmatik, mizah ve kültürel referanslardaki değişikliklere odaklanarak, YouTube’da yapay zekâ tarafından üretilen dublajın anlamı nasıl şekillendirdiğini incelemektedir. YZ tarafından otomatik olarak üretilen dublaj, genellikle kültürel bütünlükten ziyade sadece dudak senkronizasyonunun doğruluğuna ve izleyicilere çoklu dil seçenekleri sunmaya öncelik verir. Ancak bu yapay dublaj, daha fazla takipçiye ulaşmak ve onları kanala abone olmaya teşvik etmek gibi ticari hedefleri desteklerken anlam aktarımında hatalara yol açabilmektedir. Bu gerçeğe dayanılarak, mevcut çalışma, çocuk izleyicileri hedef alan “Vlad” isimli kanaldan seçilen Türkçe dublaj çevirilerinin amaçlı örneklemesinden elde edilen verilerin nitel bir incelemesini yaparak, YZ tarafından üretilen dublajın kültürel temsili nasıl etkilediğini ortaya koymayı amaçlamaktadır. Çalışmadan elde edilen veriler, Larson’ın Anlam Temelli Çeviri Kalitesi Değerlendirme (TQA) modeline göre tekrarlanan kalıplara göre sınıflandırılmıştır. Bir başka deyişle, kasıtlı olarak seçilen örnekler beş hata türü çerçevesinde sınıflandırılmıştır: (1) anlamsal hatalar, (2) pragmatik uyumsuzluklar, (3) eksiklikler ve eklemeler, (4) görsel-işitsel uyumsuzluk ve (5) konuşma senteziyle



Citation | Atıf: Çavuşoğlu, E. & Gören, E. C. (2026). Global voices, local losses: The quality assessment of AI-generated dubbing on Youtube. *Istanbul Üniversitesi Çeviribilim Dergisi–Istanbul University Journal of Translation Studies*, (24), 225–246. <https://doi.org/10.26650/iujts.2026.1950097>

This work is licensed under Creative Commons Attribution-NonCommercial 4.0 International License. ©

© 2026. Çavuşoğlu, E. & Gören, E. C.

Corresponding author | Sorumlu Yazar: Ebru Çavuşoğlu ebru.cavusoglu@samsun.edu.tr



ilgili sorunlar; bunlar, doğruluk, doğruluk ve netlik dahil olmak üzere Larson'ın çeviri kalitesini değerlendirme kriterleriyle uyumludur. Ayrıca bu çalışma, YZ dublajını kültürel manipülasyon üzerine daha geniş söylemler içinde konumlandırarak ve YZ aracılı kültürel aktarım süreçlerinde eleştirel çerçevelere ve insan müdahalesine duyulan ihtiyacı vurgulayarak, görsel-işitsel çeviri ve dijital medya etiği alanlarında giderek artan tartışmalara katkıda bulunmaktadır.

Keywords

AI-generated dubbing · Audiovisual translation · Dubbing translation · Translation Quality Assessment Model · YouTube.

Anahtar Kelimeler

Çeviri Kalitesi Değerlendirme Modeli · Dublaj çevirisi · Görsel-işitsel çeviri · Yapay zekâ destekli dublaj · YouTube.

Introduction

The advancement of technology, especially the screens with internet connections, has changed the way we spend our time. Although television was labeled an “idiot box” in the 1950s (Merriam-Webster, n.d.), it is now present in the vast majority of households worldwide. It had the power to unite families and spark discussions about a specific program at a particular time, thanks to its schedule. It may have been regarded as a constant pastime activity by adults at the time. However, today, perceptions of home entertainment sources have changed, creating more active viewers who are also content generators. The emergence of the internet set the distinction between viewers. Stewart (1998, p. 5) argues that television is typically consumed passively and at a distance, whereas web-based and computer-mediated environments require greater user engagement and interaction with the content.

YouTube is currently the world's most visited video-sharing platform and the second most visited website overall, with approximately 45 billion monthly visits worldwide (Statista, 2026). It pushed the boundaries set by TV and contributed to the phenomenon of viewers and content creators. According to YouTube (n.d.-a), the platform's mission is to empower individuals to express themselves and connect with others by sharing their stories and building communities. Therefore, all users have an equal chance to create content and share it with the world. Bar et al. (2009, p. 18) identify three distinct user groups on YouTube. The first are passive users who only want to watch YouTube videos and expect seamless, high-quality streaming. The second group consists of active users who are primarily driven by the possibility of having their opinions heard and who participate on YouTube by leaving comments or rating videos. Lastly, the third type consists of content-contributing users who create videos for the website and are mostly driven by the possibility of gaining more visibility for their channels and videos.

Looking back across the growth story of such websites, complexity skyrocketed in the late 1990s in a number of ways, including increased content quantities, dynamic content, visitor numbers, and more sophisticated supporting hardware and software (McKeever, 2003, p. 687). Today, the advancement of multi-functional devices has prompted users of all ages to interact with YouTube through a range of easy-to-use interfaces, including televisions, personal computers, laptops, tablets, and smartphones (Balakrishnan & Griffiths, 2017, p. 364). As parents increasingly used screen-based media as a means of occupying children during meals, before bedtime, or while attending to other tasks, YouTube became a popular source of entertainment for young children (Elias & Sulkin, 2017). However, the rapid global expansion of such content necessitates efficient strategies to reach diverse audiences. To meet this demand, creators increasingly rely on new features introduced by the developers.

For instance, AI-powered accessibility features such as automatic captions, caption translation, translated video metadata, and automatic dubbing enable content to be accessed by both people with disabilities and audiences from different linguistic backgrounds (YouTube, n.d.-b, n.d.-c, n.d.-d, n.d.-e). We can say that YouTube has greatly increased subcultures to the point that there are none. Today, everyone watches and

knows the same thing. Even babies watch the same videos that babies abroad watch. A few years ago, before natural language processing and other related AI models emerged, web content generally aimed at its local audience. If they were successful and had the budget like the TV industry, the content would get adapted with either subtitles or dubs. One example of those channels is Vlad and Niki (n.d.).

Traditionally, channels like Vlad and Niki (n.d.) achieved global reach by manually creating dozens of separate channels and employing teams of translators and voice actors to dub their content into various languages. However, the integration of YouTube's native AI-generated multilingual audio tracks has fundamentally transformed this landscape, allowing creators to automatically localize a single video into multiple languages. This shift from human-led localization to AI-driven automation removes the need for local intermediaries but introduces significant risks to content integrity.

The primary aim of this study is to examine how AI-generated automatic dubbing affects cultural and linguistic representation by conducting a qualitative review of Turkish dubbing translations from the children's YouTube channel "Vlad". Therefore, in accordance with this aim, the study seeks to answer the following research questions:

- 1) How does AI-generated automatic dubbing influence the transfer of semantic and pragmatic meaning in audiovisual content?
- 2) To what extent does the prioritization of technical synchronization lead to the neglect of linguistic accuracy?
- 3) How does AI dubbing vary based on the channel and genre of the content?

To provide a comprehensive analysis of these shifts, the paper is organized into several key sections. It begins with a literature review that situates AI within audiovisual translation and discusses the technical evolution of dubbing. The methodology section then details the purposive sampling of three specific videos and explains the application of Larson's (1998) Meaning-Based Translation Quality Assessment (TQA) model as the primary evaluative framework. This is followed by a detailed analysis of the collected data, classified into five systematic error patterns: semantic inaccuracies, pragmatic mismatches, omissions and additions, audiovisual misalignment, and speech synthesis-related problems. Finally, the study discusses the broader implications of these findings for digital media ethics and global monoculture, concluding with a call for critical human intervention in AI-mediated cultural transfer.

Literature Review

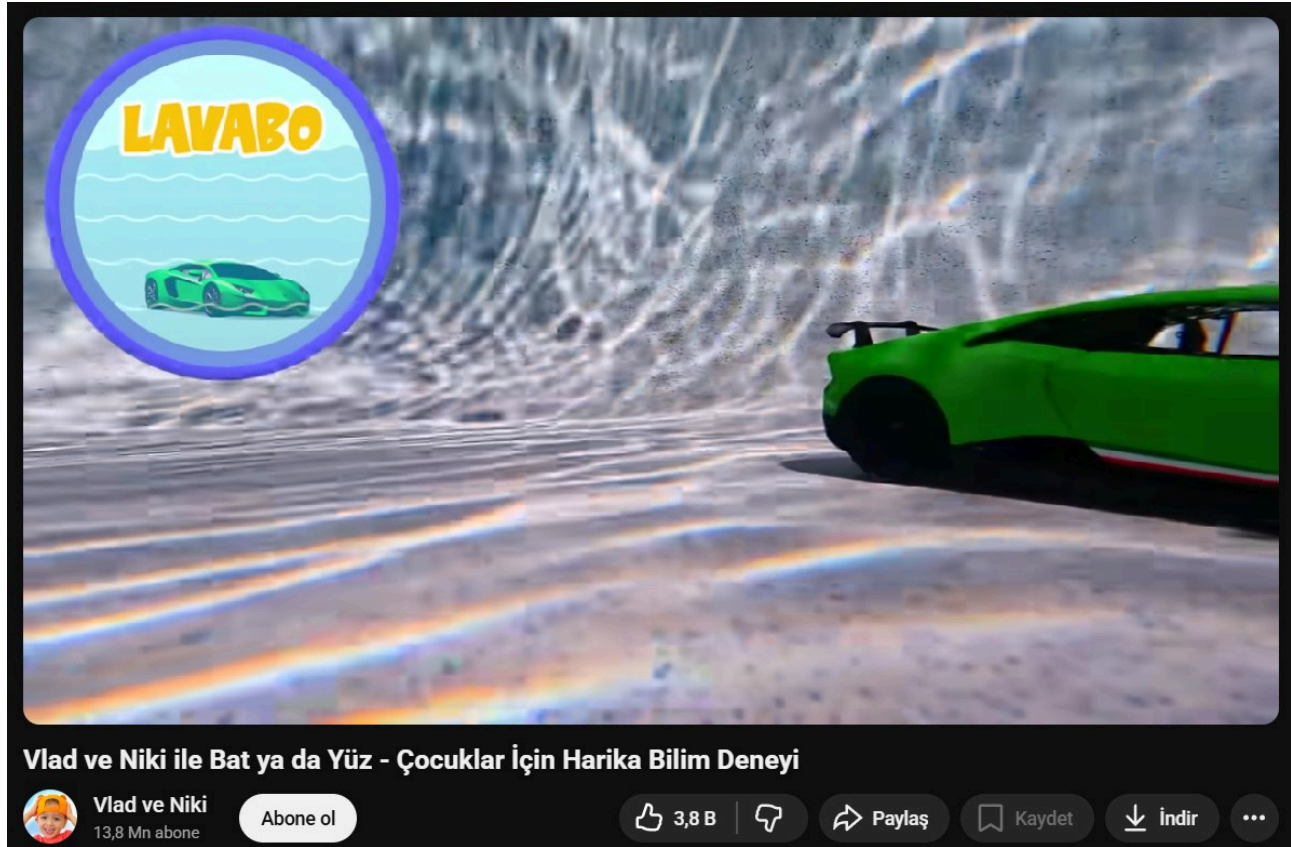
AI in Audiovisual Translation

From today's theoretical perspective, studies about translation emerged around the 1960s. Interestingly, such studies were inspired by the first attempts to make translators unnecessary by creating automatic translation devices (Nord, 2016). Moore's law states that the number of transistors on computer chips doubles approximately every 2 years, which means technology will become cheaper and thus more accessible over time. Another reason hardware and software are becoming public is the opportunity to commercialize them. Machine translation (MT) technology, for example, is a potent tool that everyone can use today for content creation and distribution in the age of artificial intelligence (AI). While there are various types (rule-based, neural, etc.) and degrees (human-assisted vs. machine-assisted translation) of MT use, it all comes down to cutting costs and saving time for those who use it. Especially when used without a quality assessment procedure, it may yield confusing, out-of-context renderings. One example comes from the Turkish account of the famous YouTube channel "Vlad and Niki", where one of the videos uses a homonym. The channel has many videos in different languages, and the video titled "Vlad ve Niki ile Bat ya da Yüz -

Çocuklar İçin Harika Bilim Deneyi" on their Turkish account shows children throwing toys into a pool to see if they swim or sink. It is important to note that most videos targeting children use sound effects, music, and animations to keep children engaged. However, in the video shown in Figure 1 below, when the toys sink, the Turkish word "lavabo" appears on screen, meaning "sink" in the bathroom. Bilingual viewers may recognize the homonymy and realize that it stems from either context-insensitive translation or unsupervised MT.

Figure 1

A translation error in a video on the Vlad ve Niki Turkish channel. From "Vlad ve Niki ile Bat ya da Yüz - Çocuklar İçin Harika Bilim Deneyi," by Vlad ve Niki, 2023 (https://youtu.be/5o_arN_9o3o).



There are many such cases found online, although MT has been thoroughly researched in textual contexts, its use on multimodal texts is still in its infancy (Jin and Yuan, 2023, p. 81). According to studies, MT has encountered difficulties in audiovisual translation (AVT) due to the complexity of audiovisual texts, which combine visual and aural components with linguistic information. For instance, on streaming platforms, you can enable automatic captions and (if provided) MT-generated caption translation on demand, allowing you to follow along in your own language directly on the site where you're watching the video. In some cases, even if the translation isn't top-notch, you can still grasp the general meaning of the content (Şahin, 2022, p. 48). However, Karakanta and her colleagues (2020) explain that dubbing is a form of audiovisual translation in which MT has not yet found a firm footing due to the genre's particularities, which require approximation to orality and the representation of oral dialogue. They address the dichotomy between on-screen dubbing (constrained by synchronization requirements) and off-screen dubbing (unconstrained), suggesting that distinguishing between the two is necessary for successful MT strategies. The findings reveal that a Neural Machine Translation (NMT) system tuned to the dubbing domain achieves significantly worse results for on-screen dubbing than for off-screen dubbing, highlighting synchronization constraints as a challenge for MT quality.

Özkaya Marangoz (2023) provides a brief overview of such prominent technologies in her study. Although companies try to equip their computer-assisted translation (CAT) tools with the dizzying array of technologies described in the study, in practice, a tool optimized for technical localization may fail to capture the creative nuance required for transcreation. Over the years, an exponential increase in computational speed and the rise of coding companies have brought about various types of translation technologies. These tools include speech-to-text, text-to-speech, and even chatbots (famous examples include ChatGPT and Gemini), powered by neural networks that streamline the entire localization workflow. Collected under the label “AI”, the specified tools play different roles in a translation pipeline. According to Suleyman (2025, p. 70), critics argue that buzzwords like AI are poorly defined or intellectually misguided and should not be taken seriously. To add to the confusion, new terms come up every day; AI, AGI, and GenAI, to name but a few.

Establishing a precise definition for AI is surprisingly tricky, largely because we still haven't reached a consensus on what “intelligence” actually is. In broad terms, AI refers to the creation of computer systems designed to execute tasks that usually demand human-level cognition. These systems use a combination of sophisticated algorithms, massive datasets, and high-performance computing. The reason the boundaries of AI remain so fluid is its interdisciplinary foundation. Aluvihara et al. (2025) state that “ambiguity in defining AI stems from its interdisciplinary nature, drawing from computer science, mathematics, philosophy, psychology, and neuroscience, each contributing its own perspective and influencing the scope of what constitutes intelligent behavior.” For instance, in the field of translation and interpreting, machine translation systems were not as effective as they are today. Google itself notes that the shift to neural networks was intended to move beyond word-for-word translation and improve fluency and naturalness (Yao, 2026). Before this update, Google Translate was strongly associated with low-quality translations in popular discourse in Türkiye (Sargın, 2024, p. 14). In contrast, DeepL, on its website (n.d.), has consistently positioned itself as an AI-powered language technology company, describing its services as AI-driven language solutions and emphasizing the role of artificial intelligence in translation and writing assistance. Therefore, although both Google Translate and DeepL rely on neural machine translation technologies, DeepL is more explicitly branded and marketed as an AI tool.

Fast forward to today, these other tools with different functions are fused in a pipeline, and new tools have emerged, such as speech-to-speech automatic translation and interpreting, which is the object of this study. This pipeline is nothing novel, as it only puts existing items in order and imports one of their outputs into the latter. The AI dubbing process includes: automatic speech recognition (ASR) to transcribe the source audio, MT to convert the text into the target language, and text-to-speech synthesis to generate the final localized vocal track. Gallo and Lordache (2026) state that Amazon began testing AI-generated dubbing in 2025, initially rolling it out in English and Latin American Spanish for several series. The limited scope of this initiative, targeting content that previously lacked localization, suggests that Prime Video is cautiously exploring AI dubbing as a potential cost-saving measure.

According to official reports, the primary goal of this expansion was to increase the volume of localized content available to international audiences (Amazon, 2025). Yet the experimental nature of these automated pipelines became evident through the titles they received. Following the release of AI-generated English dubs for series such as *Banana Fish*, Amazon eventually withdrew the audio tracks amid significant viewer backlash over the quality of the synthesis (Tate, 2025). This removal highlights the current limitations of AI dubbing in capturing the nuanced performances audiences expect, suggesting that while the technical pipeline is functional, its output often falls short of the qualitative benchmarks set by traditional human-led localization.

Both in literature across several fields and in the day-to-day operations of the entertainment industry, dubbing video content from one language into another has received considerable attention (Brannon, Virkar, & Thompson, 2023, p. 419). Yet, the analysis of constraints associated with accuracy, voice naturalness, synchronization, and cultural sensitivity in automatic dubbing is lacking in the literature. This gap is particularly significant given that AI-generated dubbing became publicly available only around February 2026, making empirical quality assessments both timely and necessary. To the best of the authors' knowledge, no prior study has systematically categorized errors arising from AI-generated dubbing, which is a gap this study directly addresses by applying Larson's (1998) Meaning-Based TQA model to Turkish dubbing on a major YouTube channel.

Quality in Translation for Dubbing

An examination of the literature on AVT, and specifically dubbing practices, reveals that studies concerning the transfer of cultural elements into the target language hold a significant position. In this context, a recent study by Çavuşoğlu (2023), one of the authors of the present research, examined the strategies employed by the translator to transfer humorous expressions in the Turkish dubbing of the animated film *Shrek*. Utilizing a qualitative descriptive approach, Çavuşoğlu analyzed how humor, wordplay, and irony, which are inherently bound by cultural and linguistic constraints, are adapted into the target language using Pedersen's (2005) translation strategies. Çavuşoğlu's study emphasizes that the dubbing process is not merely a mechanical exercise in lip-syncing or literal word-for-word translation.

As previously discussed, the advent and advancement of technology have made fan translation, machine translation (MT), machine translation post-editing (MTPE), and large language model (LLM) translation increasingly prominent areas of scholarly inquiry. To the best of the authors' knowledge, there is no study about errors stemming from AI-generated dub. This novelty is mostly due to the feature being made public in February 2026. Given the novelty of this feature, identifying and classifying errors has been deemed an appropriate preliminary step before a comprehensive quality assessment of the resulting translations is undertaken.

Translation quality in audiovisual translation, and particularly in dubbing, is a complex issue that is frequently discussed in both academic research and professional practice (Brannon et al., 2023; Chaume, 2020; Spiteri Miggiani, 2022; Spiteri Miggiani, 2023). Traditional dubbing quality standards generally focus on elements such as lip and kinetic synchronization, dialogue credibility, coherence, and fidelity to the source text. Although audiences can quickly notice mismatches in lip synchronization and dialogue length, excessive prioritization of these technical constraints during translation may lead to the neglect of the text's semantic and cultural integrity (Varela, 2007, pp. 74-75). Within the field of Translation Quality Assessment (TQA), various models have been proposed to date (Reiss, 1981; Nida, 1964; House, 2014). However, many of these models contain significant limitations in addressing the complex nature of audiovisual texts.

In contrast, Mildred L. Larson's (1998) Meaning-Based Translation model offers a practical alternative for evaluating the quality and cultural representation of AI-generated dubbing. Larson (1998, p. 3) defines translation fundamentally as a process in which form changes while meaning must remain constant. According to this approach, translation should employ the natural forms of the receptor language, convey the communicative purpose of the source text to the target audience in the same manner, and preserve the dynamics of the original text. Larson's model evaluates translation quality based on three primary criteria (Larson, 1998, p. 49): (1) accuracy, (2) clarity, and (3) naturalness. The criterion of (1) accuracy requires that the message of the source text be conveyed correctly into the target language and that any necessary additions or omissions be made to preserve the intended meaning fully. (2) Clarity refers to the use of forms in the target language that make the source message as easily understandable as possible for readers or

viewers. (3) Naturalness aims for the translation to employ normal grammatical forms appropriate to the sociocultural context of the target language, allowing the text to sound not like a translation, but like a naturally produced text in that language.

Larson (1998, pp. 485-503) asserts that the testing phase is a vital component of the quality assessment, designed to verify that the final product adheres to the core principles of accuracy, clarity, and naturalness. This process involves a collaborative effort between the translator, who does the translation; translation consultants, who provide oversight; and most importantly, the target audience, whose reactions reveal whether the translation is idiomatic. To achieve these goals, Larson outlines several practical testing methods that serve different purposes. One method is back-translation, in which a person unfamiliar with the original text renders the translation back into the source language, allowing the translator to detect shifts. Furthermore, comprehension tests assess the audience’s grasp of the narrative and its nuances through targeted questioning, while naturalness and readability tests evaluate whether the translation flows according to the idiomatic patterns of the target language.

Such systematic testing is essential for ensuring that the content functions as authentic communication within the target environment. Larson’s model is thus deemed particularly effective by the authors for identifying the kinds of errors frequently encountered in AI dubbing, where excessive emphasis on lip synchronization or literal translation often damages cultural transfer. Therefore, Larson’s framework provides the theoretical and critical foundation necessary for categorizing cultural shifts in meaning within AI-generated audiovisual translation processes.

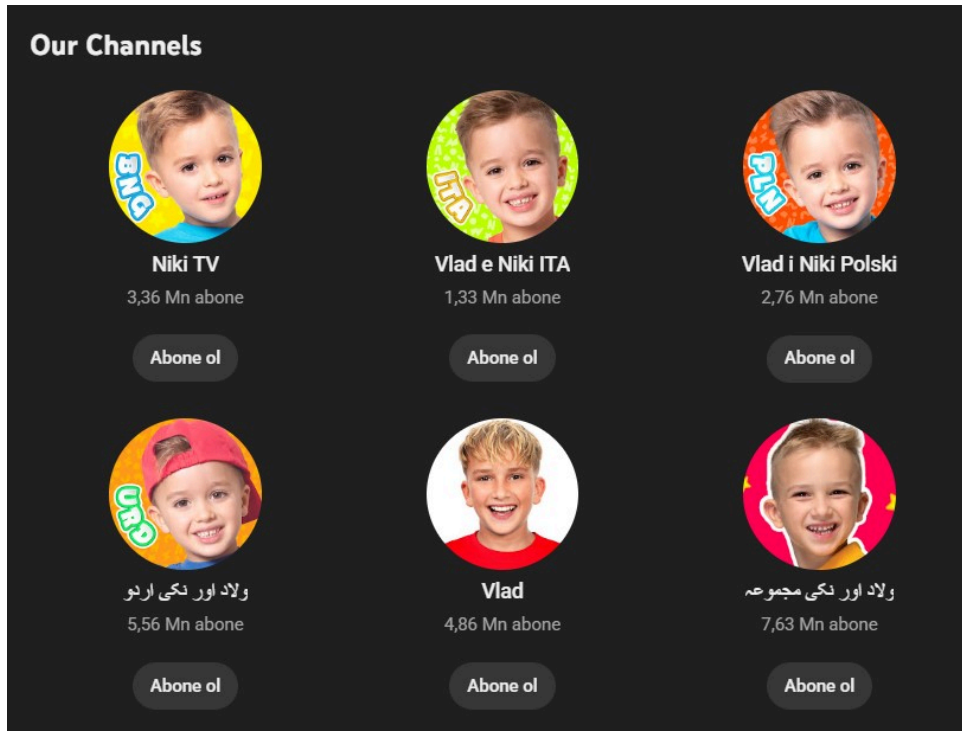
For a demographic as impressionable as young children, who now consume this AI-generated content as a primary source of auditory input, the consequences may be significant. The exposure to fragmented sentences and tonally distorted language raises urgent questions about the long-term impact on speech development and language acquisition in the next generation.

The Channel “Vlad”

The evolution of global content is best exemplified by the trajectory of massive family brands like “Vlad and Niki”, which emerged during the early explosion of YouTube Kids as a leading force alongside pioneers like Ryan’s World and Like Nastya, evolving from simple home videos of brothers Vladislav and Nikita Vashketov playing with their mother, Victoria, into a commercial powerhouse that now commands its own global toy lines and branded merchandise. In their early years, achieving global dominance required a massive logistical undertaking reminiscent of the TV industry; they had to manually create dozens of separate channels for different languages (Figure 2), employing teams of video editors, translators, and voice actors to dub their content into Spanish, Arabic, or Turkish. This “human-made” localization was the gold standard, but it still relied on a fragmented approach that kept different audiences in separate digital silos.

Figure 2

Screenshot of the first six channels under the “Our Channels” section showing alternative languages of the Vlad and Niki YouTube page. From “Vlad and Niki,” by Vlad and Niki, n.d., (<https://www.youtube.com/@VladandNiki/>).



However, as the digital landscape shifted and the children grew, the strategy evolved alongside the technology. When the eldest brother, Vlad, branched out with his own content, the necessity for a human-heavy dubbing department began to vanish. YouTube’s integration of native AI-generated multi-language audio tracks has revolutionized this process. Instead of managing a network of localized channels, a creator can now upload a single video and let YouTube’s AI handle translation and dubbing automatically. This technology doesn’t just translate words; it aims to mimic the original voice’s tone and energy, thereby removing the voice actor from the equation. It even preserves the background sound effects and music. It also translates the video title and description.

This shift marks a fundamental change in how content is consumed. While the original Vlad and Niki channels relied on paid human effort to bridge cultural gaps, the new era of AI-generated content allows a single upload to reach a global audience simultaneously. The local target audience has become a relic of the past because the platform itself now ensures that a child in Brazil and a child in Türkiye are not just watching the same video but are hearing it in their own language through an automated process. This eliminates the opportunity for others to “borrow” or adapt ideas for local markets, as the original creator now holds a monopoly over every language at once, further solidifying the global monoculture in which everyone, everywhere, is watching the exact same source. This results in a high number of views, which in turn enables the video to appear on more screens and generate more revenue.

This technological shift necessitates a rigorous examination of the practical implications of automated localization on linguistic and cultural nuances. To analyze how these nuances manifest in specific content, the following section outlines the methodological framework and process used to investigate AI-generated dubbing within this new digital landscape.

Methodology

To investigate how AI-generated automatic dubbing affects cultural representation, the current study employs a qualitative and descriptive research design. Qualitative research is adopted to feasibly categorize the errors in the rendering of language and culture stemming from AI-generated dubbing. To generate detailed insights, a purposive sampling strategy is employed. The corpus data to be analyzed consists of three videos broadcast on the YouTube channel Vlad, namely (1) *I Spent 24 Hours with the World's Shortest Woman!*, (2) *I Built a SECRET POOL in My Room!*, and (3) *Exploring the Scariest Google Maps Locations! Part 3*. All three videos feature AI-generated Turkish dubbing translations layered onto English-language audiovisual content. To gather rich samples, the videos are selected to differ in genre and communicative purpose, including lifestyle vlogging, challenge-based entertainment, and exploration content.

Data collection follows video selection. One of the two authors took on this part. The process could be described respectively with the following steps:

- 1) Watching three of the videos one by one with Turkish dub turned on.
- 2) Typing in all kinds of semantic, prosodic, and syntactic anomalies with time codes, Turkish translation (written by ear), and context.
- 3) Obtaining source text (English) from online subtitle downloader websites. Because the videos do not have official English captions, mistakes caused by automatic speech recognition were fixed manually after listening to the relevant part of the videos.
- 4) In the end, three tables, one for each video, were designed with 4 columns: timecode, translation, source text, and context.

The study employs an inductive approach to data classification, in which the evaluative categories were derived directly from observed linguistic and technical failures in the Turkish dubbing. Initially, a primary dataset was constructed by manually logging instances of anomalies that sounded unnatural or broken to the ear during a comprehensive auditory review. Following this identification of errors, the study utilized thematic analysis to group recurring patterns. While the framework is grounded in the foundational criteria of Larson's (1998) Meaning-Based TQA, namely accuracy, clarity, and naturalness, the specific classification system emerged from the qualitative examination of the data collected from the corpus. This classification comprises five types of error: (1) semantic inaccuracies, (2) pragmatic mismatches, (3) omissions and additions, (4) audiovisual misalignment, and (5) speech synthesis-related problems.

For easier reading, new tables are designed for the findings and discussion sections. To understand how this automation affects the linguistic and cultural integrity of the viewing experience, it is essential to scrutinize these columns as a whole. The following section provides a detailed analysis of this phenomenon by examining specific instances of AI-generated Turkish dubbing to determine its impact on cross-cultural communication.

The Analysis of Examples

The analysis of three YouTube videos featuring AI-generated Turkish voice-over translations reveals a recurring set of translation and speech synthesis issues. These issues are not isolated errors but appear as systematic patterns across the dataset. The findings are organized under five main categories: semantic inaccuracies, pragmatic mismatches, omissions and additions, audiovisual misalignment, and speech synthesis-related problems.

While the long titles of the videos under the methodology are self-explanatory, context will be provided here for each to facilitate understanding of the examples in this section. In the first video, which is titled

“World’s Shortest Woman,” Vlad, Rocco, and Michael, the three main characters in the selected videos, host Jyoti Amge. Throughout the day, they engage in various fun activities, such as gifting her a custom robo-elephant, eating spicy Indian food, and riding in a giant G-Wagon. The kids take Jyoti shopping for clothes, introduce her to Vlad’s family, and set up miniature experiences, including a mini-city where they film a Bollywood-style car chase. The day concludes with a surprise red-carpet event where she receives an Oscar for best actress, followed by a backyard tiki party with a cake.

In the second video called “Secret Pool”, Vlad’s friends take advantage of his vacation to secretly construct a hidden pool beneath his bedroom floor. While Vlad is away enjoying a resort, his friends cut through the floorboards to build a pool area with water slides, LED lights, and a “chill zone” featuring snacks, drinks, a TV, and water blasters. Despite being caught by Vlad’s cousin Ally, who was tasked with watching the house, they manage to finish the building. When Vlad finally returns home, he is completely shocked and thrilled by the ultimate secret pool room they built for him.

For the third video, Vlad and his friends explore abandoned locations discovered on Google Maps. Their journey begins at a prison, where they investigate solitary confinement cells, inmate records, and old uniforms. Next, they go to an abandoned secret ocean lab on an island, where they discover old scuba gear, a working microscope, and what appears to be a shark’s jaw. They also hike through a forest to reach the ruins of a mansion that burned down, uncovering an old diary and a sword, before finishing their trip outside a heavily locked military bunker where they play with an old Morse code machine.

The dataset contained 72 errors across five categories: semantic inaccuracies (41), pragmatic mismatches (21), omissions and additions (3), audiovisual misalignment (2), and speech synthesis-related problems (5). Due to the abundance of examples, only striking examples have been selected. An ‘Expected Proper Translation’ column has been added to the table to present corrected versions of the mistranslations. The accurate translations were prepared with lip-sync requirements and everyday language use in mind. Furthermore, only the first example for each video in the table will be discussed in detail.

Findings and Discussion

Semantic Inaccuracies

Semantic inaccuracies are the most frequent error type, in which the AI fails to grasp the lexical meaning of words or misinterprets polysemous terms. In many cases, the AI relies on literal translation, leading to a complete reversal or distortion of the facts.

Video Title	Source Text	AI-Generated Turkish Dub	Expected Proper Translation
World’s Shortest Woman	“I’m 24.”	“Bende 60 var.”	“Boyum 60 santimetre”
	“Scale, what have you got?”	“Terazi, orada ne var?”	“Bakalım tartı ne gösterecek?”
	“It’s a little too big for me.”	“Benim için biraz küçük.”	“Benim için biraz fazla büyük.”
Secret Pool	“We need a normal hoop.”	“İyi bir yüzük gerek.”	“Bize normal boyutta bir pota lazım.”
	“What are you thinking?”	“Ne taşıyorsun?”	“Aklında ne var?” or “Ne düşünüyorsun?”
Scariest Locations	“This is sick!” (Positive Slang)	“Bu bir felaket.”	“Bu manyak bir şey.”
	“...early 2010.”	“... 2000’lerin başında...”	“... 2010’ların başında...”
	“...pending trial.”	“Baksana, mahkeme bekliyor.”	“Baksana, tutuklu yargılanıyor.”

In the first video, the guest is the world’s shortest woman. At one point, kids ask her height, and she answers, “I’m 24.” The AI understood “24” as an inch and correctly converted it to centimeters, most likely

by gathering context from previous sentences and guessing what “24” meant. However, speech recognition understood “I am” as “I have” and translated it as “I have 60 centimeters” instead of “I am 60 centimeters”.

The second example further illustrates how the AI mishandles direct address and pragmatic register. When the host directs the phrase “Scale, what have you got?” toward a weighing scale as a playful, personified question, the AI renders it as “Terazi, orada ne var?”, meaning “Scale, what is there?”. This translation preserves the surface form but loses the rhetorical intent entirely. In Turkish, the expected natural equivalent would be something closer to “Bakalım tartı ne gösterecek?” (“Let’s see what the scale shows”), which conveys the anticipatory tone of the original. The AI’s output reads as a genuine question posed to an inanimate object rather than a performative utterance meant to build suspense for the viewer.

In the third example, when Jyoti remarks, “It’s a little too big for me,” referring to an oversized item, the AI produces “Benim için biraz küçük” (“It’s a little too small for me”), reversing the meaning entirely. The correct translation, “Benim için biraz fazla büyük,” would have preserved the original complaint. This type of error wouldn’t happen with an MT. However, there is a possibility of confusing themselves and hallucination with AI translation tools. Here, the audience can see the oversized item on screen, making the mismatch between image and dubbed dialogue immediately apparent and disorienting.

The second video offers a more interesting example. They furnish the pool area, and Rocco tries to put up a basketball hoop, but it looks like it is designed for toddlers. Michael comes up to Rocco and tells him they need a bigger hoop, saying, “We need a normal hoop.” AI translates the word “hoop” as “yüzük”. Interestingly, “yüzük” is the least used equivalent for “hoop” in dictionaries. AI completely ignores the most popular Turkish equivalents “halka” and “çember”, which could at least be meaningful in this context.

Another striking example from the second video involves a verb confusion that results in a complete semantic shift. When Michael asks, “What are you thinking?”, the AI produces “Ne taşıyorsun?” (“What are you carrying?”), substituting the mental verb think with the physical verb carry. This likely stems from acoustic similarity or ASR misrecognition of the source audio. The expected translation, “Aklında ne var?” or “Ne düşünüyorsun?”, would have preserved the conversational intent of the question.

Rounding out the sample, the third video reveals an outdated yet very interesting example. AI failed to recognize the connotation of the word “sick” in the sentence “This is sick!”. Several English words with historically negative, intense, or frightening connotations have evolved to mean “very good” or “excellent,” similar to how “terrific” originated from the sense of causing terror. These words often serve as slang intensifiers to describe something exceptionally impressive. While most MTs nowadays understand the connotation and translate accurately, the output by AI was not even a word-for-word translation because then it would “literally” mean “This (person/thing) is sick”. In reality, back-translation of the target text means “This is a catastrophe!”.

The second example from the third video reveals a numerical misinterpretation with factual consequences. When the narrator references “early 2010,” referring to the early years of the 2010s decade, the AI renders this as “2000’lerin başında” (“the early 2000s”), shifting the timeline back by an entire decade. The correct translation, “2010’ların başında,” would have accurately conveyed the intended period. In a video exploring abandoned locations and their histories, such a chronological error misleads the audience about when the depicted events or structures date from, undermining the factual integrity of the content.

The third example from the third video demonstrates the AI’s difficulty with legal terminology. The phrase “pending trial,” used in the context of an inmate record displayed on screen, is translated as “mahkeme bekliyor” (“the court is waiting”). The accurate equivalent, “tutuklu yargılanıyor,” conveys both the legal status of the person and the ongoing nature of the proceedings. This word-for-word translation leaves the audience with a phrase that is technically intelligible but terminologically wrong.

These errors suggest that while the AI is capable of recognizing speech, it lacks the “world knowledge” necessary to distinguish between a basketball hoop and a jewelry ring, or to identify when a negative word is used as positive slang.

Pragmatic Mismatches

Pragmatic mismatches occur when the AI translates the words correctly but fails to convey the social context, the speaker’s intent, or the intended illocutionary force. This is particularly evident in how the AI manages address forms and interpersonal relationships.

Video Title	Source Text	AI-Generated Turkish Dub	Expected Proper Translation
World’s Shortest Woman	“Go look at this.”	“Hızla buna bakın.”	“Bakın kim geldi!”
	“Oh, how much do you weigh if you don’t mind me asking?”	“Ne kadar kilon var, eğer sır değilse?”	“Ayıp olmazsa, kaç kilo olduğunu söyleyebilir misin?”
Secret Pool	“She is..chilling in the pool.”	“Havuzda chill yapıyor.”	“Havuzda keyif çatıyor.”
	“Guys, we have finally arrived.”	“Herkes, nihayet otelimize geldik.”	“İşte otele geldik çocuklar.”
Scariest Locations	“What are my parents going to think?”	“Atalarım bunun hakkında ne der?”	“Annem ve babam ne diyecek acaba? / çok kızacak.”
	“Tour time.”	“Tur rehberliği zamanı.”	“Hadi gezelim.”

In the first video, when the host introduces the guest to the viewers, the phrase “Go look at this” is used as an informal, friendly invitation to draw attention to the guest. However, the AI translates this as “Hızla buna bakın,” which literally means “Look at this (thing) quickly”. Pragmatically, this leads to a significant shift in tone; in Turkish, referring to guests as ‘bu’ is socially inappropriate, as this demonstrative often carries derogatory undertones when applied to people. A further example from the first video highlights the AI’s failure to transfer politeness strategies across cultures. The English phrase “Oh, how much do you weigh if you don’t mind me asking?” employs a hedging construction to soften what might otherwise be an intrusive question about a person’s body. The AI renders the hedge as “eğer sır değilse” (“if it’s not a secret”), which carries a somewhat conspiratorial tone in Turkish rather than the face-saving function of the original. The expected translation, “Ayıp olmazsa, kaç kilo olduğunu söyleyebilir misin?”, uses the culturally appropriate hedging expression. This substitution reveals that the AI can recognize a pragmatic softener but lacks cultural competence.

The second video offers a different kind of pragmatic failure involving code-mixing. During a scene where the characters are relaxing by the pool, the AI translates the action description as “He’s making chill”. Here, the AI fails to localize the English slang “chill” and instead creates a hybrid sentence that combines an English loanword with a Turkish auxiliary verb. For a platform like YouTube Kids, target text fails to convey the intended mood to a Turkish-speaking child who may not be familiar with English Gen-Z slang, making the dialogue feel artificial and unlocalized.

An additional example from the second video illustrates a mistranslated address. When the speaker says “Guys, we have finally arrived,” the AI translates the informal address “guys” as “herkes” (“everyone”), a neutral, impersonal term in Turkish rather than the warm, group-inclusive address the original conveys.

Lastly, the third video provides a humorous example of how AI can lose the interpersonal context of a sentence. When Michael jokingly asks, “What are my parents going to think?” regarding the scary locations he is exploring, the AI translates “parents” as “atalar”. In Turkish, “atalar” refers specifically to ancestors or forefathers, usually in a historical or epic context. By choosing a term associated with lineage and history

rather than the domestic and immediate “anne-baba” (mom and dad) or even the unnatural equivalence “ebeveyn” (parent), the AI completely strips the sentence of its relatable humor.

The final example from the third video conveys wrong meaning for the phrase “tour time”. AI rendered this phrase as “tur rehberliği zamanı” (“it is time for tour guiding”). The kids are not guides. Instead, they have a guide with them. Therefore, the expected equivalent, “Hadi gezelim” (“Let’s go explore”), captures both the brevity and the casual enthusiasm of the original. This example demonstrates that AI can hallucinate and add additional word (tour + guide) or choose secondary meaning for a term.

Such mismatches result in a robotic social interaction where the warmth and humor of the original creators are replaced by cold, literal statements. This pragmatic collapse strips the content of its cultural nuances, contributing to the local losses mentioned in this study.

Omissions and Additions

The AI occasionally hallucinates information or omits crucial parts of the dialogue, disrupting the narrative flow.

Video	Source Text	AI-Generated Turkish Dub	Expected Proper Translation
World's Shortest Woman	"...list of 9 activities..."	"10 maddelik liste..."	"tümü tamına 9 aktivite."

In the first video, the AI demonstrates a clear case of hallucination, generating information that contradicts the source audio. When the narrator mentions a “list of nine activities” to be completed with the guest, the Turkish dub translates this as “10 maddelik liste” (a list of 10 items). While this might seem like a minor numerical discrepancy, it highlights a fundamental reliability issue: the AI prioritized a common linguistic pattern or guess over the actual factual data provided in the English source. Such hallucinations can confuse young viewers who are following the visual cues of the list being presented on screen.

Audiovisual Misalignment

Audiovisual translation relies on the harmony between the image and the sound. The AI often fails to synchronize with visual cues or ignore nonverbal auditory cues.

Video	Source Text	AI-Generated Turkish Dub	Expected Proper Translation
World's Shortest Woman	[Laughter] “Her shoes, it give her the boost.”	[Silence] “Onun ayakkabıları boyu—”	[Laughter] “Ayakkabıları onu uzun gösteriyor.”

In the first video, the AI fails to capture the emotional atmosphere of the interaction between the hosts and their guest. While the original English audio features spontaneous laughter and emotional reactions, the Turkish dub remains completely silent during these moments. Audiovisual translation is not merely about replacing words; it involves replicating non-verbal cues such as laughter, gasps, or sighs that signify human connection. By ignoring these auditory elements (known as ‘nida’ in the Turkish dubbing industry), the AI strips the content of its warmth and spontaneity, leaving the Turkish audience with a sterile, uncanny version of the original scene.

The second example of audiovisual misalignment in the first video involves an abrupt truncation that leaves the sentence syntactically incomplete. When the narrator remarks “Her shoes, it give her the boost,” referring to how Jyoti’s shoes add to her height, the AI produces “Onun ayakkabıları boyu...” and cuts off mid-sentence, failing to deliver the predicate entirely. The expected translation, “Ayakkabıları onu uzun

gösteriyor,” would have completed the thought in a natural and concise manner. The result is a moment where sound and image, rather than reinforcing each other, actively work against comprehension.

Speech Synthesis-Related Problems

Even when the translation is semantically acceptable, the quality of the Text-to-Speech (TTS) synthesis often degrades the user experience through prosodic errors and phonetic distortions.

Video	Source Text	AI-Generated Turkish Dub	Expected Proper Translation
World's Shortest Woman	“Wait, wait...”	“Bekli beklı.”	“Bekle bekle.”
	“Lights, camera, action.”	“Kemıra, modur, başladı.”	“Işık, kamera, oyun!”
Scariest Locations	“Huge jaw”	“Dev siine.”	“Dev çene.”

In the first video, the AI demonstrates a significant phonetic breakdown during a moment of high energy. When the character says “Wait, wait” the speech synthesis engine attempts to output the Turkish equivalent “Bekle, bekle,” but produces “Bekli beklı” instead. This error is not a failure of translation but of vocalization, where the TTS model applies an incorrect phonetic structure to Turkish vowels. For young viewers, this distorted pronunciation makes the dialogue feel unnatural.

The second example of speech synthesis failure in the first video involves the classic filmmaking cue “Lights, camera, action.” The AI produces “Kemıra, modur, başladı”, a rendering that is wrong on multiple levels simultaneously. First, the word “camera” is phonetically distorted into an unrecognizable approximation that does not correspond to the standard Turkish loanword “kamera.” Second, “action” is rendered as “modur,” which appears to be a garbled attempt to pronounce “motor.” Third, and most strikingly, “lights” is omitted entirely, reducing a three-part sequence to two items. The expected translation, “Işık, kamera, oyun!” (or motor instead of oyun), is a well-established Turkish equivalent of this phrase that any viewer would recognize.

Last but not least, the second video reveals an even more severe case of phonetic distortion, leading to a total semantic collapse. While trying to describe a “huge jaw” (intended as dev çene), the AI synthesis produces the nonsensical phrase “dev siine”. This instance shows how a failure in the speech synthesis pipeline can render the content unintelligible; the intended word is so heavily distorted during the vocalization process that it ceases to be a recognizable Turkish term. Such errors go beyond sounding robotic, as they actively strip the narrative of its meaning, leaving the audience with a string of garbled sounds.

These linguistic and technical inconsistencies prompt an inquiry into the platform’s implementation of this technology. Understanding the gap between the intended function of AI dubbing and the observed errors requires a closer look at YouTube’s public policies and the unstated criteria for its high-level features.

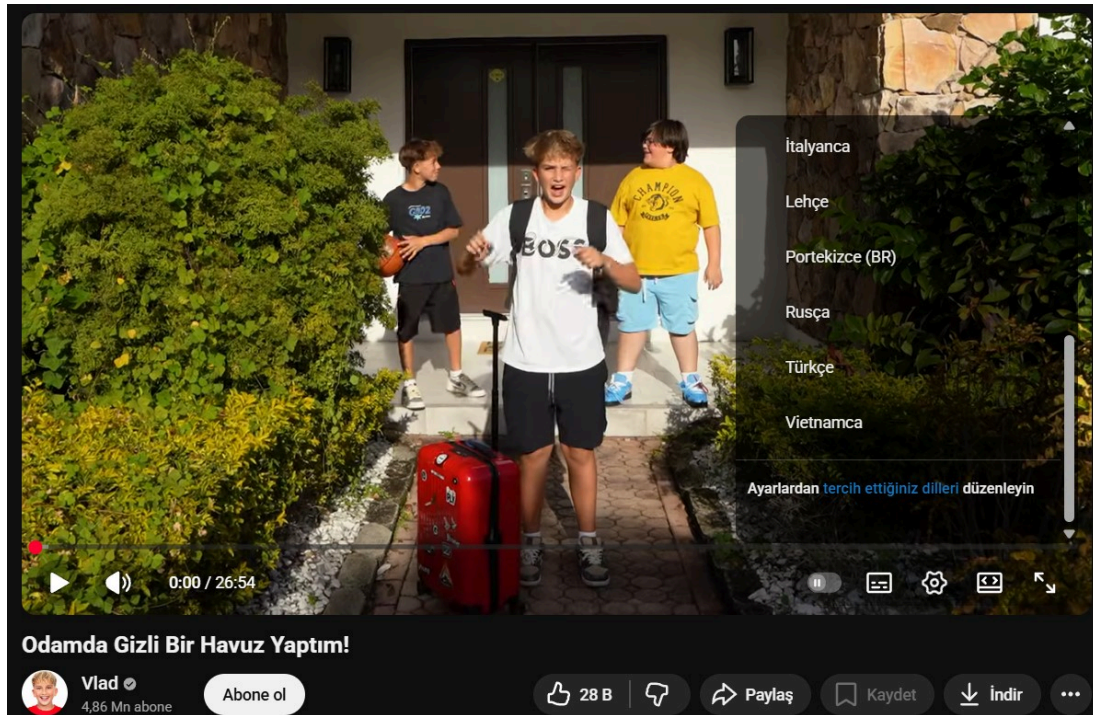
According to the YouTube Help page, YouTube publicly outlines certain technical conditions under which automatic dubbing will not be applied, including videos exceeding 60 minutes, content with little to no speech or only music, unsupported source languages, undetectable source language, overly fast speech pacing, and videos with Content ID claims due to copyrighted material. On top of these, the page states that only Turkish videos offer automatic English dubbing, but not vice versa. Additionally, this “feature is enabled by default for eligible creators,” and YouTube is “gradually expanding support for new languages and improving dub quality. As a result, some channels might have access to new languages sooner.” This brings the questions to mind: what is the unstated eligibility, how is the situation of channel Vlad different, and are there other differences, or rather, advantages of being a bigger channel? To answer the question about differences, the authors enabled automatic dubbing and tested it on a channel with approximately 80,000 subscribers, a fair number to test the newest functions, as the channel received the option in its

early phases. The channel has videos in Turkish only, ranging from gameplays with commentary to podcasts. In one of its videos, where two men and one woman talk, automatic dubbing is enabled. However, unlike Vlad's videos, the AI in this channel cannot seem to distinguish between the two men and uses the same voice for their dialogues. Also, it does not replicate the intonation quality seen in Vlad's videos. In a newer video where a commentator talks while playing a video game, the AI is much better at replicating stress, pauses, and emotions.

All of these findings reveal that automatic dubbing technology on YouTube is still in its developmental stages and that its quality is not uniform across content types and channel sizes. The performance of the AI dubbing system appears to be influenced by multiple factors: the complexity of the audio environment (e.g., the number of speakers), the nature of the content (commentary versus multi-person conversation), and potentially the channel's scale and reach. Larger channels like Vlad's seem to benefit from higher-quality dubbing, including better speaker differentiation and more accurate intonation replication, suggesting that YouTube may be prioritizing or allocating more advanced dubbing models to creators with higher subscriber counts or greater global reach. Meanwhile, the smaller channel's experience demonstrates that while access to the feature may be granted relatively early, the quality of that access is not equivalent. The single-speaker gameplay commentary performing better than the multi-speaker podcast also indicates that the AI currently handles isolated voices more competently than overlapping or alternating dialogue. These differences may also stem from the operational costs of the different automatic dubbing models working in the background.

Figure 3

Screenshot of language options from the audio track list. From "I built a secret swimming pool in my room!" by Vlad, 2026, January 30 (<https://youtu.be/e8PwGx0M6zM>). Copyright 2026 by Vlad).



Speaking of models, there is also a notable lack of transparency surrounding the feature. The name of the underlying model or its version is not disclosed anywhere, neither in video descriptions nor on any YouTube Help page, making it impossible for researchers or curious users to reference or scrutinize the technology being used. More strikingly, Figure 3 above shows that AI voice language options look like the options seen on authentic dubbed YouTube videos. There is no label or indication whatsoever that a video has been AI-

dubbed, meaning viewers may consume content without knowing that its voices are artificially generated. This is particularly concerning for younger audiences, such as children, who may sense something unusual about the voice yet remain unaware of its AI origin. They may even internalize it as a real human voice. At a time when other platforms are beginning to introduce AI-content labeling, YouTube appears to have yet to adopt a clear transparency policy for this feature, leaving both viewer awareness and creator accountability in an ambiguous space.

Conclusion

This study investigated the implications of AI-generated Turkish dubbing on the children’s YouTube channel “Vlad” to assess its impact on cultural and linguistic representation, applying Larson’s Meaning-Based Translation Quality Assessment (TQA) model to a purposive sample of video content. Overall, in light of the frequent errors mentioned, it is possible to say that the three criteria of Larson’s model, accuracy, naturalness, and clarity, are not maintained in the automatic dubbing.

Addressing the first research question, the findings indicate that AI dubbing severely distorts both semantic and pragmatic meaning. The AI’s inability to distinguish polysemous terms, such as mistranslating a basketball “hoop” as a ring (jewelry), or its failure to recognize social contexts, such as using the historical term “atalar” for “parents,” results in unnatural and often nonsensical Turkish output. Regarding the second research question, the results suggest that prioritizing lip sync and failed automatic speech recognition results in neglecting cultural integrity and linguistic accuracy. This creates a phenomenon of “Global Voices, Local Losses,” where content travels globally but remains fundamentally foreign. Finally, answering the third question, while larger channels like “Vlad” appear to benefit from more advanced AI models that offer better speaker differentiation and intonation, technology still struggles with complex, multi-person dialogues and experiences significant phonetic breakdowns regardless of channel scale. Aside from the channel variable, content topic and genre reveal a significant difference. In content focusing on social interaction and personal statistics, the AI struggles with unit conversions (e.g., mistranslating height) and experiences phonetic breakdowns during high-energy moments. In challenge videos, the AI fails to identify domain-specific objects (e.g., a hoop and a ring) and creates artificial hybrid sentences when localizing modern slang (e.g., “chilling”). Lastly, for narrative-driven exploration, the AI often misinterprets slang connotations (e.g., viewing “sick” as a catastrophe) and loses context, leading to the use of historical terms like “ancestors” for “parents”.

The persistent phonetic distortions, such as “bekli” or the nonsensical “dev siine,” highlight a critical gap in human oversight. Although YouTube allows eligible creators to preview these automatic tracks, the abundance of blatant errors suggests the feature is often used without rigorous review. This reflects a broader industry trend, as depicted by Way (2013), in which the “degree of human involvement required – or warranted – in a particular translation scenario will depend on the purpose, value and shelf-life of the content”. Currently, the high frequency of YouTube content leads creators to prioritize speed over quality, treating it as having a short shelf life that does not warrant human intervention. However, the future utility and ethical standing of AI dubbing could be significantly improved if platforms integrated features for manual script editing and audio regeneration, allowing creators to fix specific issues before publishing.

Future research should employ quantitative methodologies to establish a statistical baseline for AI dubbing quality. While this research focused on qualitative, meaning-based inductive categories, other frameworks offer different evaluative paths. Integrating quantitative tools would enable more precise comparisons of error density across various AI models and language pairs. Interdisciplinary reception studies could also show the effects of AI-generated content for children. Furthermore, there is a pressing need for

longitudinal studies to evaluate the long-term impact of prolonged exposure to distorted AI voices on the speech development and language acquisition of young children.



Peer Review	Externally peer-reviewed.
Author Contributions	
Conflict of Interest	The author has no conflict of interest to declare.
Grant Support	The author declared that this study has received no financial support.

Hakem Değerlendirmesi	Dış bağımsız.
Yazar Katkısı	
Çıkar Çatışması	Yazar çıkar çatışması bildirmemiştir.
Finansal Destek	Yazar bu çalışma için finansal destek almadığını beyan etmiştir.

Author Details	Ebru Çavuşoğlu (Assistant Professor,)
Yazar Bilgileri	¹ Samsun University, Department of English Translation and Interpreting, Samsun, Türkiye  0000-0003-1666-6002  ebru.cavusoglu@samsun.edu.tr
	Erdem Celal Gören (Research Assistant,)
	² Samsun University, Department of English Translation and Interpreting, Samsun, Türkiye  0000-0001-6535-5812

References

- Aluvihara, S., Pestano-Gupta, F., Alqasi, N., Al-Ani, I., Karimkhani, M., Radfar, M., . . . Hamdi, M. (2025). The importance of artificial intelligence (AI) tools in modern science, engineering, and technological research and innovations: A review. *American Journal of Artificial Intelligence*, 9(2), 229–241. <https://doi.org/10.11648/j.ajai.20250902.23>
- Amazon. (2025). *Prime Video tests AI-powered dubbing in English and Spanish*. Amazon News. Retrieved from <https://www.aboutamazon.com/news/entertainment/prime-video-ai-dubbing-english-spanish>
- Balakrishnan, J., & Griffiths, M. D. (2017). Social media addiction: What is the role of content in YouTube? *Journal of Behavioral Addictions*, 6(3), 364–377. <https://doi.org/10.1556/2006.6.2017.058>
- Bar, S., Barzilai, A., Elias, I., Fink, M., Frumar, J., Ho, H., . . . Tavory, R. (2009). YouTube's collaborative annotations. In *WEBCENTIVES '09: 1st International Workshop on Motivation and Incentives*. Retrieved from <https://static.googleusercontent.com/media/research.google.com/en//pubs/archive/36735.pdf>
- Brannon, W., Virkar, Y., & Thompson, B. (2023). Dubbing in practice: A large-scale study of human localization with insights for automatic dubbing. *Transactions of the Association for Computational Linguistics*, 11, 419–435.
- Chaume, F. (2020). *Audiovisual translation: Dubbing*. Routledge.
- Çavuşoğlu, E. (2023). Is Shrek Still Funny in Turkish? An Analysis of The Strategies Employed During The Turkish Dubbing Translation of Humor in The Animated Film Shrek. *IU Journal of Translation Studies*, 19, 95–106. <https://doi.org/10.26650/iujts.2023.1362904>
- DeepL. (n.d.). Retrieved June 3, 2026, from <https://www.deepl.com/en>
- Elias, N., & Sulkin, I. (2017). YouTube viewers in diapers: An exploration of factors associated with amount of toddlers' online viewing. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, 11(3).
- Gallo, P., & Iordache, C. (2026). Localisation at scale. In *Streaming Europe: How transnational streaming is reshaping European film and television markets*. Taylor & Francis.
- Yao, R. (2026, April 28). Celebrating 20 years of Google Translate: Fun facts, tips and new features to try. Retrieved June 3, 2026, from Google website: <https://blog.google/products-and-platforms/products/translate/fun-facts-google-translate-20-years/>
- House, J. (2014). *Translation quality assessment: Past and present*. Routledge.
- Jin, H., & Yuan, Z. (2023). The application of machine translation in automatic dubbing in China: A case study of the feature film Mulan. *Linguistica Antverpiensia, New Series: Themes in Translation Studies*, 22, 80–94.



- Karakanta, A., Bhattacharya, S., Nayak, S., Baumann, T., Negri, M., & Turchi, M. (2020, December). The two shades of dubbing in neural machine translation. In *Proceedings of the 28th International Conference on Computational Linguistics* (pp. 4327–4333).
- Larson, M. L. (1998). *Meaning-based translation: a guide to cross-language equivalence*. University Press of America.
- Merriam-Webster. (n.d.). Idiot box. In *Merriam-Webster.com dictionary*. Retrieved June 2, 2026, from <https://www.merriam-webster.com/dictionary/idiot%20box>
- McKeever, S. (2003). Understanding web content management systems: Evolution, lifecycle and market. *Industrial Management & Data Systems*, 103(9), 686–692. doi:10.1108/02635570310506106
- Nida, E. (1964). *Toward a Science of Translation*. Leiden: Brill
- Nord, C. (2016). Function + loyalty: Theology meets skopos. *Open Theology*, 2(1).
- Özkaya Marangoz, E. (2023). The transformative role of artificial intelligence and machine learning in interpreting and language services. *RumeliDE Dil ve Edebiyat Araştırmaları Dergisi*, (36), 1591–1598.
- Pedersen, J. (Ed.). (2005, May). Mutra 2005. Proceedings of Marie Curie Euroconferences on challenges of multidimensional translation, Saarbrücken, Germany: EU-High-Level Scientific Conference Series.
- Reiss, K. (1981). Type, Kind and Individuality of Text: Decision Making in Translation. *Poetics Today*, 2(4), 121–131. <https://doi.org/10.2307/1772491>
- Sargın, M. (2024). Google Çeviri Programı Üzerine Bir İnceleme: Sıfat Yantümcelerinin Çevirisindeki Hatalar. *Elektronik Sosyal Bilimler Dergisi*, 23(92), 1525–1543.
- Spiteri Miggiani, G. (2022). Measuring quality in translation for dubbing: a quality assessment model proposal for trainers and stakeholders. *XLinguae*, 15(2), 85–102.
- Spiteri Miggiani, G. (2023). Quality in translation and adaptation for dubbing: Applied research in a professional setting. *The Journal of Specialised Translation*, 40, 297–321.
- Statista. (2026). *Most visited websites worldwide as of February 2026, by monthly traffic*. Retrieved from <https://www.statista.com/statistics/1201880/most-visited-websites-worldwide/>
- Stewart, J. K. (1998). Interactive television at home: Television meets the Internet. In *Television of the future—or: the future of television*. Aalborg, Denmark: Aalborg University Press.
- Suleyman, M. (2025). *The coming wave: AI, power, and our future*. New York, NY: Random House.
- Şahin, M. (2022). Çeviri, teknoloji ve insan. In *BAIBU-ICASTIS Book of Congress Proceedings*.
- Tate, M. (2025). Amazon halts AI anime dub “beta” after widespread ridicule. Engadget. Retrieved from <https://www.engadget.com/entertainment/amazon-halts-ai-anime-dub-beta-after-widespread-ridicule-141501051.html>
- Varela, F. C. (2007). Quality standards in dubbing: a proposal. *TradTerm*, 13, 71–89.
- Vlad and Niki. (n.d.). Home [YouTube channel]. YouTube. Retrieved from <https://www.youtube.com/@VladandNiki/>
- Vlad ve Niki. (2023, April 21). *Vlad ve Niki ile Bat ya da Yüz - Çocuklar İçin Harika Bilim Deneyi* [Video file]. Retrieved from https://youtu.be/5o_arN_9o3o
- Vlad. (2026, January 24). *I spent 24 hours with the world's shortest woman!* [Video file]. Retrieved from <https://youtu.be/rldtJt9f1wY>
- Vlad. (2026, January 30). *I built a secret swimming pool in my room!* [Video file]. Retrieved from <https://youtu.be/e8PwGx0M6zM>
- Vlad. (2026, January 31). *Exploring the scariest Google Maps locations! Part 3* [Video file]. Retrieved from <https://www.youtube.com/watch?v=92YILoGiiY4>
- Way, A. (2013). Traditional and emerging use-cases for machine translation. *Proceedings of Translating and the Computer*, 35, <https://aclanthology.org/2013.tc-1.12.pdf>
- YouTube. (n.d.-a). *About YouTube*. Retrieved June 2, 2026, from <https://about.youtube/>
- YouTube. (n.d.-b). *Use automatic captioning*. Retrieved June 2, 2026, from <https://support.google.com/youtube/answer/6373554?hl=en>
- YouTube. (n.d.-c). *Add audio descriptions*. Retrieved June 2, 2026, from https://support.google.com/youtube/answer/16166822?hl=en&ref_topic=7296214
- YouTube. (n.d.-d). *Use automatic dubbing*. Retrieved June 2, 2026, from https://support.google.com/youtube/answer/15569972?hl=en&ref_topic=7296214
- YouTube. (n.d.-e). *Add Multi-language features to your videos*. Retrieved June 2, 2026, from https://support.google.com/youtube/answer/13338784?hl=en&ref_topic=7296214

Appendix

Table 1

I Spent 24 Hours with the World’s Shortest Woman!

Timecode	Translation (AI-Generated Turkish Dub)	Source Text	Category
0:10	“Onun gözleriyle dünyayı göreceksiniz.”	You will see the world through her eyes	Speech synthesis–related problems
0:38	“O burada, hızla buna bakın.”	She's here. Go look at this.	Pragmatic mismatches
0:47	“10 maddelik liste hazırladım.”	I prepared a list of nine activities...	Omissions and additions
1:07	“Aman tanrım, sesini çok seviyorum.”	I love your voice	Pragmatic mismatches
1:21	“Jyoti, burada olduğuna çok sevindik ama biliyor muydun ki Guinness Rekorlar Kitabı’nda bile var?”	Jodie, this is such a pleasure to have you here. Actually, did you know that she holds Guinness World Records?	Pragmatic mismatches
1:53	“Bende 60 var.”	I’m 24 (inch)	Semantic inaccuracies
1:58	“Onun ayakkabıları boyu—”	Her shoes, it give her the boost.	Audiovisual misalignment
2:14	“Ne kadar kilon var, eğer sır değilse?”	Oh, how much do you weigh if you don't mind me asking?	Pragmatic mismatches
2:27	“Terazi, orada ne var?”	Scale, what have you got?	Semantic inaccuracies
3:11	“Umarım bu kesinlikle bir basketbol potası değil.”	I hope it's not basketball hoop.	Semantic inaccuracies
3:16	—	—	Audiovisual misalignment
4:13	“Sonra şef aşçıyı çağırdık.”	After we invited a chef...	Semantic inaccuracies
5:21	“Çok ilginç bir tadı var.”	That’s an exciting taste	Semantic inaccuracies
5:57	“Peki bu nasıl?”	How do you like this?	Pragmatic mismatches
6:33	“Benim için biraz küçük.”	Little too big for me	Semantic inaccuracies
6:51	“Senin en sevdiğin ne?”	And which one is your favorite?	Semantic inaccuracies
6:58	—	—	Omissions and additions
7:08	“Hindice biliyor musun?”	Do you know Hindi?	Semantic inaccuracies
7:32	“Onun boyu nedeniyle seçimi çok daha sınırlı, sıradan bir insandan.”	Because of her size, her options are way more limited than an average person.	Pragmatic mismatches
7:41	“Onun için bir güvenlik aldık.”	So, we brought a security guard.	Semantic inaccuracies
8:10	“Dur dur dur, beklı beklı.”	Wait, wait, hold on.	Speech synthesis–related problems

Timecode	Translation (AI-Generated Turkish Dub)	Source Text	Category
8:59	"Gerçek Vipi kabulümüz olacak..."	This is a true VIP	Speech synthesis–related problems
11:46	"Bu da Lamborghini Leo gerçekten."	This is Leo's Lambo.	Semantic inaccuracies
12:15	"Jyoti kesinlikle hoşuna gidecek."	Jody is 100% going to love this.	Semantic inaccuracies
12:20	"Bizim bile DJ'imiz var."	WE EVEN GOT A personal DJ.	Pragmatic mismatches
12:50	"Kemıra, modur, başladı."	Lights, camera, action.	Semantic inaccuracies / Speech synthesis
13:42	"Ehliyetlerinizi alabilir miyim?"	You got a license?	Semantic inaccuracies
16:40	"Seni geldiğin için teşekkür etmek istiyoruz."	We would like to thank you for coming...	Semantic inaccuracies
16:45	"Beğeni bırakın."	Please like...	Pragmatic mismatches

Table 2*I Built a SECRET POOL in My Room!*

Timecode	Translation (AI-Generated Turkish Dub)	Source Text	Category
0:08	"Ne yapıyorsunuz? Dalga mı?"	Are you kidding me?	Semantic inaccuracies
0:44	"Şimdi sır havuzunu yapma zamanı."	Now it's time to make the secret pool.	Semantic inaccuracies
0:51	"Hazır mısınız inşa etmeye başlamak?"	You guys ready to get this thing built?	Pragmatic mismatches
5:10	"-lar alışveriş yaparken, ben otele yerleşiyorum."	While the guys are out shopping, I'm checking into my resort.	Omissions and additions
5:17	"Herkes, nihayet otelimize geldik."	Guys, we have finally arrived.	Pragmatic mismatches / Omissions
9:45	"Çocuklar iskelet toplarken, ben lezzetli yemekler yiyorum."	While the guys are building the pool frame, I'm enjoying some delicious food.	Semantic inaccuracies
12:04	"Vlad için harika bir mini yüzük aldık."	We got this awesome mini hoop for Vlad.	Semantic inaccuracies
12:08	"Ne taşıyorsun?"	What are you thinking?	Semantic inaccuracies
12:11	"İyi bir yüzük gerek."	...we need a normal hoop.	Semantic inaccuracies
20:15	"Eğer bir şey kırarlarsa, işlerini bitirirler."	...if they broke something, they will be done for.	Semantic inaccuracies
20:29	"Nerede onlar mı?"	Where are they	Semantic inaccuracies
21:46	"Yeterince gücü var mı çocukların ben tatilden döndüm taze."	Well, the guys used all their energy I gave from my vacation fresh.	Semantic inaccuracies
23:00	"Aman tanrım bu ne nerede ben?"	Oh my god, what is this, where am I?	Pragmatic mismatches
23:56	"Sonunda evdeyim ve beni bir devamın beklediğini beklemiyordum."	I'm finally home and I did not expect what was to come.	Semantic inaccuracies
24:50	"Bende bir soru var Allie nerede."	I just have one question. Where's Ally?	Semantic inaccuracies
25:32	"Allie nerede? Havuzda chill yapıyor."	Oh, she's at the pool in your backyard.	Pragmatic mismatches

Timecode	Translation (AI-Generated Turkish Dub)	Source Text	Category
—	“Hadi gidelim.” (repeated)	Let’s go	Pragmatic mismatches

Table 3*Exploring the Scariest Google Maps Locations! Part 3*

Timecode	Translation (AI-Generated Turkish Dub)	Source Text	Category
1:02	“Çok ileri gittiniz. Şu anda ne kadar tehlikede olduğunuzu gerçekten bilemezsiniz.”	You are so far off the beaten path, you have no clue.	Pragmatic mismatches
1:21	“Atalarım bunun hakkında ne der?”	What are my parents going to think about this?	Pragmatic mismatches
1:37	“Artık ebeveynlerimin ne kadar sinirleneceğini hayal etmiştim.”	I was already imagining how furious my family would be.	Semantic inaccuracies
1:46	“Tur rehberliği zamanı.”	Tour time	Pragmatic mismatches
2:03	“Arkadaşlar terk edilmiş bir hapisteyiz. Sizi eğlendirmek için neler yapmadığımızı görün.”	...guys. Right now, we are in this abandoned prison...	Semantic inaccuracies
2:14	“Vay dostum bu çok havalı.”	Wow, dude. This is so awesome.	Pragmatic mismatches
2:27	“Onu 1976 yılında yaptılar ve 2000’lerin başında kapattılar.”	early 2010	Semantic inaccuracies
2:46	“Kalınlığı birkaç inç.”	This right here is like an inch and a half...	Semantic inaccuracies
2:55	“Baksana, mahkeme bekliyor.”	Pending trial	Semantic inaccuracies
2:59	— / “75000 mi?”	They got bail, 75000	Omissions and additions
3:39	“İşte tek kişilik hücre böyle görünüyor.”	So this is how the solitary confinement room would look like	Audiovisual misalignment
3:50	—	—	Semantic inaccuracies
4:01	“Michael neredesin? Ses kes.”	Michael where are you? Copy.	Semantic inaccuracies
4:45	—	—	Speech synthesis-related problems
5:00	“Busi bırak.”	Let us out	Pragmatic mismatches
5:45	“Şimdi chill yapıyorum. Vlad çalışıyor.”	—	Pragmatic mismatches
6:04	“Gürlükten geçiyoruz.”	See the forest	Semantic inaccuracies
6:33	“Bu hükümetin bir komplosu, her şeyi gizliyorlar...”	There’s a government conspiracy that...	Pragmatic mismatches
9:57	“Burada bir kostüm buldum.”	This is where I found the costume	Semantic inaccuracies
11:39	“Bu bir çatal dişi.”	That’s a head for a pitchfork	Semantic inaccuracies
12:42	“Dev siine.”	Huge jaw	Speech synthesis-related problems
13:59	“Bu bir felaket.”	This is sick	Semantic inaccuracies
14:57	“Vahşi görünüyor.”	This looks sick	Semantic inaccuracies
16:10	“O artifactları topluyordu.”	Oh, so he collected artifacts.	Semantic inaccuracies
18:23	“Çekim yapmamız için gerekli izni aldı ama maalesef sunucu göndermedikleri için...”	Host	Semantic inaccuracies
18:32	“Bak burada ne kadar beton var...”	Look how huge that concrete is...	Pragmatic mismatches
19:19	“Duymak ama hiçbir şey görmüyorum.”	I can hear but I can’t see	Semantic inaccuracies

Timecode	Translation (AI-Generated Turkish Dub)	Source Text	Category
21:47	“Bu bir morse.”	This is morse code	Semantic inaccuracies