

# Assessing Quality in Machine-Translated English-Turkish Climate Communication Subtitles: A Longitudinal Study

**Makine Çevirisiyle Üretilen İngilizce-Türkçe İklim İletişimi Altyazılarında Kalite Değerlendirmesi: Boylamsal Bir Araştırma**

Araştırma/Research

**Rumeysa Nur ÇAKMAK\*, Muhammed BAYDERE\*\***

\*Res. Asst., Avrasya University, Faculty of Arts and Sciences, Translation and Interpreting, [rumeysanur.cakmak@avrasya.edu.tr](mailto:rumeysanur.cakmak@avrasya.edu.tr), ORCID ID: [orcid.org/0009-0003-1743-0613](https://orcid.org/0009-0003-1743-0613)

\*\*Asst. Prof. Dr., Karadeniz Technical University, Faculty of Letters, Western Languages and Literature, [muhammedbaydere@ktu.edu.tr](mailto:muhammedbaydere@ktu.edu.tr), ORCID ID: [orcid.org/0000-0001-6827-2027](https://orcid.org/0000-0001-6827-2027)

## ÖZET

İklim değişikliği alanındaki iletişim giderek dijital medyaya daha fazla bağlı hâle gelirken, makine çevirisi (MÇ) altyazılarının kalitesi yanlış bilgilendirmenin önlenmesinde kritik rol oynamaktadır. Bu çalışma, teknik terminoloji açısından yoğun olan ve anlamsal açıdan yüksek hassasiyet gerektiren bu alanda İngilizce-Türkçe MÇ altyazılarının kalitesini incelemektedir. Bu kapsamda Hükümetlerarası İklim Değişikliği Paneli'nin (IPCC) 2013 ve 2022 yıllarında yayınlanan iki videosunun 2022 ve 2025 yıllarına ait otomatik altyazıları karşılaştırmalı ve boylamsal olarak analiz edilmiştir. Jan Pedersen'in (2017) işlevsel eşdeğerliği (anlamsal ve biçimsel doğruluk), kabul edilebilirliği (dilbilgisi, yazım ve dilin ifade özellikleri açısından uygunluk) ve okunabilirliği (teknik özellikler) ölçen FAR modeli kullanılarak altyazı kalitesi sistematik biçimde değerlendirilmiş; MÇ'deki gelişmelerin altyazı doğruluğu, akıcılığı ve kullanılabilirliği üzerindeki etkisi incelenmiştir. Bulgular, altyazı kalitesinde belirgin bir iyileşme olduğunu göstermektedir: ilk videoda toplam hata oranı %63,04, ikinci videoda ise %53,40 oranında azalmıştır. Bu iyileşmeler eşit dağılmamıştır; birinci videoda en yüksek azalma okunurlukta (%73,77), ikinci videoda ise kabul edilebilirlikte (%76,25) gözlemlenmiştir. İşlevsel eşdeğerlik ise daha sınırlı bir ilerleme göstermiş (birinci videoda %53,18, ikinci videoda %27,91) ve 2025 altyazılarında da en yüksek hata puanına sebep olan tür olmaya devam etmiştir; bu durum, incelikli anlam aktarımında kalıcı sorunların varlığını göstermektedir. Elde edilen bulgular, otomatik altyazılar, çevirmen eğitimi ve sürdürülebilirlik

odaklı medya stratejileri açısından önemli çıkarımlar sunmaktadır. Çalışma ayrıca, iklim konusunda içerik üreten kurumlar için yanlış bilgilendirme riskini azaltacak ve iletişimin etkinliğini artıracak “insan merkezli iyileştirilmiş çeviri” modellerine (O’Brien, 2023) yönelmenin gerekliliğini savunmaktadır. Öte yandan, kullanıcılar açısından çeviri okuryazarlığı, özellikle de MÇ okuryazarlığı, yalnızca teknik bir beceriye değil, erişilebilir ve güvenilir iletişimi sağlamak için gerekli olan temel bir toplumsal edince karşılık gelmektedir. Bu edincin, çeviri programlarının ötesine geçen eğitim uygulamaları aracılığıyla (örn. hedefe yönelik atölyeler, ilgili derslere entegre edilen kısa modüller ve bölümler arası sertifika seçenekleri) teşvik edilmesi önerilmektedir. Böylelikle geniş kitlelerin, genel olarak MÇ aracılığıyla sunulan dijital içerikleri özel olarak da iklim konusundaki bu yolla aktarılan bilgileri eleştirel biçimde değerlendirip anlamalarına katkıda bulunabilecektir. Çalışma, MÇ değerlendirmesini çevre iletişimiyle ilişkilendirerek yüksek çeviri kalitesinin küresel iklim okuryazarlığını güçlendirmedeki ve bilinçli politika katılımını desteklemedeki rolünü vurgulamaktadır.

**Anahtar Sözcükler:** makine çevirisi, görsel-işitsel çeviri, kalite değerlendirmesi, makine çevirisi okuryazarlığı, sürdürülebilirlik.

#### ABSTRACT

As climate change communication increasingly relies on digital media, the quality of machine-translated subtitles becomes a critical factor in preventing misinformation. This study investigates the quality of English–Turkish machine-translated subtitles in climate change communication, a domain characterized by technical terminology and the need for high semantic precision. Focusing on two Intergovernmental Panel on Climate Change (IPCC) videos released in 2013 and 2022, we conduct a longitudinal analysis comparing their automatic subtitles from 2022 and 2025. Using Jan Pedersen’s (2017) FAR model—which assesses functional equivalence (semantic and stylistic accuracy), acceptability (grammatical and idiomatic correctness), and readability (technical presentation)—we offer a systematic assessment of translation quality and analyze how advances in machine translation (MT) have affected subtitle accuracy, fluency, and usability. Results demonstrate substantial improvement, with total error scores reduced by 63.04% for the first video and 53.40% for the second. These gains were uneven: the largest reduction in errors was in readability for the first video (73.77%) and in acceptability for the second (76.25%), while functional equivalence showed more limited progress (27.91%-53.18%) and remained the leading source of errors in 2025, with persistent issues in conveying nuanced meaning. The findings have implications for automated subtitle practices, translator training, and sustainability-focused media strategies. We argue that for producers of climate communication, a shift towards “a human-centered augmented translation” (O’Brien, 2023) model is necessary to reduce the risk of misinformation and improve the effectiveness of climate communication. Furthermore, for consumers, we contend that translation literacy, specifically MT literacy, is not merely a technical skill but a fundamental public competency, crucial for ensuring accessible and trustworthy communication. We recommend fostering this skill through educational initiatives beyond translation programs such as targeted workshops, short modules in relevant courses, and cross-departmental certificate options. This, in turn, will facilitate the capacity of non-specialist audiences to critically assess and comprehend MT-mediated digital content broadly, and information pertaining to climate conveyed via this modality specifically. The study highlights the role of high-quality translation in strengthening global climate literacy and supporting informed policy engagement by linking the evaluation of MT with environmental communication.

**Keywords:** machine translation, audiovisual translation, quality assessment, machine translation literacy, sustainability.

## 1. Introduction

Climate change is an unprecedented global challenge. The World Meteorological Organization (WMO, 2025) confirmed that “the past ten years 2015–2024 are the ten warmest years on record,” highlighting a warming trend. Similarly, the Intergovernmental Panel on Climate Change (IPCC) stresses that action is essential: government and societal measures “play a crucial role in enabling and accelerating shifts in development pathways towards sustainability and climate resilient development” (IPCC, 2023). These findings underscore the need to disseminate accurate climate information to global audiences. Digital media platforms have emerged as key channels for such communication. YouTube is one of the dominant global platforms. DataReportal’s Digital 2024 (Kemp, 2024) report shows that the typical user spends over 28 hours 45 minutes per month on YouTube, making it the second-most-used social app after TikTok. This viewing time indicates YouTube’s reach for environmental content. Climate science organizations like the IPCC increasingly use YouTube to share knowledge, relying on subtitles and translations to overcome language barriers and reach non-English speakers. The IPCC plays a foundational role in climate change communication for policymakers and the global public. As a body created to provide assessments and inform mitigation and adaptation responses, the accessibility of its information is central to its impact (IPCC — *Intergovernmental Panel on Climate Change*, n.d.). With digital platforms, the IPCC and similar organizations have expanded their communicative reach beyond policy circles to broader audiences. Machine translation (MT), specifically through automatic subtitle generation, facilitates this expansion.

Meanwhile, MT and automated subtitling technologies have rapidly evolved. Bowker (2020) observes that the advent of online MT tools and neural network methods has generated “a growing interest in exploring the potential of machine translation”. Díaz Cintas (2013) documents how platforms like YouTube have integrated automatic captioning. For example, YouTube announced features in 2009–2010 that automatically generate captions and machine-translated subtitles, aiming to make video content accessible globally. Automatic captions enable people worldwide to access video content in their language. Such innovations suggest that artificial intelligence (AI)-driven subtitles could enhance the dissemination of climate messages across linguistic boundaries. However, these technological advances also raise concerns about translation quality. In particular, the reliability of machine-generated subtitles remains uncertain, especially in specialized domains and for languages that have not been studied extensively. Li (2024) reports that despite MT’s speed and accessibility, “concerns persist about the quality of machine-translated content” (p. 174). In climate change context, a field with technical terminology and critical implications, errors in auto-translated subtitles could distort messages or erode trust. This issue calls for careful assessment of automatic subtitling, especially in environmental content. Addressing these concerns also requires stronger MT literacy, the ability to critically interpret, evaluate, and apply machine-generated

translations (Bowker, 2020). In climate science, where public understanding influences collective action, MT literacy is not merely technical proficiency but a core of accessible and trustworthy communication.

This study evaluates English–Turkish automatic subtitles for the two most-watched IPCC YouTube videos, comparing their subtitles released in 2022 and 2025. Using Jan Pedersen’s FAR model (2017), a penalty-based subtitle quality assessment framework that evaluates semantic and stylistic accuracy (functional equivalence), linguistic well-formedness (acceptability), and technical readability, the study investigates longitudinal changes in quality and analyzes the benefits and limitations of MT for disseminating climate information. By applying a systematic model to real-world audiovisual data, the study contributes to translation studies methodology and assesses how effectively climate information is conveyed to environmental communication practice. The findings aim to guide content creators, inform policymakers, and help developers refine MT tools for sustainability-related media while fostering MT literacy among end-users.

The remainder of this paper is structured as follows: first, it reviews research on MT-based subtitling, its technical foundations, and the quality challenges documented in the literature; next, it critically examines the opportunities and limitations of MT in global communication, with particular attention to sustainability discourse, highlighting both reach and accuracy risks; and, it situates Pedersen’s (2017) FAR model, outlining its principles, scoring scheme, and prior applications, as the evaluative lens adopted. The methodology section then details the data selection process and the FAR model applied. The results and discussion present subtitle quality at two time points, using quantitative and qualitative evidence and interpreting it in light of current advances and challenges in automatic subtitling. Finally, the conclusion summarizes the study’s key contributions and identifies directions for future research at the intersection of translation technology and climate change communication.

## **2. MT and Automatic Subtitling Technologies**

MT has evolved from rule-based systems to statistical approaches and, most recently, neural MT (NMT), which leverages deep learning for greater fluency and contextual awareness (Bowker, 2020). Contemporary subtitling workflows typically combine automatic speech recognition, segmentation, and NMT to generate interlingual subtitles with minimal human input (Díaz Cintas, 2019). Major platforms such as YouTube, Netflix, and Bilibili now routinely offer auto-generated subtitles, illustrating MT’s role in reducing language barriers and supporting cross-cultural exchange (Hagström & Pedersen, 2022). The use of MT in subtitling dates to early projects including ALTo, a rule-based MT system for English–Spanish closed captions (Popowich et al., 2000), which achieved 70–80% accuracy. Armstrong et al. (2006) tested example-based MT for English–German/Japanese subtitles but noted significant segmentation and creativity issues. Sousa (2011), using Moses and Google Translate (SMT), found post-edited MT subtitles comparable in quality to human-produced ones by BLEU scores. Beyond technical systems, subtitling theory stresses readability and accessibility. Karamitroglou (1998)

emphasized viewer comprehension and enjoyment, especially by enhancing on-screen readability; Gottlieb (2001) argued semantic fidelity must remain central; and Díaz Cintas (2001) highlighted translators' creative and ethical responsibility in audiovisual translation. However, persistent challenges, particularly domain-specific constraints, undermine subtitle quality. Lu (2023) reports that in Chinese–English subtitles, time/space constraints, cultural mismatches, and segmentation issues often yield overly literal or verbose translations unsuited to reading speed required in subtitled media. Jin and Yuan's (2023) study of automatic dubbing in *Mulan* similarly found mistranslations of culture-specific terms and register shifts when MT lacked adaptation.

The specialized content found in climate change and sustainability communication further deepens these challenges. Such material typically includes technical terminology and multimodal elements that demand high semantic precision. Rivera-Trigueros's (2021) systematic review of MT systems and quality-assessment frameworks shows that even state-of-the-art NMT often produces literal phrasing and segmentation errors that disrupt viewer comprehension. Advances in MT architecture, like deep-transfer learning with semantic awareness, have demonstrated error-detection improvements of up to 15% over baseline NMT (Yan et al., 2024), yet persistent issues with technical jargon and context modeling remain problematic for sustainability-related videos. These flaws may reduce trust and clarity in climate-related messages. Addressing these risks requires both technical and pedagogical strategies. One effective approach is offered through subtitler training frameworks. Okyayuz and Kaya (2016) propose a five-module subtitling training model that combines theory, technology use, formal subtitling skills (timing, condensation, error correction), and project-based evaluation, emphasizing competence-oriented outcomes.

Several studies in the Turkish literature have addressed issues of subtitling quality. Yılmaz (2017), in their MA thesis, examined the transfer of culture-specific elements in subtitles and highlighted the impact of translators' decisions on the preservation of cultural items. Metin Tekin (2019) analyzed the Turkish subtitled and dubbed versions of *Frozen*, evaluating translation strategies within the framework of equivalence theory and comparing qualitative and quantitative outcomes. Uslu Korkmaz and Okyayuz (2025) investigated interlingual subtitle quality in English–Turkish through Pedersen's (2017) FAR model, comparing professional and non-professional translators and drawing implications for AVT training in Türkiye. Also, Dalmış and Yıldız (2025) explored the quality of MT in literary texts and its suitability for post-editing. While not directly focused on automatic subtitle translation, their findings contribute indirectly by revealing the limitations of MT systems and emphasizing the continued necessity of human intervention for translation quality.

### **3. MT for Global Communication: Implications for Sustainability Discourse**

Communicating climate change effectively is vital because it is not only an environmental issue but also a profound social and ethical challenge. As Chadwick

(2017) notes, the climate crisis threatens global health, security, and equity, demanding urgent and coordinated communication strategies that translate scientific evidence into collective awareness and action. Similarly, Ballantyne (2016) emphasizes that effective climate change communication must bridge the gap between scientific complexity and public understanding to motivate behavioral and policy change. MT is promoted to achieve 'language justice' in sustainability discourse. Bowker (2020) asserts that MT "has significantly reduced the language barriers that impede effective communication," facilitating rapid dissemination of technical documents and educational materials. Bibliometric studies show that most of climate change research is published in English, which creates barriers for global audiences and highlights the role of MT in ensuring equitable access to sustainability knowledge (Haunschild, Bornmann, & Marx, 2016; Hulme, 2009). Yet, MT subtitle quality remains under scrutiny. Subtitling climate communication content presents distinct challenges that go beyond semantic transfer. Ashraf (2024), in his multimodal discourse analysis of two climate documentaries (*Kiss the Ground* and *Seaspiracy*), demonstrates that syntactic differences between English and Arabic complicate the translation of structures like adjective phrases and compounding. His findings reveal that subtitlers frequently resort to omission or modifications in cutting to preserve multimodal balance across verbal and visual modes, highlighting a key point: inadequate translation of syntactic nuances can undermine the accessibility and persuasiveness of climate communication. Qiu and Pym (2024) investigated the effects of MT errors on reception in audiovisual contexts, noting that semantic distortions and timing mismatches can erode viewer trust and comprehension. Du and Lu (2024) compared human-versus-machine subtitles and reported that MT systems exhibit higher error rates when handling specialized terminology, highlighting the need for quality assessment in critical domains. Moreover, survey evidence indicates that high confidence in AI reduces critical thinking and cognitive effort during generative tasks (Lee et al., 2025), suggesting translators may overlook subtle errors when over-relying on automation.

While MT and AI can speed global sustainability dialogues by cutting costs and time, they risk misrepresenting key concepts and cultural nuances. Advances in deep-transfer learning and automated quality estimation offer partial remedies (Yan et al., 2024), yet the tension between speed and precision remains. A hybrid MT-human workflow seems necessary to balance scalability with the semantic fidelity required in high-stakes sustainability communication. O'Brien's 'human-centered augmented translation' framework stresses that MT should empower, not replace, human expertise; AI interfaces must amplify translator agency through contextual suggestions and cognitive-load monitoring (O'Brien, 2023). Öztürk Baydere (2023) likewise calls for translator competence to emphasize adaptability, lifelong learning, and entrepreneurial skill, noting Türkiye's occupational standards underplay these capacities and risk passivity toward automation. By positioning translators as proactive agents integrating MT into value-added services, this approach aligns with sustainability communication, where human expertise must guide technology to ensure accuracy, nuance, and resilience amid accelerating change. Öner and Bengi (2025) found that when graduate

students used ChatGPT to draft Turkish user-manual sections, effective prompting, domain knowledge, and iterative revision determined usability. While AI excelled at efficiency and drafting, it faltered in standards compliance, symbol use, and multimodal integration—areas where human oversight proved indispensable. They argue prompting itself is a literacy: a communicative act reflecting user priorities and skill. Their study concludes AI's best role lies in augmentation, not automation—serving as a collaborator that extends human capability rather than replacing professional expertise. In practice, a hybrid MT-human workflow, guided by human-centered AI principles, is vital to scale multilingual sustainability communication without losing the nuance and accuracy climate-critical content demands.

#### **4. Pedersen's FAR Model**

The FAR model was introduced by Jan Pedersen in 2017 as a systematic tool for evaluating the quality of interlingual subtitles. The model is rooted in Pedersen's earlier theoretical concept of the "contract of illusion" between the subtitler and the viewer (Pedersen, 2017, pp. 46–47). According to this notion, a successful subtitle is one that is so accurate, fluent, and unobtrusive that it makes the viewer forget the presence of the original spoken language. Drawing on this premise, Pedersen developed the FAR model to enable more structured and targeted assessments of interlingual subtitle quality. The acronym FAR represents the model's three principals: functional equivalence, acceptability, and readability. It is a penalty-based error analysis model that incorporates a scoring system to help evaluators detect where and how improvements in subtitle quality can be made (Pedersen, 2017). Inspired by Romero-Fresco's (2011) NER model, originally designed for assessing intralingual subtitles, FAR applies a similar system of weighted penalties to errors, categorized as minor (0.25 points), standard (0.50 points), and serious (1 point), but semantic errors, which Pedersen (2017) assigns higher weights of 0.50, 1, and 2 points given their critical impact on meaning fidelity. These scores are summed to produce a quantitative index of quality, allowing for comparative evaluation across subtitle versions and timeframes.

The functional equivalence category is divided into semantic and stylistic errors. Semantic errors involve distortions or omissions altering the meaning and are thus the most severe type of error in the model. This category includes the failure to subtitle non-verbal but meaningful sounds like crying, shouting, or creaking, which are critical for narrative comprehension. These errors receive the highest penalty scores: minor semantic errors are scored at 0.50 points, standard at 1 point, and serious semantic errors at 2 points. Stylistic errors, in contrast, may not lead to complete incomprehensibility but result in the loss of nuances and register, affecting the pragmatic fidelity of the translation. The acceptability dimension encompasses errors related to the target linguistic system and norms. This includes grammar errors, where the subtitles violate syntactic rules; spelling errors, which may involve incorrect letters that hinder readability; and idiomaticity errors, in which unnatural collocations or inappropriately chosen expressions disturb the fluency and naturalness of the target language. Such errors result in subtitles that may feel linguistically foreign to the viewer, interrupting

their engagement with the audiovisual material. The third pillar, readability, assesses technical and ergonomic dimensions of subtitle delivery. It includes segmentation and synchronization errors that cause misalignment between spoken language and text, punctuation errors including incorrect or missing marks as well as failure to italicize off-screen sounds like telephones or ambient noises and reading speed and line-length issues. These issues may manifest as subtitles that pass too quickly to be read comfortably or that linger on the screen after the dialogue has finished (Pedersen, 2017). Unlike the NER model, which is based on words per second (wps), the FAR model adopts characters per second (cps) to ensure a more nuanced cross-linguistic readability measure, particularly suitable for assessing translations across structurally diverse languages. While Pedersen (2017) acknowledges limitations like the model's exclusively penalty-based structure, the absence of a reward mechanism, its emphasis on fluency over absolute accuracy, and the risk of subjective judgment in implementation, the FAR model remains a significant contribution to subtitle quality assessment. Its specificity for interlingual subtitles, rather than general translation or intralingual captioning, positions it as an essential tool in evaluating the increasingly widespread and fast-paced practice of automatic subtitle generation in contemporary audiovisual media.

Applications of the FAR model in AVT research have steadily expanded across diverse language pairs and content types. Early examples include Abdelaal (2019), who evaluated Arabic subtitles of *American Pie 2* and examined how cultural references were rendered in terms of functional equivalence, acceptability, and readability. Abdelaal and Al Sahrani (2021) analyzed taboo expressions in *Training Day*, finding that euphemized or omitted renderings often incurred acceptability penalties due to lost idiomaticity and register mismatch. Lassila (2019) applied FAR to Finnish subtitles of the Netflix series *Outlander*, highlighting segmentation and timing issues that impaired readability despite otherwise fluent translations. In the English–Indonesian context, Ningtias (2020) assessed *Shakespeare in Love*, finding that while general semantic fidelity was preserved, idiomatic and structural challenges persisted, particularly in acceptability and readability.

Yao (2022) evaluated machine-translated English–Chinese TED Talk subtitles and reported high error rates that compromised comprehension and engagement, underscoring the risks of relying solely on unedited MT in educational content. Li (2022) examined *The Good Wife*, showing that a hybrid MT plus post-editing workflow can yield time-efficient, high-quality results. Koglin et al. (2022) combined FAR with translator and audience assessments to evaluate post-edited interlingual subtitles, identifying persistent functional equivalence lapses even after human revision. Alaa and Al Sawi (2023) applied FAR to Egyptian-Arabic subtitles of *Feathers*, reporting generally high quality but occasional stylistic mismatches in culturally specific renderings. More recently, Ludera et al. (2024) found no significant overall score differences between professional and trainee subtitlers using FAR, though professionals exhibited superior condensation skills. Building on this line of inquiry, Zhou and Zhou (2024) integrated MQM with FAR to propose a “FAR 2.0” model applied to Chinese–English subtitle translations across four MT tools; their results showed Baidu outperforming

competitors—attributed to richer training data—yet persistent functional equivalence and readability challenges remain, underscoring the need for expanded corpora and closer human–machine collaboration.

Although various frameworks exist for evaluating machine-translated subtitles (e.g., FAR, NER, MQM), to the best of our knowledge, no published study has applied the FAR model or any comparable quality assessment framework to English–Turkish subtitles, particularly in the climate change and sustainability domain. Longitudinal analyses of MT subtitle quality remain rare, with most studies offering only single-point snapshots. (Hagström & Pedersen, 2022). Given the widespread use of automatic subtitles among Turkish-speaking viewers and the growing importance of audiovisual sustainability communication, this study conducts a longitudinal FAR-based evaluation of English–Turkish subtitles in IPCC videos from 2022 and 2025, providing domain-specific insights while contributing to translation studies and informing global sustainability communication.

## **5. Methodology**

### **5.1 Research Design**

This study adopts a longitudinal (repeated cross-sectional) comparative design to evaluate subtitle quality over time. The same two videos from IPCC’s YouTube were analyzed at two points: first in 2022 and again in 2025. This design allows for a direct comparison of subtitle translation quality at different stages of MT development. By revisiting the same video content three years apart, the research can observe whether improvements or regressions have occurred in subtitle accuracy, fluency, and readability. The longitudinal element allows assessment of both subtitle quality at a given time and its evolution. The comparative aspect ensures consistency in content, enabling the observed changes to be attributed to subtitle rendering system rather than to variations in source material.

### **5.2 Data Collection**

The dataset comprises the two most-watched videos on the IPCC’s YouTube channel, which had 45.9k subscribers at the time of collection. These videos were chosen for their popularity, scientific significance, and alignment with the IPCC’s mission: “to provide policymakers with regular scientific assessments on climate change, its implications and potential future risks, as well as to put forward adaptation and mitigation options” (*IPCC – Intergovernmental Panel on Climate Change*, n.d.). As a UN body of which the Republic of Türkiye is a member, the IPCC plays a pivotal role in shaping global awareness and sustainability discourse through comprehensive climate reports (<https://www.ipcc.ch/>). The channel was selected for its combination of scientific credibility, neutrality, and wide digital reach. As high-visibility materials communicating climate science to policymakers and the public, these videos provide a meaningful

dataset for examining how MT subtitles affect comprehension, accessibility, and the communication of climate knowledge across languages.

The first video, *Climate Change 2022: Impacts, Adaptation & Vulnerability* (released 7 June 2022), had 375,325 views in 2022 and 425,478 by 2025, underscoring its ongoing relevance. Based on the IPCC Sixth Assessment Report (AR6, Working Group II), it assesses climate impacts on ecosystems, biodiversity, and human communities, as well as vulnerabilities and adaptive capacities (Intergovernmental Panel on Climate Change [IPCC], n.d.). The 13:45-minute video is structured into three segments— “Climate Change Impacts,” “Tipping Points,” and “Maladaptation.” The second video, *Climate Change 2013: The Physical Science Basis* (released 21 November 2013), had about 302,000 views in 2022 and 307,871 by 2025. It summarizes the Fifth Assessment Report (AR5, Working Group I), providing “a comprehensive assessment of the physical science basis of climate change” (Intergovernmental Panel on Climate Change [IPCC], n.d.). At 9:19 minutes, it condenses findings on global warming, atmospheric change, and climatic drivers. Their inclusion ensures both a recent and a foundational IPCC report are represented, allowing analysis of subtitle quality across temporal, stylistic, and scientific contexts. Given the study’s focus, the dataset intentionally centers on a single platform and a single language pair (English–Turkish). While this controlled design supports internal consistency, it also means that the subtitling patterns identified here may not fully generalize across other genres, platforms, or language combinations. In addition, because YouTube’s proprietary MT engine was the sole source of the subtitles, the analysis reflects the behaviour of an MT system whose architecture and training data are not publicly disclosed, which inherently constrains cross-system comparability.

### 5.3 Data Analysis

Each subtitle was assessed for error frequency and penalty scores across Pedersen’s FAR dimensions (functional equivalence, acceptability, and readability). To ensure full comprehension of the content, every video was first watched in its original format without subtitles; the same videos were then viewed with YouTube’s automatic Turkish subtitles, with each successive viewing devoted to a single FAR dimension. All subtitles were annotated by the first coder, then independently reviewed by a second coder. The coders were professional translators trained in FAR coding. Any disagreement was discussed until consensus was reached; unresolved cases were settled by a third translator trained in FAR coding. This multi-coder structure ensured systematic cross-checking. The scope of the analysis was linguistic outputs rather than real-time viewer cognition, as the methodological framework did not incorporate audience-based reception tests.

Errors were quantified by category: functional equivalence (semantic and stylistic), acceptability (grammar, spelling, idiomaticity), and readability (segmentation & spotting, punctuation & reading speed, line-length) issues. Following the scoring system in the FAR model, minor, standard, and serious errors were assigned 0.25, 0.50, and 1 point, respectively. An exception applies to semantic errors, which Pedersen (2017) weights more heavily, ranging from 0.50 up to 2 points, due to their greater impact on meaning fidelity. Functional equivalence, acceptability, and readability errors were

totalled and normalized according to the length of videos, allowing for consistent comparisons, and summary statistics (total errors, mean penalty points per minute, and overall quality scores) were generated for the 2022 and 2025 datasets to enable longitudinal comparison. While Pedersen’s FAR model offers a robust tool for identifying linguistic and functional issues in subtitles, it does not capture dimensions such as viewer comprehension, emotional engagement, or usability in real-time contexts. Thus, the analysis presented here reflects textual quality indicators rather than experiential audience responses, which were beyond the methodological scope of this study.

## 6. Results and Discussion

Error scores declined substantially between 2022 and 2025 across both IPCC videos, reflecting improvements in machine-translated subtitle quality. In Table 1, the first video (*Climate Change 2022: Impacts, Adaptation & Vulnerability*) saw total scores fall from 207.00 to 76.50, a 63.04% reduction. The second video (*Climate Change 2013: The Physical Science Basis*) improved from 95.50 to 44.50, a 53.40% reduction. As it is demonstrated in Table 1, functional equivalence caused most penalties in both videos across both years. In Video 1, the score for functional equivalence decreased from 86.50 in 2022 to 40.50 in 2025, while acceptability dropped from 39.50 to 14.75 and readability from 32.50 to 8.75. These results indicate that although functional equivalence remained the most error-prone dimension, all categories improved between 2022 and 2025.

**Table 1**

*Total Error Scores of Automatic Subtitles in IPCC Videos (2022 vs. 2025)*

Video	Year	Functional Equivalence	Acceptability	Readability	Total Score
Video 1	2022	86.50	39.50	81.00	207.00
Video 1	2025	40.50	14.75	21.25	76.50
Video 2	2022	43.00	20.00	32.50	95.50
Video 2	2025	31.00	4.75	8.75	44.50

**Note.** Functional equivalence includes semantic and stylistic errors; acceptability includes grammar, spelling, and idiomaticity errors; readability includes segmentation & spotting, punctuation & reading speed, line length errors.

Percentage changes in Table 2 underscore these gains. Video 1 recorded reductions of 53.18% in functional equivalence, 62.66% in acceptability, and 73.77% in readability. Video 2 showed smaller, though still considerable, reductions in functional equivalence (27.91%), alongside sharp declines in acceptability (76.25%), and readability (73.08%). These figures indicate that semantic fidelity remains the most challenging category, while grammatical, spelling, and technical formatting improved faster.

**Table 2**

Percentage Change in Error Scores of Automatic Subtitles in IPCC Videos (2022–2025)

Video	Functional Equivalence	Acceptability	Readability	Total Score
Video 1	-53.18%	-62.66%	-73.77%	-63.04%
Video 2	-27.91%	-76.25%	-73.08%	-53.40%

**Note.** Error counts are raw numbers; penalty scores reflect Pedersen’s weighted system. Percentage change is calculated as  $(2025 - 2022) / 2022 \times 100$ . Negative values indicate error reduction, showing improvement in subtitle quality between 2022 and 2025.

Sub-category breakdowns clarify these patterns. For Video 1 (Table 3), semantic errors decreased from 82.00 to 36.00 points, spelling errors from 22.00 to 1.25, and segmentation & spotting from 31.00 to 7.75. Grammar, punctuation & reading speed and line length issues also showed declines, though minor idiomaticity errors remained stable or slightly increased, indicating persistent challenges in collocation and natural phrasing.

**Table 3**

Sub-Category Error Scores of Automatic Subtitles in Video 1 (2022 vs. 2025)

Category	Error Level	2022 Score	2025 Score	Change
<b>Functional Equivalence – Semantic</b>	Serious	82.00	36.00	-46.00
	Standard	4.50	4.50	0.00
	Minor	–	–	–
<b>Functional Equivalence – Stylistic</b>	Serious	–	–	–
	Standard	0.50	0.50	0.00
	Minor	–	–	–
<b>Acceptability – Grammar</b>	Serious	8.00	4.00	-4.00
	Standard	2.00	1.00	-1.00

*Assessing Quality in Machine-Translated English-Turkish Climate Communication Subtitles: A Longitudinal Study*

Category	Error Level	2022 Score	2025 Score	Change
	Minor	1.50	2.25	+0.75
<b>Acceptability – Spelling</b>	Serious	1.00	–	–1.00
	Standard	2.50	1.00	–1.50
	Minor	18.50	0.25	–18.25
<b>Acceptability – Idiomaticity</b>	Serious	1.00	1.00	0.00
	Standard	2.50	2.00	–0.50
	Minor	2.50	3.25	+0.75
<b>Readability – Punctuation &amp; Reading Speed</b>	Serious	8.00	6.00	–2.00
	Standard	5.50	1.00	–4.50
	Minor	23.50	3.00	–20.50
<b>Readability – Segmentation &amp; Spotting</b>	Serious	28.00	7.00	–21.00
	Standard	3.00	0.50	–2.50
	Minor	–	0.25	+0.25
<b>Readability – Line Length</b>	Serious	13.00	3.00	–10.00
	Standard	–	0.50	+0.50
	Minor	–	–	–
<b>Total</b>	–	<b>207.00</b>	<b>76.50</b>	<b>–130.50</b>

**Note.** Error counts are raw numbers; penalty scores reflect Pedersen’s weighted system (0.25–1 points, except for semantic errors; 0.50-2 points per error). Percentage change is calculated as  $(2025 - 2022) / 2022 \times 100$ .

Video 2 results (Table 4) highlight comparable trajectories. Severe semantic penalties decreased from 42.00 to 28.00, spelling errors from 12.00 to 0.25, and

punctuation & reading speed from 15.25 to 2.50. Segmentation & spotting errors declined by 9 points. Idiomaticity errors improved little, and functional equivalence penalties slightly increased, illustrating that while gross mistranslations diminished, finer semantic nuances remain unresolved.

**Table 4**

*Sub-Category Error Scores of Automatic Subtitles in Video 2 (2022 vs. 2025)*

<b>Category</b>	<b>Error Level</b>	<b>2022 Score</b>	<b>2025 Score</b>	<b>Change</b>
<b>Functional Equivalence - Semantic</b>	Serious	42.00	28.00	-14.00
	Standard	1.00	3.00	+2.00
	Minor	–	–	–
<b>Functional Equivalence - Stylistic</b>	–	–	–	–
<b>Acceptability – Grammar</b>	Serious	3.00	–	-3.00
	Standard	1.00	1.00	0.00
	Minor	0.75	1.25	+0.50
<b>Acceptability –Spelling</b>	Serious	1.00	–	-1.00
	Standard	0.50	–	-0.50
	Minor	10.50	0.25	-10.25
<b>Acceptability –Idiomaticity</b>	Serious	1.00	1.00	0.00
	Standard	1.00	1.50	+0.50
	Minor	1.25	0.25	-1.00
<b>Readability – Punctuation &amp; Reading Speed</b>	Serious	1.00	–	-1.00
	Standard	2.50	–	-2.50

Category	Error Level	2022 Score	2025 Score	Change
	Minor	11.75	2.50	-9.25
<b>Readability – Segmentation &amp; Spotting</b>	Serious	12.00	5.00	-7.00
	Standard	2.50	0.50	-2.00
	Minor	0.25	0.25	0.00
<b>Readability – Line Length</b>	Serious	2.00	–	-2.00
	Standard	0.50	0.50	0.00
	Minor	–	–	–
<b>Total</b>	–	<b>95.50</b>	<b>44.50</b>	<b>-51.00</b>

**Note.** Error counts are raw numbers; penalty scores reflect Pedersen’s weighted system (0.25–1 points, expect for semantic errors; 0.50–2 points per error). Percentage change is calculated as  $(2025 - 2022) / 2022 \times 100$ .

A layered analysis of the MT output reveals distinct, recurring problem types that progress differently between 2022 and 2025. First, semantic errors are the most consequential because they alter or remove content. For example, the full source sentence reads: “This report really brings us a new message though, that those two things are not separate.” The clause in the source sentence “that those two things are not separate” is absent from the 2022 output and remains missing in the 2025 rendering “Bu rapor bize gerçekten yeni bir mesaj getiriyor: (back-translation: This report really brings us a new message:),” which constitutes a serious semantic omission because the reported relation between the two phenomena is not transmitted to the target audience. Similarly, the source sentence “Each of the last three decades has been warmer than all preceding decades since 1850” was rendered in both years as “Son otuz yılın her biri, 1850’den beri önceki tüm on yıllardan daha sıcaktı, (back-translation: Each of the last thirty years has been warmer than all previous decades since 1850),” which misrepresents three decades as thirty individual years and thus distorts the fact. A lexical mistranslation in Video 2 shows the same risk: the 2022 output rendered “glaciers” as “camcılar,” literally means glassmakers, while the 2025 correction to “Dünya çapındaki buz tabakaları ve camı yapılar kütle kaybediyor (back-translation: Ice sheets and glass-like structures worldwide are losing mass)” recovers part of the referent but fails to provide an authoritative, accurate scientific term, leaving semantic fidelity unattained. Likewise, the mistranslation of “maladaptation” as “erkek adaptasyonu (back-translation: male adaptation)” persisted across both datasets, demonstrating the limits

of MT and signaling a risk for domain precision. A parallel case is the rendering of “coldest” as “CEST” in both 2022 and 2025 outputs, where the model mistakenly inserted the time-zone acronym instead of the intended adjective, resulting in a serious semantic error that entirely obscures the source meaning. A similar failure was observed in the first analysis of the phrase “We have three key messages,” which was rendered as “stratejilerimiz var; (back-translation: we have strategies;),” a serious semantic loss that altered the communicative intention of the speaker. In the 2025 output this speech recognition problem was corrected, yielding the accurate version “üç temel mesajımız var:(back translation: we have three key messages:)”Such errors are best explained as consequences of weaknesses in MT speech recognition, where acoustically similar strings are misidentified and mistranslated, producing distortions that undermine domain accuracy and communicative clarity. Likewise, the source phrase “The 20th century as a whole, the dominant contributions are ...” becomes “Bir bütün olarak baskın katkılar ... (back-translation: As a whole, the dominant contributions are)” in 2025. The Turkish rendering omits the historical frame “the 20th century,” so the reader is left without this frame of reference that is central to the original meaning. In addition, the phrase “baskın katkılar” is a word-for-word rendering of “dominant contributions” that sounds awkward in Turkish discourse and fails to convey the intended nuance. For these reasons, the output constitutes a serious semantic mistranslation. In the 2022 version, the same passage appears with an omission, further compounding the semantic loss. By contrast, some items show semantic recovery, like the 2022 fragment “besleyici gıda içeriyor mu ... temiz içme suyu sağlama yeteneğimizi azaltır, (back-translation: does it contain nutritious food... reduces our ability to provide clean drinking water)” which in 2025 appears as “besleyici gıda yetiştirme ve temiz içme suyu sağlama yeteneğimizi azaltır, (back-translation: reducing our ability to grow nutritious food or provide clean drinking water),” aligning more closely with the source “reducing our ability to grow nutritious food or provide clean drinking water.” Also, in both datasets, passages delivered in a second language were left untranslated (Spanish in the first video and Chinese in the second) representing serious semantic errors, as entire stretches of meaning were lost and rendered inaccessible to the target audience. In the first video, the sentence “This report is important because I think it really underscores the fact that the scientific evidence is now unequivocal” was rendered in both years as “Bu rapor önemlidir çünkü bilimsel kanıtların artık kesin olduğu gerçeğini gerçekten vurguluyor. (back-translation: This report is important because it really underscores the fact that the scientific evidence is now unequivocal)” While the translation preserves the main propositional content, it omits the phrase “I think,” which signals the speaker’s stance. In addition, the use of the copular suffix “-dir” creates a stylistic shift. While the source text reflects spoken, interactive delivery, the Turkish rendering takes on a formal, written-like tone. This transforms the communicative mode from spoken to academic, thereby neutralizing the immediacy and conversational character of the original utterance. According to Pedersen (2017), stylistic errors occur when features of register, tone, or idiomaticity are not adequately conveyed, reducing pragmatic alignment between source and target texts. In this case, the stylistic mismatch results not only from omission but also from tonal elevation through morphological marking. By contrast, no stylistic errors were identified in either year’s analysis of the second video, likely due to

its mainly informative nature, with relatively less reliance on personal stance or rhetorical nuance.

Second, acceptability errors concern the naturalness of the target-language rendering. Under grammar, the ecosystem sequence “From high mountain ecosystems to the deep ocean, from tropical coral reefs all the way to Arctic ice driven ecosystems” was produced in 2025 as “Yüksek dağ ekosistemlerinden derin okyanuslara, tropikal mercan resiflerinden Arktik buzla kaplı ekosistemlere kadar. (back-translation: From high mountain ecosystems to the deep oceans, from tropical coral reefs to Arctic ice-covered ecosystems.)” The phrase “Arktik buzla kaplı ekosistemlere (back-translation: to Arctic ice-covered ecosystems)” displays syntactic awkwardness that disrupts standard Turkish phrase structure and was scored as a noticeable grammar error because it reduces immediate clarity. Another grammatical pattern appears in existential constructions: “There are four major contributors to sea level rise” was rendered as “Deniz seviyesinin yükselmesine dört ana katkıda bulunan faktör vardır (back-translation: There are factor making four main contributions to sea level rise),” an ordering that reads as an inversion in Turkish and is classified as a standard grammar error since it is recognizable yet nonstandard. Under idiomaticity, several literal collocations and nonnative word choices persist. For example, the sentence “We have looked at all the evidences that tell us how the climate has changed in the past and presently took that evidence...” was translated in 2022 as “iklimin geçmişte nasıl değiştiğini anlatan tüm kanıtlara baktık ve şu anda bu kanıtları alıyoruz.... (back-translation: we looked at all the evidence that tells how the climate has changed in the past, and now we are taking this evidence...)” The direct rendering of “took” as “alıyoruz” produces an unnatural and inconsistent phrasing in Turkish, reducing fluency. By 2025, the output improved to “iklimin geçmişte nasıl değiştiğini anlatan tüm kanıtlara baktık ve şu anda bu kanıtları kullanarak (back-translation: We looked at all the evidence showing how the climate has changed in the past, and by using this evidence now),” which is more idiomatic and pragmatically aligned with Turkish usage. This example illustrates that while some idiomatic refinements emerged over time, literal transfer patterns continued to surface, showing that idiomaticity remains a persistent weakness in MT outputs. The 2025 rendering of “It is clear that climate change is impacting on the wellbeing of human societies, but also on the wellbeing of our planet” appears as “iklim değişikliğinin insan toplumlarının refahını ... etkilediği açıktır (back-translation: It is clear that climate change affects the wellbeing of human societies),” where the literal translation “insan toplumları (back-translation: human societies)” sounds odd in Turkish and was assessed as a minor idiomaticity error; the same clause was absent in 2022, producing a semantic loss. The phrase “This report really brings us a new message” is expressed as “Bu rapor bize bir mesaj getiriyor (back-translation: This report brings us a message)” a verb-noun collocation that is not natural in Turkish and thus incurs an idiomaticity penalty. In 2022 analysis, the sequence “yaygın ve yaygın (back-translation: widespread and widespread),” “widespread and pervasive” in source video, was successfully corrected to “yaygın ve nüfuz eden (back-translation: widespread and pervasive),” which removed redundancy and partially restored nuance, yet its initial appearance signals persistent

register control problems in 2025 analysis. In the phrase “The 20<sup>th</sup> century as a whole, the dominant contributions are...,” the 2025 subtitle rendering “Bir bütün olarak baskın katkılar ... (back-translation: As a whole, the dominant contributions)” uses the collocation “baskın katkılar (back-translation: dominant contributions),” which reads as an unusual scientific collocation in Turkish and is therefore counted as an idiomaticity error. In 2022, additional spelling and punctuation mistakes increased acceptability burdens; for instance, the ecosystem sentence combined syntactic awkwardness with spelling and punctuation errors, raising the post-editing workload.

Third, readability problems concern punctuation, connective choice, and general fluency, and they affect processing speed and real-time comprehension. Punctuation omissions and incorrect segmentation in the 2022 material produced lower readability scores and made human recovery of lost meaning more difficult, as when the omitted clause about the non-separability of two phenomena was accompanied by segmentation problems that impeded clause reconstruction. Another set of readability problems arose from surface-level formatting. For example, the source text “These are spread across parts of Africa, South Asia, South and Central America, small islands and the Arctic” appeared in the 2022 subtitles without capitalization, “afrika, güney asya, orta amerika, arktik (back-translation: africa, south asia, central america, arctic),” whereas in 2025 the proper forms were restored as “Afrika, Güney Asya, Orta Amerika, Arktik (back-translation: Africa, South Asia, Central America, Arctic),” improving clarity and aligning with standard orthography. Conversely, punctuation problems remained unresolved. In the sequence “Ocean heat uptake, melting of glaciers, reduction of ice sheets, and changes in water storage on land,” both versions failed to use commas, producing dense clusters of terms, “Okyanus ısı alımı buzulların erimesi buz tabakalarının azalması ve karadaki su depolanmasındaki değişiklikler (back-translation: Ocean heat uptake melting of glaciers reduction of ice sheets and changes in water storage on land),” which hindered fluency and increased the reader’s cognitive burden. These persistent punctuation issues demonstrate that while surface repairs like capitalization improved between 2022 and 2025, the readability dimension continues to be constrained by inadequate clause separation and connective marking.

Taken together, these layered examples demonstrate that the MT system achieved measurable semantic and grammatical recovery between 2022 and 2025 in the translation of frequent and simple propositions. However, the system continues to produce serious semantic omissions, inaccurate quantification, recurrent nonstandard grammatical constructions, idiomatic collocations that sound nonnative, and punctuation or connective choices that slow comprehension. Recurrent patterns are particularly visible in the mistranslation of terminology, the imperfect handling of domain-specific language, rushed display speeds, and disruptions caused by punctuation. These findings are consistent with Pedersen’s (2017) argument that semantic errors undermine the coherence and credibility of the target text and are “so serious as to disturb the contract of illusion” (p. 27), which explains why they received

the highest penalty scores in this study. Although some gains in readability are evident, unresolved idiomaticity errors highlight the ongoing gap.

These results resonate with broader MT developments, particularly in handling specialized vocabulary, domain adaptation, and textual consistency (Yan et al., 2024). Between 2022 and 2025, progress was evident in the ability of systems to follow context more reliably, adapt to subject-specific content, and reduce errors that previously disrupted fluency (Yan et al., 2024; Rivera-Trigueros, 2021). Improvements in semantic and grammatical dimensions mirror these innovations, while persistent idiomaticity challenges illustrate the limits of statistical alignment without deep cultural modeling. These results imply that while semantic fidelity can benefit from large-scale learning, the preservation of pragmatic nuance requires deeper linguistic and cultural integration, highlighting a key limitation in current MT architectures. This aligns with Alimen, Öner Bulut and Karadağ's (2023) argument that AI and MT are not merely tools for efficiency but transformative forces that demand a reconceptualization of translation competence. They emphasize that translators must cultivate critical awareness toward AI outputs while safeguarding cultural and linguistic specificity, since over-reliance on machine-generated text risks eroding the very human dimensions of meaning-making. In sustainability communication, this means that while MT can scale accessibility, it must be balanced with human oversight to ensure nuance, context, and inclusivity remain intact (Alimen et al., 2023).

## **7. Conclusion**

Analysis of the two IPCC videos from 2022 and 2025 shows notable improvements in subtitle quality, with error counts falling from 207.00 to 76.50 (63.04% reduction) in the first video and from 95.50 to 44.50 (53.40% reduction) in the second. Progress, however, was uneven: acceptability and readability recorded the largest gains (62.66%–76.25%), while functional equivalence improved least (53.18% and 27.91%) and remained the main source of penalties in 2025, underscoring semantic accuracy as the key challenge. Qualitative examples confirm these trends, showing improved term translation, capitalization, sentence structure, and fewer punctuation errors. Yet despite these advances, persistent problems continue to limit sustainability communication. Automated subtitles expand access, reduce costs, and widen reach, but they risk miscommunication through semantic omissions, mistranslated scientific terms, poor handling of climate-specific language, inadequate rendering of cultural and idiomatic expressions, rushed display speeds, and residual punctuation-related issues.

Yet these risks are not only linguistic but also strategic, as subtitle accuracy directly shapes how climate information is received and acted upon. In this context, the mandate to inform decision-makers depends on translation accuracy, because subtitle errors can undermine public comprehension and policy responses. Clear and accurate subtitles foster broader participation in sustainability dialogues, whereas mistranslations or synchronization failures can erode trust, delay understanding, and distort adaptation strategies in non-English contexts. Therefore, for scientific subtitles in contexts similar to

those addressed in this study, post-editing should be prioritized with the following aims: to repair semantic omissions and inaccurate quantification, to rework grammatical structure and replace literal collocations with idiomatic equivalents, and to restore punctuation and connective choices that maximize readability for a Turkish-speaking audience. Strategically, these findings suggest that subtitle workflows should first focus on correcting semantic errors and improving idiomatic expression, while also coordinating multilingual releases with major policy moments to ensure that climate messages reach audiences when they matter most.

The study demonstrates that Pedersen’s FAR model can track changes in subtitle quality over time, confirming its value for specialized domains. It situates these changes within the evolution of MT, linking technical progress to linguistic and cultural effects. In practice, the findings support hybrid workflows and targeted refinements. A human-centred AI paradigm that enhances translator agency through contextual suggestions and adaptive segmentation is advocated; O’Brien (2023) calls for systems that move beyond static, translator-centered frameworks to dynamic, sensor-informed environments that monitor cognitive load and adjust segmentation and terminology in real time. Such approaches balance efficiency with idiomatic accuracy, natural style, and readability. Promoting MT literacy is equally crucial, since high audience confidence in AI outputs without critical evaluation allows subtle inaccuracies to persist (Bowker & Ciro, 2019; Lee et al., 2025). Besides its integration into translator and subtitler training, this literacy should also be extended as a public competency that goes beyond translation and language programs. For non-professional users—students, journalists, civil servants, NGO staff, educators, and platform moderators—in the context of automatic subtitle translation, MT literacy entails (i) recognizing machine-generated captions, (ii) spotting common error types (e.g., semantic omissions, misquantification, literal idiomatic renderings, segmentation/timing issues), and (iii) applying simple mitigations (e.g., pause/rewatch, triangulate with on-screen visuals or official transcripts, consult glossaries, flag low-quality captions). To scale these skills, we recommend embedding MT-literacy modules in secondary and undergraduate media and science communication curricula; offering certificate courses in fields like environmental studies, public policy, teacher education, and journalism; and organizing targeted workshops for public agencies and NGOs. Platform-level supports—clear auto-caption labels, confidence/coverage notices, and expert review channels—can reinforce these efforts. Treating translation/MT literacy as civic infrastructure strengthens the quality and impact of automated access, extending benefits to audiences most dependent on auto-generated subtitles.

The study has some limitations. Its scope is limited to two IPCC videos and one language pair, which constrains generalizability across genres, platforms, and languages. Reliance on YouTube’s MT engine narrows comparability with other subtitle systems. The FAR model’s penalty-based scoring cannot fully capture viewer comprehension or emotional engagement, and no reception studies were conducted to measure how audiences perceive subtitle quality in real time. Future research should expand the dataset to include additional languages, genres, and platforms; integrate reception

methods, such as surveys, comprehension tests, or eye-tracking; and experiment with hybrid workflows that combine FAR-based error detection with targeted post-editing. Assessing the effect of subtitle accuracy on climate literacy, policy understanding, and local adaptation strategies would further clarify translation's role in sustainability communication.

In conclusion, MT delivers measurable accessibility and reach gains, yet unresolved semantic, idiomatic, and technical problems continue to undermine clarity and trust in climate communication. A human-centred approach to MT, pairing adaptive automation with human oversight, remains essential to ensure that sustainability messages are both widely disseminated and faithfully conveyed.

## Acknowledgements

This study was supported by the Scientific and Technological Research Council of Türkiye (TÜBİTAK) within the scope of "2209-A- Research Project Support Programme for Undergraduate Students" under the project number 1919B012223587.

---

### Author Contributions

**First Author:** Rumeysa Nur Çakmak 50% (literature review, data collection and analysis, theoretical framework, discussion)

**Second Author:** Muhammed Baydere 50% (literature review, data analysis, theoretical framework, discussion)

**Conflict of Interest:** The authors declare no conflict of interest.

**Financial Support:** This study was supported by TÜBİTAK (The Scientific and Technological Research Council of Türkiye) under project number 1919B012223587.

### Yazar Katkıları

**Birinci Yazar:** Rumeysa Nur Çakmak %50 (literatür taraması, veri toplama ve analizi, kuramsal çerçeve, tartışma)

**İkinci Yazar:** Muhammed Baydere %50 (literatür taraması, veri analizi, kuramsal çerçeve, tartışma)

**Çıkar Çatışması:** Yazarlar çıkar çatışması bildirmemiştir.

**Finansal Destek:** Bu çalışma, Türkiye Bilimsel ve Teknolojik Araştırma Kurumu (TÜBİTAK) tarafından 1919B012223587 numaralı proje kapsamında desteklenmiştir.

---

## References

- Abdelaal, N. M. (2019). Subtitling of culture-bound terms: Strategies and quality assessment. *Heliyon*, 5(4). <https://doi.org/10.1016/j.heliyon.2019.e01532>
- Abdelaal, N. M., & Al Sarhani, A. (2021). Subtitling strategies of swear words and taboo expressions in the movie *Training Day*. *Heliyon*, 7(7). <https://doi.org/10.1016/j.heliyon.2021.e07583>
- Alaa, A. M., & Al Sawi, I. (2023). The analysis and quality assessment of translation strategies in subtitling culturally specific references: *Feathers*. *Arab World English Journal*, 14(2), 155–173. <https://doi.org/10.24093/awej/vol14no2.10>

- Alimen, N., Öner Bulut, S., & Karadağ, A. B. (2023). Yapay zekâ, dil ve çeviri. [AI, Language, and Translation.] In B. Küçükcan & B. F. Yıldırım (Eds.), *Yapay zekâ: Disiplinlerarası yaklaşımlar* [AI: Interdisciplinary Approaches] (pp. 81–103). Vakıfbank Kültür.
- Armstrong, S., Caffrey, C., & Flanagan, M. (2006). Translating DVD subtitles from English–German and English–Japanese using example-based machine translation. In *MuTra 2006 – Audiovisual Translation Scenarios: Conference Proceedings* (pp. 1–10). [http://www.euroconferences.info/proceedings/2006\\_Proceedings/2006\\_Armstrong\\_Caffrey\\_Flanagan.pdf](http://www.euroconferences.info/proceedings/2006_Proceedings/2006_Armstrong_Caffrey_Flanagan.pdf)
- Ashraf, M. (2024). Subtitling climate change: A multimodal discourse analysis study of key structural differences in the subtitles of two climate documentaries on Netflix. *The International Journal of Communication and Linguistic Studies*, 23(1), 163–179. <https://doi.org/10.18848/2327-7882/CGP/v23i01/163-179>
- Ballantyne, A. G. (2016). Climate change communication: What can we learn from communication theory? *Wiley Interdisciplinary Reviews: Climate Change*, 7(3), 329–344. <https://doi.org/10.1002/wcc.392>
- Bowker, L. (2020). Machine translation literacy instruction for international business students and business English instructors. *Journal of Business & Finance Librarianship*, 25(1–2), 25–43. <https://doi.org/10.1080/08963568.2020.1794739>
- Bowker, L., & Buitrago-Ciro, J. (2019). *Machine translation and global research: Towards improved machine translation literacy in the scholarly community*. Emerald Publishing.
- Chadwick, A. E. (2017). Climate change communication. In D. L. Moyer (Ed.), *Oxford Research Encyclopedia of Communication*. Oxford University Press. <https://doi.org/10.1093/acrefore/9780190228613.013.22>
- Dalmış, E., & Yıldız, M. (2025). The limits of machine translation: A quality analysis of machine translation in literary text. *Uluslararası Dil, Edebiyat ve Kültür Araştırmaları Dergisi*, 8(1), 314–326. <https://doi.org/10.37999/udekad.1627217>
- DataReportal. (2024). *Digital 2024: Global overview report*. <https://datareportal.com/reports/digital-2024-global-overview>
- Díaz Cintas, J. (2001). *Striving for quality in subtitling: The role of a good dialogue list*. Multi(media) Translation Publishing House.
- Díaz Cintas, J. (2013). *Subtitling: Theory, practice and research*. In C. Millán & F. Bartrina (Eds.), *The Routledge handbook of translation studies* (pp. 273–287). Routledge.
- Díaz Cintas, J. (2019). The technological turn in audiovisual translation. *International Journal of Translation and Interpreting Research*, 11(1), 20–37.
- Du, J., & Lu, J. (2024). A comparative study on the translation quality between human and machine-generated subtitles. *Translation & Interpreting Studies*. Advance online publication. <https://doi.org/10.1075/tis.22008.du>
- Gottlieb, H. (2001). Texts, translation and subtitling—in theory, and in Denmark. In G. Anderman & M. Rogers (Eds.), *Translators and translations* (pp. 149–192). Multilingual Matters.
- Hagström, H., & Pedersen, J. (2022). Subtitles in the 2020s: The influence of machine translation. *Journal of Audiovisual Translation*, 5(1), 110–132.

- Hauschild, R., Bornmann, L., & Marx, W. (2016). Climate change research in view of bibliometrics. *PLOS ONE*, 11(7). <https://doi.org/10.1371/journal.pone.0160393>
- Hulme, M. (2009). *Why we disagree about climate change: Understanding controversy, inaction and opportunity*. Cambridge University Press.
- Intergovernmental Panel on Climate Change. (2013, September 27). *Climate change 2013: The physical science basis* [Video]. YouTube. <https://www.youtube.com/watch?v=6yiTZm0y1YA> (Accessed: 12 January 2025)
- Intergovernmental Panel on Climate Change. (2022, February 28). *Climate change 2022: Impacts, adaptation & vulnerability* [Video]. YouTube. <https://www.youtube.com/watch?v=SDRxfuEvqGg&t=160s> (Accessed: 12 January 2025)
- Intergovernmental Panel on Climate Change. (2023). *AR6 synthesis report: Climate change 2023*. <https://www.ipcc.ch/report/ar6/syr/>
- IPCC — Intergovernmental Panel on Climate Change. (n.d.). IPCC. <https://www.ipcc.ch/>
- Jin, H., & Yuan, Z. (2023). The application of machine translation in automatic dubbing in China: A case study of the feature film *Mulan*. *Translation Horizons*, 16(1), 45–59.
- Karamitroglou, F. (1998). A proposed set of subtitling standards in Europe. *Translation Journal*, 2(2), 1–15.
- Kemp, S. (2025, March 23). *Digital 2024: Global Overview Report — DataReportal — Global Digital Insights*. DataReportal — Global Digital Insights. <https://datareportal.com/reports/digital-2024-global-overview-report>
- Koglin, A., da Silveira, J. G. P., de Matos, M. A., da Costa Silva, V. T., & Candido Moura, W. H. (2022). Quality of post-edited interlingual subtitling: FAR model, translator’s assessment and audience reception. *Translation Spaces*, 11(2), 275–299.
- Lassila, H. (2019). *“Dinna Fash, Sassenach”: Subtitle quality and Finnish translations of cultural references in streaming services*. [Unpublished Master’s thesis], University of Turku. <https://www.utupub.fi/handle/10024/147456>
- Lee, H. P. H., Sarkar, A., Tankelevitch, L., Drosos, I., Rintel, S., Banks, R., & Wilson, N. (2025). The impact of generative AI on critical thinking: Self-reported reductions in cognitive effort and confidence effects from a survey of knowledge workers. *CHI*. <https://doi.org/10.1145/3706598.3713778>
- Li, M. (2024). A report on post-editing of machine translation of Climate Change 2023. *Region — Educational Research and Reviews*, 6(3). <https://doi.org/10.32629/rerr.v6i3.1862>
- Li, S. (2022). Quality assessment of machine-translated subtitles based on the FAR model: Exemplified by *The Good Wife*. *Journal of Audiovisual Translation Studies*, 4(1), 92–110.
- Lu, T. (2023). Applying machine translation to Chinese–English subtitling: Constraints and challenges. *Linguistica Antverpiensia: New Series — Themes in Translation Studies*, 22, 95–113. <https://doi.org/10.52034/lans-tts.v22i.752>
- Ludera, E., Szarkowska, A., & Orrego-Carmona, D. (2024). Expertise in interlingual subtitling: Applying the FAR model to study the quality of subtitles created by professional and trainee subtitlers. *Translation, Cognition & Behavior*, 7(1), 86–112.

- Metin Tekin, B. (2019). An assessment of audio-visual translation strategies in dubbing and subtitling of Frozen into Turkish. *Rumelide Dil ve Edebiyat Araştırmaları Dergisi*, (15), 342–352. <https://doi.org/10.29000/rumelide.580635>
- Ningias, N. A. (2020). *Subtitling strategies and qualities of flirting styles in Shakespeare's in Love movie*. [Unpublished undergraduate thesis], State Islamic Institute of Surakarta].
- O'Brien, S. (2023). Human-centered augmented translation: against antagonistic dualisms. *Perspectives*, 32–32(3), 391–406. <https://doi.org/10.1080/0907676X.2023.2247423>
- Okyayuz, A. Ş., & Kaya, M. (2016). A model for subtitling training. *International Journal of Languages' Education and Teaching*, 4(2), 257–275. <https://doi.org/10.18298/ijlet.601>
- Öner, I., & Bengi, Z. B. (2025). Prompt, evaluate, revise: AI-assisted user manual creation in academic practice. *transLogos Translation Studies Journal*, 8(1), 103–347. <https://doi.org/10.29228/transLogos.78>
- Öztürk Baydere, H. (2023). Rethinking translator competence in the machine translation era. *transLogos Translation Studies Journal*, 6(1), 45–74. <https://doi.org/10.29228/transLogos.54>
- Pedersen, J. (2017). The FAR model: Assessing quality in interlingual subtitling. *The Journal of Specialised Translation*, 28, 210–229.
- Popowich, F., McFetridge, P., Turcato, D., & Toole, J. (2000). Machine translation of closed captions. *Machine Translation*, 15(4), 311–341. <https://doi.org/10.1023/A:1026724505180>
- Qiu, J., & Pym, A. (2024). Fatal flaws? Investigating the effects of machine translation errors on audience reception in the audiovisual context. *Translation and Interpreting Studies*. Advance online publication. <https://doi.org/10.1075/tis.21012.qiu>
- Rivera-Trigueros, I. (2021). Machine translation systems and quality assessment: A systematic review. *Language Resources and Evaluation*, 56(2), 593–619. <https://doi.org/10.1007/s10579-021-09537-5>
- Romero-Fresco, P. (2011). *Subtitling through speech recognition: Respeaking*. St Jerome.
- Sixth assessment report — IPCC*. (n.d.). IPCC. <https://www.ipcc.ch/assessment-report/ar6/>
- Sousa, V. D. (2011). Post-editing quality in SMT subtitle translations: An analysis using BLEU. In *Proceedings of the 7th Workshop on Statistical Machine Translation* (pp. 32–39). Association for Computational Linguistics.
- State of the global climate 2024*. (2025, May 9). World Meteorological Organization. <https://wmo.int/publication-series/state-of-global-climate-2024>
- Uslu Korkmaz, A., & Okyayuz, A. Ş. (2025). Audiovisual translation training in Turkey from the perspective of interlingual subtitling quality: Findings and suggestions. *Çeviribilim ve Uygulamaları Dergisi*, (TÜÇEVAS Special Issue), 109–132. <https://doi.org/10.37599/ceviri.1614516>
- World Meteorological Organization. (2025). *State of the global climate 2025*. WMO. <https://public.wmo.int/en/resources/library/state-of-global-climate-2025>

- Yan, L., Wu, Y., & Zhu, G. (2024). Automatic rating method based on deep transfer learning for machine translation considering contextual semantic awareness. *Artificial Intelligence Review*. Advance online publication. <https://doi.org/10.1007/s10462-024-10633-x>
- Yao, G. (2022). Evaluation of machine translation in English–Chinese automatic subtitling of TED Talks. *Modern Languages, Literatures, and Linguistics*, 1(1), 1–17.
- Yılmaz, M. D. (2017). *Altyazı çevirilerinde kültürel öğelerin aktarımı ve çeviri sorunları* [Unpublished MA thesis], İstanbul University, Institute of Social Sciences.
- Zhou, X., & Zhou, J. (2024). A study on the quality assessment of the Chinese-English subtitle translation among four machine translation tools: A case study of Wreaths at the Foot of the Mountain. *transLogos Translation Studies Journal*, 7(2), 29–56. <https://doi.org/10.29228/transLogos.68>