Translating Cultures: Multimodal Literacy in Representing Indonesian Culture in Netflix Series Of Street Food Asia

Bernadine Cresentia[1], Rahayu Purbasari[2]

{bernadinecresentia29@gmail.com[1], rpurbasari@unj.ac.id[2]}

Faculty of Languages and Arts, English Literature Study Programme,
Universitas Negeri Jakarta[1,2]

Abstract. This paper examines the role of multimodal literacy in translation studies, focusing on how digital narratives bridge cultural gaps through the interplay of textual, visual, and auditory elements. Using Gunther Kress and Theo van Leeuwen's multimodal literacy theory as a foundation, the study analyses Netflix's *Street Food Asia* (Episode Yogyakarta, Indonesia) to explore the challenges and strategies of preserving cultural authenticity in subtitles and narrative voice overs. The study emphasises the significance of multimodal coherence in translation through examining the ways in which various modes interact to communicate Indonesian cuisine and cultural identity to a worldwide audience. The results show that maintaining synchronisation across modes while striking a balance between linguistic accuracy and readability is essential in ensuring effective cultural translation. This study contributes to the understanding of cultural translation in audiovisual media and emphasizes the value of multimodal literacy in global storytelling.

Keywords: Multimodal Literacy, Audiovisual Translation, Indonesian Culture, *Street Food Asia*

1 Introduction

Multimodal literacy has become an essential framework for understanding how different semiotic modes, such as visuals, sounds, and text work together to construct meaning in various media forms. Gunther Kress and Theo van Leeuwen's multimodal literacy theory highlights the dynamic interplay of these modes and how they contribute to communication across contexts, particularly in media that combine visual and auditory elements. This theory is increasingly applied in fields such as audiovisual translation (AVT), where the simultaneous use of multiple modes, including language, visuals, and sound, is key to effectively conveying meaning. Unlike traditional forms of communication that rely heavily on spoken or written text, AVT requires an intricate balance between modes to ensure that the cultural and contextual essence of the source material is preserved in the target context. The difficulties and possibilities of multimodal communication—particularly in situations involving cultural representation—are becoming more and more important as media consumption expands globally.

One area where multimodal literacy is particularly relevant is audiovisual media, which inherently combines linguistic, visual, and auditory elements. This approach is highly applicable to audiovisual content like films, TV shows, and documentaries, where meaning is not conveyed through language alone but is supported and enhanced by other modes. For example, a study has explored how audio-described films leverage multimodal elements, such as soundscapes and visual cues, to enhance accessibility for visually impaired audiences [1]. Similarly, another research has been examined gastronomic tourism websites to analyze how visual and linguistic elements work together to shape regional identity [2]. These studies underline the importance of multimodal approaches in representing and translating complex cultural narratives, but food documentaries which is a unique form of audiovisual media, remain relatively underexplored in this regard. Food documentaries present rich, multimodal narratives that use visuals of food, cooking processes, oral storytelling, background sounds, and subtitled translations to tell stories that are both deeply personal and culturally significant.

Netflix's *Street Food Asia* is a particularly compelling example of how food documentaries employ multimodal storytelling to translate culture for global audiences. The series blends vibrant cinematography, evocative soundscapes, and engaging narrative voice overs with subtitles that ensure accessibility across linguistic barriers. This integration of modes allows the series to showcase the culinary traditions of various Asian countries while immersing viewers in the broader cultural and historical contexts that shape those traditions. In the Yogyakarta episode, for instance, Indonesian culture is represented not only through its culinary practices but also through the personal stories of the individuals featured, the lively ambiance of bustling street markets, and the sounds of food preparation. These elements work together to convey a deeply immersive and authentic depiction of Indonesian cultural identity. Food is portrayed as a cultural artifact, a repository of history, a bridge between generations, and a medium through which communities sustain their social ties. While prior research has examined multimodal literacy in various contexts, including audio-described films and tourism promotion, there has been limited focus on how food documentaries use multimodal strategies to represent culture authentically. This gap is particularly significant given the unique challenges these documentaries face in translating not only language but also non-verbal cultural expressions, sensory experiences, and emotional connections. Subtitles, for example, must convey linguistic meaning while preserving cultural nuances, such as idiomatic expressions and metaphors that may not have direct equivalents in the target language. Similarly, visuals and soundscapes must complement the narrative to convey cultural depth, yet they risk oversimplifying or exoticizing the culture for international audiences. Food documentaries, therefore, require a sophisticated approach to cultural translation, balancing authenticity with accessibility and ensuring that the richness of the represented culture is not reduced to superficial stereotypes.

This study focuses on the Yogyakarta episode of *Street Food Asia* as a case study to explore how multimodal elements, specifically subtitles, narrative voice overs, and visuals interact to construct and communicate an authentic representation of Indonesian culture. By applying Kress and van Leeuwen's multimodal literacy theory [3], the research investigates how the series navigates the complex process of translating culture for a global audience. Unlike previous studies that have largely focused on the linguistic dimensions of audiovisual translation, this research takes a holistic approach by examining the synergy between linguistic and non-linguistic modes in constructing cultural narratives. The Yogyakarta episode provides a rich context for this analysis, as it highlights the intersection of food, history, and community

while employing a sophisticated blend of modes to convey the cultural significance of Indonesian culinary traditions.

In the realm of audiovisual translation, especially within the context of food documentaries like Street Food Asia, the challenge lies not only in linguistic translation but also in maintaining cultural authenticity while ensuring global comprehension. Subtitling in documentaries requires a balance between the need to make content accessible to a global audience and the necessity of preserving the cultural depth of the original material [4]. This is particularly important in Street Food Asia, where Indonesian food terms such as Mbah and Gudeg remain untranslated, allowing viewers to experience the cultural specificity of these dishes. Translation in global cinema often involves the negotiation between domesticating the content for ease of understanding and foreignizing it to retain the richness of the source culture [5]. In this way, Street Food Asia employs a strategy of foreignization, where cultural terms and practices are presented in their original form, allowing the audience to engage directly with Indonesian culinary traditions. Moreover, voice-over narration plays a pivotal role in documentary filmmaking, offering both linguistic translation and cultural contextualization [6]. Prianto's voice-over in Street Food Asia provides essential background to the food featured in the episode, enriching the viewer's understanding of Indonesian culture beyond what subtitles alone can convey. Thus, the combination of subtitles, visuals, and voice-over in Street Food Asia creates a multimodal translation that not only facilitates comprehension but also preserves the cultural integrity of the food being showcased. By analyzing the multimodal strategies used in *Street Food Asia*, this study contributes to the broader discourse on cultural translation and representation. It emphasizes the unique potential of food documentaries to bridge linguistic and cultural divides, offering global audiences an immersive and nuanced understanding of diverse cultures. At the same time, it underscores the challenges of representing culture in a way that remains authentic and respectful while ensuring accessibility for international viewers. Through this investigation, the study highlights the broader implications of multimodal literacy for media studies, cultural translation, and global communication. Ultimately, it aims to uncover how food documentaries like *Street Food Asia* operate as cultural texts, transcending language to communicate the richness of human experiences, emotions, and identities through the universal language of food.

2 Research Method

This study adopts Gunther Kress and Theo van Leeuwen's multimodal literacy theory to investigate the interaction of subtitles, narrative voice overs, and visuals in conveying cultural authenticity in Netflix's *Street Food Asia* [3]. As a major international streaming service, Netflix purposefully expands its content library by including local productions, which in turn promotes multilingual translation and encourages vibrant cross-cultural interactions. This makes Netflix a valuable subject for research in the fields of Translation and Cultural Studies [7]. The research follows a systematic methodology, beginning with an in-depth analysis of subtitle translations to identify the strategies used and evaluate their effectiveness in preserving cultural nuances. The next step involves a close examination of Kevindra Prianto's narrative, focusing on how it reflects Indonesian culture and integrates with the visual and textual elements of the series. These steps aim to uncover how the series uses multimodal storytelling techniques to construct a culturally authentic portrayal of Indonesian culinary and cultural identity, contributing to the broader understanding of multimodal approaches in audiovisual translation.

3. Result And Discussion

This section presents an in-depth analysis of the linguistic strategies and multimodal elements employed in the Netflix documentary series *Street Food Asia*, particularly in its Yogyakarta, Indonesia episode. The series' approach to translating Indonesian culture—specifically the culinary traditions of Yogyakarta—is multifaceted, combining textual translation (subtitling), visual representations, and auditory elements to convey the essence of Indonesian culture to a global audience. This analysis focuses on how linguistic, visual, and auditory strategies work together to represent food culture and its contextual significance, aligning with the research goal of exploring language and cultural translation in food documentaries.

3.1 Textual Translation

The core of translating cultural terms in *Street Food Asia* lies in the subtitling process, where the richness of Indonesian culture, especially food, needs to be communicated to a global audience while maintaining its authenticity. Many of the culturally significant terms in the Yogyakarta episode, such as *Mbah*, *Jajan Pasar*, *Gudeg*, *Bakso*, *Gado-gado*, *Nasi Goreng*, and *Mie Lethek*, remain untranslated in the subtitles. This deliberate choice reflects the tension between the need to preserve cultural identity and the necessity of making these terms accessible to a non-Indonesian-speaking audience. By leaving these terms untranslated, the series introduces viewers to the linguistic diversity of Indonesia, while also encouraging them to engage with local food culture beyond mere translation. The challenge here lies in conveying the meanings and cultural implications of these terms without providing explicit translations, forcing viewers to rely on context, visual cues, and narrative explanation to understand the deeper cultural significance.

3.1.1. Translating Cultural Terms

In the *Street Food Asia* Yogyakarta episode, food terms like *Gudeg*, *Bakso*, and *Gado-gado* remain untranslated, requiring viewers to learn about these dishes not just through the text but also through their visual and auditory presentation. Here, language becomes a key tool for cultural education, where each term carries layers of meaning and cultural significance. For example, Mbah, a term used to refer to an elder or respected figure within the community, is not translated, but its significance is explored through the interactions between local food vendors and their families. In Indonesian culture, *Mbah* is not just a term for an elder but also a sign of respect. The word becomes a cultural artifact, signaling the importance of respecting elders in Indonesian society. In the context of the episode, the word *Mbah* often appears when addressing senior figures in the culinary tradition, further emphasizing the generational transmission of culinary knowledge. Another example, Jajan Pasar, which refers to traditional snacks sold in local markets, is a culturally significant term that is left untranslated in the subtitles. This term reflects the dynamic street food culture of Indonesia, where food is not only a commodity but also a social activity. *Jajan Pasar* embodies the bustling, vibrant atmosphere of traditional Indonesian markets, where people gather to sample a wide variety of foods. By leaving the term untranslated, the series allows viewers to experience the rich cultural setting of the Indonesian market and the social practices surrounding food consumption.

Translation is a difficult process as it has differences in language systems, sentence structures, and cultural backgrounds between the source language (SL) and the target language (TL). These differences often make it hard for translators, especially when they have to deal with cultural references [8]. The decision to retain terms like *Gudeg* and *Bakso* also highlights the challenge of linguistic translation. These dishes are regional staples, each representing specific localities and histories. *Gudeg*, for instance, is a slow-cooked dish made from young jackfruit, primarily associated with Yogyakarta. In the subtitles, the series does not provide a direct translation of *Gudeg*, allowing the dish's name to remain tied to its cultural origins. However, through Prianto's voice-over and the visuals of food preparation, the audience is given context that links *Gudeg* to Yogyakarta's culinary identity and heritage.

3.1.2. Adapting Local Terms for Global Comprehension

Given the complex nature of translating food culture, *Street Food Asia* employs several strategies to ensure that global audiences can grasp the cultural significance of Indonesian food while retaining its authenticity. These strategies involve both linguistic adaptation and contextualization, which work together to ensure that viewers from different cultural backgrounds can understand the terms and connect with the culinary culture being depicted. One key strategy is the use of descriptive subtitles. Terms such as *Bakso* and *Nasi Goreng* appear alongside short descriptions in the subtitles that explain the dish's basic ingredients or preparation method. For instance, *Bakso* is described as "meatball soup," while *Nasi Goreng* is referred to as "fried rice." While these descriptions are brief, they serve to anchor the dishes in familiar culinary terms for global viewers. However, this adaptation is carefully balanced: the terms *Bakso* and *Nasi Goreng* remain in their original form to preserve their cultural identity, while the additional context helps viewers comprehend the dish.

Another adaptation strategy involves narrative contextualization through Kevindra Prianto's voice-over. His commentary does not just translate the food, but also conveys its cultural significance. For instance, while describing *Mie Lethek*, a traditional noodle dish from Yogyakarta, Prianto explains that it has historical roots tied to the region's past, made with local, often humble ingredients. His narrative helps contextualize the food within its historical and cultural background, enabling international viewers to understand its deeper significance beyond the ingredients or preparation methods. The visual elements also support the translation of cultural terms. When the subtitle *Mbah* appears, the accompanying visuals show an elderly vendor preparing food, providing a non-verbal cue to the role of elders in Indonesian society. Similarly, scenes of bustling traditional markets and vibrant food stalls help viewers understand the concept of *Jajan Pasar* without needing a direct translation. The visual representation reinforces the connection between food and local cultural practices, offering a more immersive experience.

3.1.3. Subtitling Constraints and Cultural Impact

Subtitling, as a method of translation, imposes certain constraints due to its reliance on limited space and time. These constraints often result in the omission or simplification of cultural details, which can lead to a reduction in the depth of cultural nuance. For

instance, terms like *Mie Lethek* and *Gudeg* are not explained in great detail, and some of their cultural or historical contexts may be lost if viewers do not engage with the visual or narrative elements closely. It is evident that achieving full equivalence between two languages is extremely challenging [9]. Despite these limitations, the multimodal approach of *Street Food Asia* mitigates these constraints. While subtitles alone cannot fully convey the richness of Indonesian culture, the combination of text, visuals, and narration allows for a deeper understanding. The visuals of Yogyakarta's markets, food preparation, and the locals' interactions with food help fill in the gaps left by the subtitles. Furthermore, Prianto's voice-over provides additional layers of information, giving viewers a broader understanding of the cultural significance behind the food.

3.2 Narrative Translation (Kevindra Prianto's Voice Over)

In the context of literary translation, the act of translating extends beyond mere visual interpretation, requiring the involvement of multiple sensory and emotional elements [10]. Kevindra Prianto's voice-over is an essential part of the narrative translation process in *Street Food Asia*. Through his narration, Prianto bridges the gap between the Indonesian culinary world and the global audience. His commentary not only explains the food but also positions it within its cultural, historical, and social context, helping viewers understand Indonesian food culture in a more immersive way.

3.2.1 Prianto's Narrative in Contextualizing Indonesian Culture

Prianto's narrative serves as the primary tool for contextualizing Indonesian culture within the documentary's framework. His voice provides a narrative that connects the dishes featured in the episode to their geographical and cultural origins. By doing so, Prianto helps bridge the gap between unfamiliar foods and their significance to viewers outside of Indonesia. For example, when discussing *Gado-gado*, Prianto explains that it is more than just a salad; it is a symbol of Indonesian diversity, where different ingredients from various regions are brought together to form a dish that is both harmonious and representative of the nation's multicultural identity. His voice provides insight into the social and historical layers of each dish, transforming the food from a simple meal into a cultural artifact.

3.2.2 Simplification of Cultural Details for Accessibility

To make the culinary experience accessible to global viewers, Prianto simplifies certain cultural details without sacrificing the integrity of the food's cultural significance. His explanations focus on the most salient features of each dish, such as its history, regional variation, or unique ingredients, while omitting more complex aspects that may be too niche for the average viewer. This selective simplification ensures that the documentary remains accessible without overwhelming the audience with too much detail. For example, Prianto might describe *Gudeg* as a "slow-cooked jackfruit dish from Yogyakarta," providing enough information for viewers to understand its key characteristics while avoiding an in-depth discussion of the complex variations in its preparation. Similarly, *Bakso* is framed as "meatball soup," a term simple enough for

international audiences to grasp, while still providing a window into Indonesian street food culture.

3.2.3 Cultural Identity Representation in Subtitles and Visuals

Prianto's narrative complements both the subtitles and the visuals to create a unified representation of Indonesian cultural identity. While the subtitles provide direct translations of the food terms, the visuals situate them within their cultural context, showing the markets, the people, and the culinary traditions that define each dish. Prianto's voice-over ties everything together, offering an explanation of how the food relates to Indonesian cultural values and practices. For instance, the voice-over describing *Mie Lethek* ties the dish to its historical roots in Yogyakarta, where the noodles were once made from humble ingredients by local farmers. This connection between the food and the people of Yogyakarta provides viewers with a deeper understanding of how food is intertwined with cultural identity, reinforcing the idea that food is not just about nourishment but also about heritage and social connections.

3.2.4 Auditory Elements Enriching the Narrative

The auditory elements, particularly the traditional Gamelan music and the old pop song playing in the background, also play a significant role in enriching the narrative. Gamelan, a traditional form of Indonesian music, evokes the cultural ambiance of the region and adds to the immersive experience of the documentary. The sounds of Gamelan enhance the portrayal of traditional food-making processes, underscoring the historical and cultural significance of the food. Additionally, the use of old pop song brings a nostalgic and modern layer to the narrative, bridging the past and present of Indonesian food culture. These auditory elements work in tandem with the visual and textual components to offer a rich, multisensory experience that helps contextualize Indonesian culinary traditions for a global audience. The music becomes a vehicle for conveying the mood and essence of Yogyakarta, making the food not just something to be tasted but something to be experienced through multiple senses.

However, the balance between local specificity and global accessibility is where the translation strategy faces challenges. While leaving cultural terms untranslated can enrich the viewing experience, it also assumes a certain level of cultural literacy on the part of the audience. Not all viewers may be familiar with the cultural concepts tied to the foods being presented, which could create a barrier for understanding. By simplifying complex cultural terms or omitting more intricate details of food history, the series ensures that its content remains accessible to a wider audience, but this comes at the cost of losing some of the nuances that define Indonesian food culture. While Prianto's voice-over mentions the dish's humble origins, there is a limit to how much can be conveyed in a short documentary format. In this case, the balance between global accessibility and local specificity is maintained through the use of simple descriptions. These simplifications ensure that the dishes are understood by a wide audience but might lose some of the richness of their local cultural significance. This approach represents a broader issue within translation studies on ways to balance the preservation of cultural identity with the need for global comprehension. In the case of *Street Food Asia*, the simplifications made in the subtitles and the choice to focus on the most salient features

of the dishes reflect an understanding that a documentary's purpose is not only to preserve culture but also to make it accessible to a global audience. The challenge lies in determining where to draw the line between authenticity and accessibility, ensuring that the local culture is represented without overwhelming or alienating viewers who may be unfamiliar with it.

Finally, this analysis leads to a broader consideration of the implications for multimodal literacy in translation studies. Multimodal literacy refers to the ability to interpret and produce meaning through multiple modes of communication, including text, image, sound, and gesture. In translation studies, it emphasizes the importance of considering all modes of communication, especially in audiovisual translation. *Street Food Asia* exemplifies the power of a holistic approach, where language is just one part of the translation process, and meaning is constructed through the interaction of text, visuals, and sound. The success of this approach is evident in how the documentary conveys the cultural significance of Indonesian food while also ensuring that the translation is accessible to a global audience. By using subtitles, voice-over narration, and visual storytelling together, the series demonstrates the potential of multimodal translation strategies to preserve cultural depth while simultaneously making it accessible. This strategy is particularly useful in food documentaries, where the subject matter is inherently tied to local customs, traditions, and practices, and cannot be fully understood through text alone. Multimodal translation allows for a richer, more nuanced representation of culture that engages viewers on multiple levels, enriching their understanding and appreciation of the food and the culture it represents. Moreover, *Street Food Asia* highlights the growing importance of multimodal literacy in translation studies, especially in the context of globalized media. As more content is consumed across borders, it becomes increasingly important to develop translation strategies that consider the interplay between different modes of communication. This holistic approach is crucial not only for the effective translation of cultural content but also for the preservation and dissemination of cultural identity in a globalized world.

4. Conclusion

The analysis of Netflix's *Street Food Asia* (Yogyakarta episode) demonstrates the pivotal role of multimodal literacy in bridging linguistic and cultural divides in audiovisual translation. By leveraging a combination of textual (subtitles), auditory (narration and background music), and visual elements, the documentary effectively communicates Indonesian culture to a global audience while maintaining authenticity. The deliberate choices to leave culturally significant terms untranslated, such as *Gudeg* and *Jajan Pasar*, and contextualize them through narrative and visual cues, illustrate how food documentaries can transcend mere language translation to become cultural texts. This research underscores that multimodal coherence, achieving balance among linguistic accuracy, cultural authenticity, and audience accessibility, is essential for successful cultural translation. However, the study also highlights inherent challenges, such as the potential oversimplification of cultural nuances due to subtitling constraints and the risk of cultural exoticism in representing local traditions for international viewers.

References

[1] C. &. S. G. S. Jiménez Hurtado, „Multimodality, translation and accessibility: a corpus-based study of audio description," *Perspectives,* zv. 21, %1. vyd.4, pp. 577-594, 2013.

[2] V. &. K. S. Eisenlauer, „Multimodal literacies: Media affordances, semiotic resources and discourse communities.," *Journal of Visual Literacy,* zv. 39, %1. vyd.3-4, pp. 125-131, 2020.

[3] G. &. V. L. T. Kress, Reading images: The grammar of visual design, Routledge, 2020.

[4] J. &. R. A. Díaz Cintas, Audiovisual Translation: Subtitling., Manchester: St. Jerome Publishing., 2007.

[5] A. M. Nornes, Cinema Babel: Translating Global Cinema, University of Minnesota Press, 2007.

[6] J. &. O. P. Díaz Cintas, „Voiceover and dubbing," rev. *Handbook of translation studies*, 2010, pp. 441-445.

[7] S. Kuşçu Özbudak, „Subtitling in local original series of Netflix: Is "The Protector"," *RumeliDE Dil ve Edebiyat Araştırmaları Dergisi,* %1. vyd.17, pp. 383-394, 2019.

[8] M. N. L. U. &. S. M. A. Farkhan, „Translation strategies of food-related culture specific items in Indonesian subtitle of Netflix series the final table," *Insaniyat: Journal of Islam and Humanities,* zv. 4, %1. vyd.2, pp. 140-153, 2020.

[9] L. Huber, „Key cultural texts and translation as a creative act of cultural mediation," *Translation studies: theory and practice,* zv. 4, %1. vyd.2, pp. 32-40, 2024.

[10] R. &. C. M. Vidal, „The Translator's Gaze: Intersemiotic Translation as Transactional Process," rev. *Translating across Sensory and Linguistic Borders: Intersemiotic Journeys between Media*, Palgrave Macmillan, 2019, pp. 1-36.