

## SUBTITLE TRANSLATION FROM A MULTIMODAL PERSPECTIVE: TEXT AND VISUAL INTEGRATION IN ANIMATED FILM

Sri Widyarti Ali

*Universitas Negeri Gorontalo, Indonesia*

widyartiali@ung.ac.id

First Received: October 1<sup>st</sup>, 2024

Final Proof Received: December 31<sup>st</sup>, 2024

### Abstract

The global appeal of animated films has increased significantly, leading to an urgent demand for high-quality and easily comprehensible subtitle translations. This research investigates the complexities of subtitle translation in the animated film *Frozen* (2013) by exploring the interaction between visual elements and dialogue. Employing multimodal discourse analysis, the study evaluates how subtitles can enhance or diminish the narrative experience by aligning with or diverging from the film's visual context. The findings indicate that while translations generally capture the intended meaning and tone, there are critical moments—especially in humorous scenes—where subtitles fail to fully resonate with the accompanying visuals. This misalignment underscores the importance of a refined approach to translation that considers the interplay of various expressive modes. The research ultimately aims to contribute to best practices in audiovisual translation, emphasizing the need for coherence between subtitles and visual storytelling to enrich audience engagement.

**Keywords:** animated films, subtitle translation, multimodal discourse analysis, visual elements

### INTRODUCTION

The global surge in the popularity of animated films has significantly increased the demand for high-quality, easily comprehensible subtitle translations. Unlike live-action movies, animated films rely heavily on enhanced visual storytelling that intertwines dialogue, imagery, and sound to convey meaning. In this context, subtitles play a crucial role in overcoming language barriers, enabling international audiences to fully appreciate these cinematic experiences. However, translating subtitles for animated content presents unique challenges; it requires careful consideration of the interplay between spoken words, visuals, and audio components. Animated films represent a complex multimodal discourse, where text and visuals interact to shape the

narrative, necessitating that subtitles effectively account for both spoken dialogue and the visual context to accurately convey the intended meaning.

Despite the growth of research in audiovisual translation, existing studies predominantly focus on the linguistic aspects of subtitle translation. This narrow focus overlooks the critical role that visual elements play in animated films, where imagery often conveys as much meaning as the spoken word. The absence of comprehensive studies that integrate both visual and textual analyses in the context of animated films highlights a significant gap in the current literature. Addressing this gap is crucial because the visual component of animated films, with their unique features such as fantastical settings, exaggerated character expressions, and symbolic imagery, plays an indispensable role in shaping narrative understanding and audience engagement.

To address these complexities, this study aims to investigate how the integration of visual elements and dialogue in animated films influences subtitle translation, utilizing Professor Zhang Delu's (2009) multimodal discourse analysis framework as a foundational basis. This approach is particularly suitable for analyzing how visual and textual elements can be harmonized in subtitle translation, which is essential for maintaining the coherence and emotional depth of animated films. By applying Zhang Delu's comprehensive multimodal framework, this study intends to bridge the gap in current research and contribute new insights into the intricacies of translating subtitles for animated content, where visual storytelling is paramount.

The 2013 Disney film *Frozen* serves as the primary data source for this analysis due to its rich visual and auditory elements, exemplified by Elsa's magical powers and the expressive emotions of the characters. Examining how these elements influence subtitle translation can reveal specific strategies that translators use—or overlook—in aligning visual cues with textual content. The findings are expected to show patterns in how visual components impact subtitle translation, offering practical recommendations for translators to improve their work. This study also seeks to refine Zhang Delu's multimodal discourse analysis, adjusting its application to meet the distinct challenges posed by animated films.

The significance of this study lies in its potential to enhance subtitle translation practices, ensuring that subtitles align more effectively with visual storytelling, thereby enriching the audience's viewing experience. Improved subtitle translation that considers multimodal elements can increase the accessibility and cultural resonance of animated films for diverse audiences. By contributing to best practices in subtitle translation, this research promotes better cross-cultural understanding and appreciation through media. The primary objective of this research is to explore how text and visual elements interact in *Frozen* (2013) and how this interaction influences subtitle translation quality. The key research question guiding this study is: How does the interaction of text and visual elements in *Frozen* (2013) impact subtitle translation?

Recent studies in subtitle translation, such as those by Pérez-González (2021) and Matamala and Orero (2019), have highlighted the importance of a multimodal approach that integrates textual and visual elements to improve viewer engagement. Yet, there remains a noticeable gap in research that directly addresses how unique visual narratives in animated films influence subtitle translation. This study aims to fill that void by conducting an in-depth multimodal discourse analysis focused on synchronizing text with intricate visual storytelling.

This research will refine existing theories, particularly Zhang Delu's framework, and provide targeted recommendations to enhance subtitle translation practices for animated films.

Foundational works by Kress and van Leeuwen (1996) introduced the concept of multimodal discourse analysis (MDA), emphasizing the interplay between various communication modes in audiovisual content. However, their approach has not been extensively applied to subtitle translation for animated films, an area that requires unique consideration due to its reliance on visual storytelling. By integrating Zhang Delu's multimodal discourse analysis framework, which examines the synergy of cultural, contextual, content-based, and expressive elements, this study addresses the gap and expands the application of multimodal theory to animated film translation.

## METHOD

This study adopts a qualitative research paradigm, focusing on the interplay between visual elements and dialogue in subtitle translation within the animated film *Frozen* (2013). The choice of this film is rooted in its rich integration of visual and auditory elements, making it particularly suitable for examining how multimodality influences subtitle translation. The analysis concentrates on key scenes where visual storytelling significantly contributes to the narrative, utilizing English-to-Indonesian subtitles as the primary focus for translation evaluation.

To achieve the research objectives, this study employs a Multimodal Discourse Analysis (MDA) approach as developed by Professor Zhang Delu (2009). The methodology unfolds in three distinct stages:

1. **Visual Analysis:** This stage involves a detailed examination of key visual elements within the selected scenes of *Frozen*. The analysis focuses on aspects such as facial expressions, gestures, background imagery, and color schemes. Each visual element is scrutinized to determine its contribution to the overall narrative and its interaction with the dialogue, thereby providing insight into how these elements work in tandem to shape viewer understanding.
2. **Subtitle Analysis:** Following the visual analysis, a linguistic examination of the subtitles is conducted. This is focused on how the spoken dialogue is translated into text, assessing whether the subtitles accurately reflect the visual cues present in the corresponding scenes. This analysis helps identify any discrepancies between the spoken dialogue and its visual representation, as well as the implications of these discrepancies for audience comprehension.
3. **Multimodal Integration:** The final stage involves a comparative analysis of the text (subtitles) and visuals. This analysis evaluates how effectively the two modes integrate to create a coherent viewing experience. Particular attention is given to whether the subtitles account for visual elements, ensuring that meaning is conveyed effectively to the audience. By synthesizing insights from both the visual and subtitle analyses, this stage provides a comprehensive understanding of how multimodal elements interact in the context of subtitle translation.

This methodology is feasible given the availability of *Frozen* as a rich data source, allowing for an in-depth exploration of the subtitled content. The MDA approach directly aligns with the research questions by focusing on how the integration of text and visuals influences subtitle translation. By analyzing both the linguistic and visual elements of the film, this study aims to provide a thorough understanding of how multimodality shapes the subtitle translation process, ultimately contributing valuable insights to the field of audiovisual translation.

## **FINDINGS AND DISCUSSION**

### **Findings**

This analysis examines the interplay between visual elements and dialogue in *Frozen* (2013) through a multimodal lens, focusing on how subtitle translations enhance or detract from the narrative experience. The table 1 below summarizes key data points, including visual analysis, subtitle analysis, and multimodal integration for each scene analyzed.

Table 1 below summarizes key data points, including visual analysis, subtitle analysis, and multimodal integration

<b>Data Point</b>	<b>Time</b>	<b>Character/Utterance</b>	<b>Subtitle</b>	<b>Visual Analysis</b>	<b>Subtitle Analysis</b>	<b>Multimodal Integration</b>
1	00:05:36	Elsa: 'You are okay, Anna. I got you.'	Kau baik-baik saja, Anna. Aku bersamamu.	Elsa holds Anna with a fearful and concerned expression amidst an icy environment, emphasizing danger and protectiveness.	Maintains reassurance but shifts focus from physical support ('I got you') to emotional comfort ('Aku bersamamu').	Emotional weight captured but physicality not fully reflected, affecting the perception of Elsa's protective role.
2	00:05:45	Elsa's Father: 'Elsa, what have you done? This is getting out of	Elsa, apa yang telah kamu lakukan? Ini di luar kendali.	Elsa's father shows shock and concern, with gestures emphasizing panic amidst a contrast of a warm family	Accurately mirrors original dialogue's urgency, reflecting emotional and	The subtitle aligns well with visual elements, effectively emphasizing the gravity of the scene.

		hand.'		setting and dangerous ice.	contextual tension.	
3	00:11:02	Anna: 'We only have each other. Just you and me. What are we gonna do?'	Kita hanya memiliki satu sama lain. Hanya ada kau dan aku. Apa yang harus kita lakukan?	Anna shows vulnerability with sad expressions and open gestures in dim lighting, highlighting emotional depth.	Faithful translation that emphasizes mutual dependence without significant meaning loss.	Strong alignment with visuals, enhancing the emotional resonance of the scene.
4	00:18:05	Anna (to Prince Hans): 'Lucky you, it's just me.'	Kau beruntung, aku yang tertabrak.	Anna's playful demeanor and budding romantic interest in a bright, colorful background.	Literal translation reduces the playful, self-deprecating humor of the original.	Disconnect between playful visuals and more serious subtitle tone, impacting humor.
5	00:23:32	Anna: 'Elsa and I were really close when we were little, and then one day she just shut me out, and I never knew why.'	Elsa dan aku sangat dekat saat kecil. Lalu suatu hari dia mengacuhk anku dan aku tak pernah tahu alasannya.	Anna's expressions of sadness and confusion in a calm setting emphasize emotional hurt.	Accurate reflection of the original dialogue, maintaining emotional depth.	Strong integration between visuals and subtitle, enhancing viewer understanding of character dynamics.
6	00:29:49	A man: 'Are you a	Apa kamu seorang	Man's fearful expression and	Literal translation	Maintains serious tone but

		monster too?' / Anna: 'No, no, I'm completel y ordinary.'	monster juga? / Tidak, tidak. Aku hanya manusia biasa.	Anna's serious, defensive body language highlight tension.	but omits playful self- deprecation of original, impacting humor.	misses subtle humor, affecting multimodal balance.
7	00:39:31	Anna: 'I want you to take me up the North Mountain. Look, I know how to stop this winter.'	Aku ingin kau mengantark anku ke gunung utara. Dengar, aku tahu caranya menghentik an musim dingin.	Dynamic visuals of Anna's determination, animated gestures, and snow-filled background.	Subtitle captures original urgency and meaning effectively.	Strong alignment between text and visuals, enhancing urgency and emotional weight.
8	00:43:18	Svan: 'But I've just paid it off.'	Aku baru saja melunasiny a.	Exaggerated disappointment in Svan's expression and body language against a bright, comedic background.	Simple and effective translation, maintaining the humor.	Effective multimodal integration that strengthens comedic impact.
9	00:47:00	Olaf: 'Summer? Oh, I don't know why, but I've	Musim panas? Oh aku tak tahu itu, tapi aku selalu suka gagasan tentang	Olaf's wide- eyed, joyful expressions and contrasting cold environment emphasize	Maintains humor but more formal than original playful tone.	Supports humor but could improve playful tone to match visuals.

---

always	musim	irony.
loved the	panas, dan	
idea of	mentari,	
summer,	dan segala	
and sun,	kehangatan	
and all	nya.	
things		
hot.'		

---

### Multimodal Discourse Analysis Overview:

Across all the data points, the interaction between visual elements and subtitle translation in *Frozen* (2013) demonstrates how multimodal integration shapes the viewer's experience. Using Zang Delu's (2009) multimodal discourse analysis framework, which considers cultural context, contextual factors, content analysis, and expressive modes, we see several key patterns:

1. **Cultural Context:** The subtitles largely remain neutral and accessible across different cultures, with no major shifts in meaning required for an Indonesian audience. However, some subtle cultural undertones, such as humor or playfulness, could be more effectively captured to mirror the original tone.
2. **Contextual Factors:** The physical and emotional context of the scenes is closely reflected in the subtitles. Tension, humor, and emotional depth are preserved, although minor shifts in tone—such as the downplaying of humor or physical actions—sometimes affect the impact.
3. **Content Analysis:** The translation generally retains the core meaning of the dialogue. However, moments of linguistic or emotional nuance, especially in humorous or action-driven scenes, occasionally shift due to the difference in languages. In particular, playful or casual expressions in English may be translated in a slightly more formal tone in Indonesian.
4. **Expressive Modes:** Visual elements, including facial expressions, gestures, color schemes, and background imagery, play a crucial role in conveying the narrative. In most cases, the subtitles align with these visual cues, but there are moments when the translation could better reflect the specific actions or emotions present on screen. For example, translations could incorporate more playful language to match the character's body language or expressions, enhancing the overall multimodal experience.

### Discussion

The integration of visual elements and dialogue in *Frozen* (2013) is crucial for effective subtitle translation. Using Zhang Delu's (2009) multimodal discourse analysis framework, this study demonstrates that while Indonesian subtitles often preserve the meaning and tone of the original

dialogue, certain areas show a need for improvement. For instance, in Data 1, Elsa's phrase "I got you" is translated as "Aku bersamamu," shifting from the physical action of catching Anna to a more abstract notion of emotional support. This creates dissonance with the visual of Elsa holding Anna, highlighting the need for subtitles to reflect physical actions accurately. This finding resonates with Díaz Cintas and Remael's (2007) emphasis on the importance of coherence between audio-visual elements and text for a unified narrative, which ultimately enhances viewer immersion. Their work underscores that subtitles should not only convey linguistic meaning but also align with visual cues to support the overall narrative experience.

The findings further illustrate the importance of contextual factors in multimodal integration. In Data 2, Elsa's father's line, "This is getting out of hand," is translated as "Ini di luar kendali," effectively mirroring the original dialogue's urgency. This translation, combined with the visual elements of the scene—such as Elsa's father's panicked expressions and spreading ice—emphasizes the emotional stakes and escalating tension. This observation supports Gambier's (2013) argument that effective audiovisual translation requires a careful balance between visuals and text to enhance the emotional impact and maintain storyline coherence. Gambier's analysis reinforces that contextual factors, such as character authority and the visual portrayal of stress, contribute to how audiences interpret translated subtitles.

However, seamless integration between text and visuals is not always achieved, particularly in humorous contexts. In Data 4, Anna's playful remark, "Lucky you, it's just me," is translated as "Kau beruntung, aku yang tertabrak," which shifts the original self-deprecating humor to a more literal and serious tone. This diminishes the comedic effect and fails to align with Anna's playful demeanor on screen. This result aligns with Chaume's (2018) findings, which emphasize the importance of preserving the original tone in subtitles, especially in humor translation. Chaume highlights that humor is often culturally specific and dependent on subtle linguistic nuances, making it challenging to translate without losing its essence. Zhang Delu's (2009) framework also underscores the necessity of considering character expressions and body language to maintain the intended comedic impact. The findings indicate that more attention to these non-verbal cues in subtitle translation could significantly enhance audience engagement by preserving the humor's original tone and context.

Furthermore, the study's analysis of Data 8 supports the work of Kress and van Leeuwen (1996), who established that multimodal discourse analysis should account for the interplay of visual, linguistic, and auditory modes. In Svan's humorous remark, "But I've just paid it off," translated as "Aku baru saja melunasinya," the subtitle effectively captures the simplicity and humor, aligning with the exaggerated disappointment on his face. This supports Kress and van Leeuwen's theory that successful integration of text and visual modes can amplify comedic timing and strengthen audience connection. The alignment here shows how subtitles can preserve and even enhance the visual storytelling in animated films when approached with attention to multimodal principles.

Data 9, involving Olaf's whimsical line about summer, underscores the need for cultural and tonal considerations in translation. While the subtitle captures the overall meaning, its more formal tone does not fully reflect Olaf's playful and naive character. This highlights a nuanced



challenge discussed by Matamala and Orero (2019), who emphasized that technological and theoretical advances in multimodal translation should facilitate better integration of visual and linguistic cues. The current finding suggests that incorporating a more playful tone in subtitle translations could enhance audience perception and engagement, particularly when dealing with characters known for their unique humor and expressions.

In conclusion, this study reinforces previous research findings, such as those by Gambier (2013), Díaz Cintas and Remael (2007), and Chaume (2018), which collectively highlight the importance of multimodal alignment in subtitle translation. By addressing visual and contextual elements alongside linguistic meaning, translators can ensure a cohesive viewing experience. This research contributes to refining Zhang Delu's multimodal discourse analysis framework by demonstrating its applicability to animated films and emphasizing the need for a balanced approach that respects both the visual narrative and the linguistic integrity of the subtitles. Enhanced multimodal practices will ultimately support better cross-cultural understanding and richer audience engagement with animated media.

## CONCLUSION AND SUGGESTION

### Conclusion

In conclusion, this study emphasizes the crucial role of multimodal elements in the subtitle translation process, particularly in animated films like *Frozen* (2013). By applying Zhang Delu's multimodal discourse analysis framework, the research highlights the interplay between dialogue and visual elements, illustrating how this interaction significantly affects the viewer's experience. While the Indonesian subtitles often succeed in preserving the original meaning and tone, there are notable instances where greater attention to visual cues could enhance coherence and emotional engagement. It is imperative for subtitle translators to consider the visual context and expressive modes to improve the overall narrative experience.

### Suggestion

To enhance the effectiveness of subtitle translations in animated films, several recommendations are proposed. Subtitle translators should undergo specialized training focused on integrating visual and textual elements, with workshops on multimodal discourse analysis to develop their skills. Additionally, encouraging collaboration between translators and filmmakers can help ensure that subtitles accurately reflect visual storytelling. Leveraging technology, such as AI-assisted tools, could assist in identifying key visual cues that need to be incorporated into subtitles. Moreover, ongoing research into the impact of multimodality on subtitle translation, along with mechanisms for audience feedback, will provide valuable insights. By implementing these recommendations, translators can significantly improve the alignment of text with visual cues, enriching the viewer's experience and fostering better cross-cultural understanding through animated films.

**REFERENCE:**

- Chaume, F. (2018). *Audiovisual Translation: Dubbing*. Routledge.
- Díaz Cintas, J., & Remael, A. (2007). *Audiovisual Translation: Subtitling*. Routledge.
- Gambier, Y. (2013). *Multimodality and Audiovisual Translation*. *Translation Studies*, 6(1), 29-45.
- Garcia, S., & Thompson, R. (2020). Multimodal Strategies in Subtitle Translation for 3D Animated Movies. *Translation and Media Studies Journal*, 8(2), 112-130.
- Jia, L. (2019). *The Role of Semiotic Interplay in Subtitling: A Multimodal Perspective*. Semantics Scholar.
- Kim, H.-J. (2019). Visual Consistency and Textual Clarity: Balancing Elements in Animated Film Subtitling. *International Journal of Multimodal Communication*, 5(1), 78-95.
- Kress, G., & van Leeuwen, T. (2006). *Reading Images: The Grammar of Visual Design*. Routledge.
- Liu, M., & Chen, Y. (2021). Enhancing Comprehension in Animated Films through Multimodal Subtitling Techniques. *Journal of Audiovisual Translation*, 19(3), 245-263.
- Martinez, A., & Lopez, E. (2022). *Semiotic Integration in Subtitling: A Multimodal Approach to Animated Narratives*. *Semiotics and Translation Quarterly*, 14(4), 310-329.
- Orero, P., & Matamala, A. (2019). *Eye Tracking in Audiovisual Translation*. Routledge. DOI: 10.4324/9780429429427
- Pérez-González, L. (2021). *Multimodal Translation with the Blind: A Practical Guide to the Subtitling of Visual and Tactile Information for a Blind Audience*. John Benjamins Publishing Company. DOI: 10.1075/ata.xix
- Singh, P., & Das, R. (2023). *Adaptive Subtitling Techniques for Interactive Animated Media*. *Journal of Digital Media Translation*, 10(1), 50-68.
- Zhang, D. (2009). *Multimodal Discourse and Foreign Language Teaching*. Shanghai Foreign Language Education Press.