

# Multimodal Discourse Analysis of Nobody Based on Subtitle Corpus

Yan Zhang, Keyu Li

School of Foreign Languages, Beijing Forestry University, No. 35 Tsinghua East Road, Beijing, 100083, China

---

**Abstract:** The animated film *Nobody* constructs a unique workplace metaphor in the world of Chinese monsters from a distinctive perspective. Through the survival dilemmas and spiritual struggles of Chinese monsters, it reflects the growth predicaments of ordinary people in reality. From a cross-cultural perspective, the study of film subtitle should focus on how to enhance the cross-cultural adaptability of Chinese films and traditional culture while preserving the original cultural connotations of the film. This paper explores the application of multimodal discourse in the English subtitle of *Nobody* from four dimensions: culture, context, content and expression. It aims to provide insights for the application of multimodal discourse analysis in cross-cultural communication and references for subtitle translation of domestic films, so as to promote the creation of local culture and the understanding and integration of diverse cultures.

**Keywords:** Corpus, *Nobody*, Multimodal Discourse Analysis, Subtitle.

---

## 1. Introduction

As a key component of films, subtitles are not only the direct presentation of verbal interactions between characters, but also the collaborative interpretation of multimodal symbols such as images and sound effects. Since film subtitles need to convey multiple meanings of works within a limited time and space, and their translation quality profoundly affects the audience's understanding and emotional resonance with the film content, film subtitle translation is a research field worthy of in-depth exploration.

As an emerging field of linguistics, multimodal discourse analysis focuses on the collaborative meaning-making process of symbols such as words, images and sounds, providing a new perspective for the study of film subtitle translation. The subtitles of *Nobody* contain both textual information and convey meanings with the support of character actions and scene sound effects, so multimodal theory can be used to analyze its translation strategies. Although multimodal discourse analysis has been widely applied in many fields, research on subtitle translation of animated films featuring "small-character narratives" and "localized metaphors" is still in the exploratory stage. Taking multimodal discourse analysis as a theoretical tool, this study deeply explores the translation strategies of English subtitles of *Nobody*, the effect of cross-cultural communication of local culture and the acceptance of target audiences. It analyzes how translation can convey meanings and achieve emotional resonance in cross-cultural contexts on the basis of retaining the core connotation of the struggle and growth of small characters, so as to provide theoretical and practical references for subtitle translation and cross-cultural communication of similar localized films.

## 2. Theoretical Framework of Multimodal Discourse

### 2.1. Cultural Dimension

The cultural dimension is the key to the realization of multimodal discourse communication, determining the strategies, traditions and forms of meaning transmission.

Without this dimension, the construction of situational context will lose persuasiveness. This dimension consists of values and themes. Values cover deep cognitions such as thinking patterns, life philosophies, living habits and potential rules in a specific culture, while themes are the specific carriers of values, presenting cultural characteristics through communication or structural potential. Cultural considerations are particularly important in film subtitle translation. Due to differences in thinking patterns and living habits between Chinese and Western cultures that may form cross-cultural communication barriers, translators need to fully explore the cultural connotations of the original film, such as traditional cultural symbols and value orientations, and help target audiences overcome cultural barriers through targeted translation strategies to accurately convey the cultural characteristics of the source language.

### 2.2. Contextual Dimension

The contextual dimension is one of the core dimensions of multimodal discourse analysis, referring to the specific environment of communication activities, including time, place, participants, social and cultural background and other elements, and is restricted by communication themes, emotional attitudes and expression channels. The use logic of linguistic and non-linguistic symbols varies significantly in different contexts. As a linguistic symbol, film subtitles need to collaborate with non-linguistic symbols such as images and sounds to construct contextual meaning [1]. Specifically in subtitle translation practice, translators need to thoroughly grasp the contextual information of the original film, including background setting, character relationships and plot development, and combine the cognitive and aesthetic habits of target audiences. By adjusting language expressions, translators ensure that subtitles can accurately restore the contextual atmosphere of the original film and achieve contextual collaboration with non-linguistic symbols.

### 2.3. Content Dimension

The content dimension includes two sub-dimensions: meaning and form, which is the core carrier of meaning transmission in multimodal discourse. The meaning layer

focuses on the information connotation conveyed by symbols, including conceptual meaning, interpersonal meaning and textual meaning, all of which are restricted by field, tenor and mode of discourse. The formal layer focuses on modal types, namely specific modalities such as language, visual images, sounds and touch, as well as the relationships between these modalities. The complementary relationship between modalities emphasizes that when a single modality cannot fully express meaning, other modalities are needed to supplement for complete meaning expression. In film subtitle translation, content dimension analysis should take information accuracy and logic as the core. It is necessary to ensure the complete transmission of conceptual meanings such as dialogues and plot development, strengthen the expression of interpersonal meaning and textual meaning through language adjustment, and coordinate the complementary relationship between subtitle text and other modalities to ensure the audience can smoothly understand the film content.

## 2.4. Expressive Dimension

The expressive dimension refers to the specific presentation form of the symbol system in multimodal discourse, covering linguistic and non-linguistic aspects. The linguistic aspect includes paralinguistic and pure language. Pure language focuses on the propositional meaning of language itself, such as the literal meaning of vocabulary and grammar, while paralinguistic conveys emotions and contexts through auxiliary features such as intonation, speed and vocal characteristics. As an important dimension of multimodal discourse meaning-making, the non-linguistic aspect covers various meaning-making elements beyond linguistic symbols. In addition to body movements, it also includes facial expressions, gestures, spatial distance and other specific forms. These elements convey emotional attitudes, character relationships and other implied meanings through visual or behavioral features, and collaborate with linguistic symbols to completely construct context [2]. From the perspective of semiotics, the primary media of language meaning transmission are acoustic semiotics transmitted by sound waves and graphic semiotics presented by words. Film subtitles belong to graphic semiotics, and their expressive effect needs to be comprehensively considered with non-linguistic factors such as tone, audio, intonation and screen layout. In subtitle translation, expressive dimension analysis should focus on language style, word selection and sentence structure. Translators should choose appropriate expressions according to the overall style and context of the film, and take into account the reading habits and aesthetic needs of target audiences to achieve expressive collaboration between subtitle text and other non-linguistic modalities and improve the readability and acceptance of subtitles.

## 3. Experience of Subtitle Translation in Nobody under Multimodal Discourse

The animated film *Nobody* takes the workplace experience of the Little Pig Monster in the King's Cave on Langlang Mountain as the main line, telling the story of his struggle for growth after encountering rule constraints, resource shortages and identity dilemmas in grassroots positions. From the unique perspective of the monster workplace, the film not only conveys the universal emotion of small characters

breaking through dilemmas, but also integrates distinctive Chinese characteristics. From cultural symbols such as “monsters” and the dilemma metaphor of “Langlang Mountain” to workplace expressions such as “connections” and “secured jobs”, these elements carry traditional Chinese cultural genes and observations of real society, providing a unique window for cross-cultural audiences to understand Chinese cultural characteristics and social mentality.

In addition, the meaning system of the film is deeply consistent with the core dimensions of multimodal discourse: culturally, symbols such as “monsters” and “Langlang Mountain” carry local Chinese cultural connotations; contextually, the workplace competition scenes and character interactions in the film construct a specific communication atmosphere; in terms of content, the survival struggles and growth demands of small characters convey the common aspirations of people in modern society; expressively, the colloquial dialogues and doggerel-style lines of characters have a unique linguistic style, which deserves careful consideration when converted into English subtitles. These elements interact through multimodal symbols such as language, images and sounds, jointly building a meaning network with both Chinese cultural characteristics and cross-cultural narrative power.

As a key carrier of cross-cultural communication of local culture, the subtitle translation of *Nobody* accurately conveys the core information of the film in four dimensions: culture, context, content and expression, ensuring that foreign audiences can understand the cultural connotation and narrative logic of the film. In addition, it achieves two-way perspective in the translation process, helping foreign audiences break through the barriers of cultural cognition and enabling Chinese audiences to reflect on cultural differences between China and the West through translation strategies. Therefore, multimodal analysis of its subtitle translation is not only a practical need to explore the path of cross-cultural communication of local characteristics, but also an important entry point to deepen the understanding of cross-cultural communication. Based on the theory of multimodal discourse analysis, this paper analyzes the subtitle translation in *Nobody* from four dimensions of culture, context, content and expression to explore the path of balancing the retention of local characteristics and cross-cultural communication.

### 3.1. Subtitle Translation at the Cultural Dimension

The cultural dimension is the foundation of meaning-making in multimodal discourse, whose core is to realize the cross-context transmission of cultural connotations through the symbol system [3]. As a key carrier of cultural meaning-making, film subtitles need to balance the characteristics of the source language and the cognition of target audiences in cross-cultural communication. Western audiences have no direct cognitive understanding of cultural symbols such as “Yao (monsters)” and “wild monsters” in *Nobody*, and there are understanding barriers due to differences in cultural contexts. As a dynamic cultural expression form, film subtitles need to undertake the function of cultural decoding, presenting the cultural and social background of the source language to the audience, adapting to the cognitive habits of Western audiences, and achieving a two-way balance between cultural transmission and acceptance.

Example 1: 妖 → Yao (Chinese monster)

In the film, the translation of culturally specific symbols

accurately conveys connotations through the collaboration of linguistic and non-linguistic modalities. As a unique supernatural symbol in traditional Chinese culture, “Yao” carries the complex meaning of “non-human beings, both good and evil”. In Western culture, “monster” focuses more on the negative attributes of “terror and evil”, with a natural deviation between the two. The subtitle adopts the strategy of transliteration combined with annotation, using “Yao” to retain the recognition of culturally specific names and “Chinese monster” to supplement basic concepts. Meanwhile, it collaborates with visual modalities such as the cute appearance of the Pig Monster and the fantasy scenes of the King’s Cave to weaken the terrifying attribute of “Yao”, helping Western audiences establish the cognition that “Yao is a unique supernatural existence in China” and avoid cultural misreading.

### 3.2. Subtitle Translation at the Contextual Dimension

Context, as the situational support of multimodal discourse, is a meaning field constructed by scene setting, character relationships and emotional attitudes, restricted by communication themes, interaction logic and communication forms. Chinese subtitles are often deeply bound to image scenes, character actions and sound intonations. Chinese audiences can instantly capture the emotional tension behind hesitant eyes and blank dialogues combined with the image atmosphere. However, this “language + non-language” collaborative logic is a blind spot of cultural cognition for Western audiences. English subtitles need to take language as the fulcrum, convert contextual clues hidden in images and actions into expressions decodable by the target language, reconstruct multimodal collaborative relationships, and help Western audiences perceive the atmosphere of communication scenes and character connections [4].

Example 2: 就你们几个小妖怪啊/ 就你们这几根葱/ 无名之辈 → Nobody

In the film, when grassroots monsters such as the Little Pig Monster try to join the King’s Cave, the Crow Monster, as an insider, mocks them with words and expressions to belittle their identity value. In the Chinese context, this mockery is realized through three levels of progression: “just a few of you” weakens the other’s sense of presence in quantity; “a few scallions” uses material metaphor to compare them to cheap and humble things; “nobodies” directly negates identity significance. In addition to the three-level progressive mockery, the film presents non-linguistic modal resources such as the Crow Monster’s contemptuous expression of raising eyebrows and curling lips and sharp mocking tone. Chinese audiences can quickly recognize the crushing contempt of the Crow Monster, as a powerful monster, for grassroots monsters. However, Western audiences lack cognition of the metaphor of “scallions”, and even if they see the same images and expressions, it is difficult to experience the progressive contempt in Chinese.

The English subtitle uniformly translates it as “Nobody” and strategically strengthens the ironic effect through repetition. In English, “Nobody” means an insignificant person. Repeating it not only simulates the sense of progressive mockery in Chinese, but also collaborates with non-linguistic modal resources such as the Crow Monster’s impatient expression and sharp tone in the image. Western audiences do not need to understand the metaphor of scallions; they can perceive the continuous and crushing contempt from

the repetition of Nobody, naturally capture the power gap between characters, and realize the cross-cultural transmission of contextual meaning.

#### Subtitle Translation at the Content Dimension

Discourse meaning at the content dimension covers conceptual meaning, interpersonal meaning and textual meaning, all of which are generated through the collaboration of multimodal symbols such as language, images and performances [5]. In Nobody, Chinese expressions such as the workplace metaphor of pulling strings and the ironic humor of “seriously” need to be nested in plot scenes and character actions for Western audiences to understand their meaning logic. English subtitles are reconstructed through linguistic strategies and deeply echo image and performance modalities to realize the cross-cultural transmission of the meaning layer.

Example 3: 不会走关系托人 → No strings to pull

In the film, the Little Pig Monster wants to enter the “King’s Cave” to get a secured job. The Toad Monster tips him that he “needs to pull strings to find connections”, but the Little Pig Monster sighs with his head down: “I don’t know how to pull strings”, expressing the dilemma of grassroots characters lacking network resources. “Pulling strings” is a representative metaphorical expression in Chinese workplace culture, directly referring to the social phenomenon of obtaining convenience through personal connections. This culture-specific expression has natural understanding barriers for Western audiences unfamiliar with Chinese workplace logic.

The English subtitle chooses “No strings to pull” for accurate correspondence, whose ingenuity lies in the cross-cultural adaptation of cultural metaphors. In Western culture, “pull some strings” is a conventional idiom, literally meaning “pulling ropes”, but actually metaphorically referring to “mediating relationships through personal connections”, which is highly consistent with the core connotation of “pulling strings” in Chinese. The subtitle directly conveys the conceptual meaning of “lacking network resources” in the negative form “No strings to pull”, enabling Western audiences to understand the core of Chinese workplace metaphors through familiar linguistic symbols.

### 3.3. Subtitle Translation at the Expressive Dimension

The expressive dimension of multimodal discourse covers linguistic and non-linguistic aspects. The linguistic aspect includes tone, rhythm and stylistic style, while the non-linguistic aspect includes actions, sound effects and image rhythm. The two collaborate to shape the rhythm and tension of meaning transmission. Chinese subtitles often create atmosphere through rhythmic rhetoric or metaphors in the form of doggerels, but Western audiences lack innate cognition of Chinese rhythmic logic and metaphorical associations. This requires English subtitles to reconstruct linguistic forms, adapt to Western expression paradigms, and resonate with image actions and sound intonations to balance naturalness and acceptance. In Nobody, the metaphor of lightning-fast eating and the rhythmic humor of made-up mantras achieve cross-cultural expression through the deep coupling of linguistic strategies and non-linguistic modalities [6].

Example 4: 闪电式吃法 → Blitz-feast

When monsters hurry to eat to rush to make mechanisms for capturing Tang Monk, the image constructs a race-against-time scene rhythm through characters’ quick eating actions

and high-frequency sound effects of bowls and chopsticks colliding. The source language “lightning-fast eating” uses “lightning” to metaphorize speed, but “lightning” in Western culture is mostly associated with natural phenomena and weakly related to the context of “rushing for time”. The English subtitle borrows the root “blitz” from the German loanword “blitzkrieg” (blitzkrieg, a well-known war metaphor in the West, implying the meaning of fast raid), combining with “feast” to form “Blitz-feast”. The plosive sound of “blitz” naturally fits the urgent rhythm of bowls and chopsticks colliding, and the war metaphor of “blitzkrieg” helps Western audiences associate with the urgency of “racing against time”, making the rhythm of linguistic symbols resonate with image actions and sound effects, strengthening the tense atmosphere of the scene.

## 4. Conclusion

Based on the framework of multimodal discourse analysis, this study systematically sorts out the cross-cultural communication logic of English subtitles of *Nobody* and refines the core strategy of four-dimensional collaboration: culture, content, context and expression. Culturally, it retains the uniqueness of local cultural symbols such as “Yao” and “Langlang Mountain” through transliteration plus annotation combined with metaphor reconstruction, making it understandable to Western audiences. Contextually, it strengthens scene atmosphere through language repetition and simulates character states through rhythmic dislocation, enabling emotional atmospheres such as “mockery” and “comedy” to be perceived across cultural differences. In terms of content, it conveys core meanings through idiom borrowing and homophonic reconstruction, and deepens emotional resonance with character actions, making Chinese expressions such as “pulling strings” and “seriously” understandable across cultures [7]. Expressively, it coordinates language rhythm with image actions and sound effects through root transformation and tongue twister adaptation, conveying the core of Chinese rhythmic humor. These strategies reveal the important role of multimodal symbols in the cross-cultural communication of localized animations and provide practical methods for similar works.

However, there is still room for further improvement in the research. On the one hand, the analysis scope of the multimodal collaboration mechanism can be expanded. Existing research mainly focuses on the interaction between language, visual and auditory modalities, but insufficiently explores spatial modalities. For example, the closed scene of “Langlang Mountain” has a dilemma metaphor, and the analysis of how subtitles strengthen this perception is insufficient. In terms of tactile modalities, details such as the rough hand movements and clumsy physical efforts of the Little Pig Monster when working can convey the state of grassroots struggles, but the existing research does not deeply explore how the stylistic style of subtitles, such as simple or heavy word choice, corresponds to these action details, failing to form an analysis framework covering more modal interactions [8]. On the other hand, the evaluation methods of translation effects need to be enriched. Existing analysis mainly judges the accuracy of translation by comparing the text content and semantic logic of the original and translated texts, lacking feedback evidence from the actual audience after watching, and unable to test the effectiveness of translation strategies with empirical data [9].

Future research can be promoted from three aspects: first,

expand the typological research of local animations, focusing on narrative types such as folk fantasy (e.g., Yao-Chinese Folktales) and workplace fables (e.g., *Nobody*), analyze multimodal translation strategies of different meaning-making methods such as mythological metaphors and real mapping [10], and summarize universally applicable communication models to help local works go international more accurately. Second, deepen the research on multimodal collaboration mechanism, break the traditional analysis framework of language as the main, audio-visual as the auxiliary, systematically analyze how the symbolic meaning of scenes corresponds to subtitle language and how character action details cooperate with subtitle style, and build a more comprehensive interactive analysis model [11]. Third, establish a simple and feasible empirical system, analyze audience feedback through bullet screen content in fan communities, or design targeted questionnaires focusing on humor perception and cultural understanding to test the effectiveness of translation strategies with low-cost methods and provide data support for practical optimization.

In summary, this study reveals the core laws of subtitle translation of localized animations. In the future, it is necessary to continue exploring in expanding the scope of modal analysis, enriching effect evaluation methods and refining typological research, so as to build a more systematic theoretical and practical system for the cross-cultural communication of local animations.

## References

- [1] G. O. Young, “Synthetic structure of industrial plastics (Book style with paper title and editor),” in *Plastics*, 2nd ed. vol. 3, J. Peters, Ed. New York: McGraw-Hill, 1964, pp. 15–64. [1] J. Boase-Beier, “Stylistic Approaches to Translation,” in *Stylistic Approaches to Translation*, New York: St. Jerome Publishing, 2014.
- [2] M. Jobert, “Phonostylistics and the written text,” in *The Cambridge Handbook of Stylistics*, 2nd ed., P. Stockwell and S. Whiteley, Eds. Cambridge: Cambridge University Press, 2014, pp. 231-240.
- [3] M. Lee, B. Roksoos, and D. R. Ewoldsen, “The Impact of Subtitles on Narrative Film Comprehension,” in *Media Psychology*, 2014, vol. 3, pp. 412-440.
- [4] E. Perego, “The Codification of Non-verbal Information in Subtitled Texts,” in *New Trends in Audiovisual Translation*, J. Díaz-Cintas, Ed. Bristol/Buffalo/Toronto: Multilingual Matters, 2009.
- [5] Y. P. Chen and C. H. Zhang, “Chinese Translation of English Film Subtitles—A Multimodal Analysis Focusing on Image-Text Relations,” in *Chinese Translators Journal*, 2017, vol. 38, no. 05, pp. 105-110.
- [6] S. L. Dai, “A Study on the Relationships Between Modalities in Multimodal Discourse,” in *Foreign Language Research*, 2017, no. 06, pp. 1-7.
- [7] J. Fang and J. R. Guo, “An Analysis of the Influencing Factor System and Its Action Path of Translation Behavior from a Cross-Cultural Perspective,” in *Foreign Languages in China*, 2020, vol. 17, no. 01, pp. 94-103.
- [8] J. B. Hou, “A Study on the Intertextual Function of Multimodal Web News Cartoons—Taking the ‘Big Fish Cartoon’ of Ifeng.com as an Example,” in *Foreign Language Education*, 2019, vol. 40, no. 03, pp. 39-43.
- [9] Y. Sun, L. B. Tian, and L. Chen, “FILMIP: A New Approach to Multimodal Metaphor Recognition in Films,” in *Journal of Beijing Film Academy*, 2022, no. 09, pp. 67-77.

- [10] J. G. Wang and J. X. Dai, "The Impact of Chinese-English Pragmatic Differences on Film Subtitle Translation—Taking the Subtitle Translation of *Dreaming under the 天籁* as an Example," in *Chinese Translators Journal*, 2023, vol. 44, no. 03, pp. 148-155.
- [11] Q. Wu, "Translation Challenges and Multimodal Discourse Analysis of Modern Pop Music Songs," in *Journal of Southeast University (Philosophy and Social Science)*, 2025, vol. 27, no. S1, pp. 135-138.
- [12] A. M. Nonis and Z. Su, "'Visible' Subtitles—An Interview with Professor Abel Marcus Nonis on 'Abusive' Film Translation Studies," in *World Cinema*, 2015, no. 03, pp. 147-156.
- [13] M. M. Yin, "A Study on Subtitle Translation Strategies of Classic French Literary Films from the Perspective of Multimodal Discourse Analysis," in *Contemporary Cinema*, 2020, no. 02, pp. 164-168.
- [14] X. F. Yin, "Subtitle Translation Strategies from the Perspective of Skopos Theory," in *Film Literature*, 2024, no. 18, pp. 130-134.
- [15] L. Zang and Y. Zhang, "Influencing Factors and Improvement Paths of Chinese Online Literature's External Communication from the Perspective of Multimodal Communication," in *Publishing Science*, 2025, vol. 33, no. 04, pp. 94-105.
- [16] N. N. Zhang, "Multimodal Discourse Analysis of Humorous Language in *Mansfield Park* from the Perspective of Sociolinguistics," in *Film Review*, 2020, no. 15, pp. 107-109.