



Multimodal Discourse Analysis and Communication of "Digital Dunhuang" in the Age of Artificial Intelligence

Lulu Zhan

Northwest Normal University, Lanzhou, Gansu, China

*2922022341@qq.com

Abstract. The virtual culture project of Digital Dunhuang covers symbols such as grotto exhibits, text, images, audio, etc., which is a typical multimodal discourse, and this feature is especially prominent in the age of artificial intelligence. This paper focuses on the method of multimodal interaction analysis and absorbs other two types of multimodal analysis in modal application and interaction to evaluate key clips of digital Dunhuang and its overall construction. This study found that each modality is not only affected by its own unique but also complements and collaborates with other modalities, and ultimately develops the Digital Dunhuang by harmonious multimodal coexistence through the visual, auditory, tactile and other sensory modalities. The result intensifies the audience's visual experience, thereby communicating Dunhuang culture and building up confidence in Chinese culture.

Keywords: Digital Dunhuang; multimodal discourse analysis; communication

1 Introduction

With the booming development of digital technology, we are stepping into a new era—digital intelligence (Wang & Yang, 2019)^[1]. The channel of communicating culture is changing from the traditional offline to the online. As a country with a long history and rich culture, the communication of China's traditional culture is of great significance for enhancing the country's soft power and strengthening international influence (Meng & Zhou, 2019)^[2]. Technologies such as digitalization, informatization, and intelligence are no longer isolated but have begun to integrate and form a complex but efficient technological ecosystem (Hu, 2022)^[3]. Digital intelligence technology has not only changed the way traditional culture is communicated but also provides potential for the innovation of its content (Jia, Cui & Zhang, 2024)^[4]. Empowered by modern technology, "online exhibition" has become an important form of "imbuing culture with life". Modern technology, such as virtual reality, augmented reality and three-dimensional panorama, shows Digital Dunhuang rich collection and offers a strong support for communicating excellent traditional culture. In the era of digital intelligence, we need to use digital intelligence technology to combine local traditional culture to create cultural products while maintaining local features (D'Silva & Atay, 2020)^[5].

2 The Literature Review

Multimodal discourse analysis is an academic focus in the branch of discourse analysis today, and a number of researchers have gradually formed three main theoretical approaches, including systemic-functional multimodal discourse analysis, multimodal interaction analysis and corpus linguistic multimodal discourse analysis and analyzed multimodal discourse in a systematic and coherent way. The first one, represented by Kress and Leeuwen (2006)^[6], starts from the perspective of social semiotics, constructs a theoretical framework and determines the scope of analysis on the basis of functional linguistic research, which is usually called "systemic functional semiotics" or "systemic functional Multimodal Discourse Analysis". Next, "multimodal interaction analysis" is mainly put forward by Norris (2002)^[7]. Last, Gu (2006)^[8] proposes "multimodal discourse analysis in corpus linguistics", which is guided by social-behavioral psychology, behavioral ecology, and perceptual ecology, to analyze and model the vivid situations of interpersonal interactions for servicing language project.

Chinese scholars' research on multimodal discourse analysis began with the introduction of foreign theories concerned. The first paper to introduce the theory of multimodal semiotics to China was "A Social Semiotics Analysis of Multimodal Discourse" (Li, 2003)^[9]. This paper detailed the visual grammar and image analysis methods constructed by Kress and Leeuwen (1996)^[10], and discussed the significance of this analytical method for deepening the understanding of the social semiotic characteristics of language, the production and understanding of multimodal discourse, and English teaching. Hu (2011b)^[11], Zhu (2007)^[12], Zhang (2009)^[13], and others reviewed the theoretical foundations, research paths, and practical significance of multimodal research, which played a macro-guiding role in promoting multimodal discourse research in China. Especially, Hu, who not only introduced the concept of "multimodal semiotics" to China and explained the differences and the relationships between multimodality and multimedia (Hu, 2007a)^[14], but also studied the genre characteristics of multimodal discourses such as PPT and multimodal skits (Hu, 2011a)^[15]. At the same time, his research is closely combined with the reality of modern education, integrating the multimodality of social practice with the requirements of humanistic education, and proposing that the "Literary" in today's world has transcended the traditional scope of "literacy," and the humanistic literacy in a pluralistic society is the "multiple intelligences" capable of reading and judging language, culture, technology, and information (Hu, 2007c)^[16].

There are still potential research in multimodal discourse analysis, namely, 1) Many scholars and experts only focus on the application of a specific modal theory, i.e., a single modal theory with additional modal forms and relations; 2) The existing multimodal discourse research has not illustrated its modal interaction features and functions in virtual cultural engineering through the construction of specific data models; 3) There are many articles on the theme of "visual culture" from the perspective of communication and journalism, but there are very few studies on the field of cultural virtual engineering based on the multimodal interaction analysis; 4) At present, there is no research on a certain research subject from the three major theoretical methods of multimodal discourse analysis, research perspectives, types of research; 5) Many scholars

study content analysis and have not made a detailed explanation of the function between visual image modalities. This article attempts to explore the multimodal forms of visual images in the Digital Dunhuang project from the three major methods of multimodal discourse analysis and analyze the interaction between visual images and their functions.

3 Research Methodology

Each modality is a set of symbolic system, and verbal modality does not always play a leading role. This study discusses the multimodal interaction analysis and adopts some of the methods of the other two types of multimodal analysis methods in the use of modality and interaction research, exploring the modal forms and relationships of each segment of Digital Dunhuang and their impact on cultural communication. Specifically, the multimodal interaction analysis is used to observe and analyze three key clips of digital Dunhuang culture, namely, the introduction clip, the highlight clip and the display clip. In addition, each clip consists of a number of high-level actions, and each high-level action consists of a number of low-level actions, i.e., modality-mediated actions.

3.1 Corpus Selection

In this article, we have recorded three key clips on the web page of Digital Dunhuang, with a total recording time of 63 minutes and 27 seconds. The three key clips are the introduction clip of Digital Dunhuang, mainly the page layout, with a recording time of 1 minute and 50 seconds; the narration clip of digital Dunhuang, with a recording time of 60 minutes and 33 seconds; and the display clip of digital Dunhuang, with a recording time of 2 minutes and 44 seconds.

3.2 Segmentation and Annotation

Human interaction has a purpose whose realization needs clips. The Digital Dunhuang virtual project can be divided into the "introduction clip", "highlight clip " and "display clip ". In these clips, it can be seen that the most exciting clip of digital Dunhuang should be the explanation clip. This clip lasted nearly more than 60 minutes, and the modal interactions reached 10 times, which can show the obvious interactions between the major modalities and the powerful visual effects of Digital Dunhuang. Following Table 1 illustrates the overview of the key clips in Digital Dunhuang.

Table 1. Overview of the Key Clips in Digital Dunhuang

Key Clips	Frequency and Time in Modal Interactions
Introduction clip in Digital Dunhuang	Eight times /one minute and fifty seconds
Highlight clip in Digital Dunhuang	Ten times /one hour and thirty-three seconds
Display clip in Digital Dunhuang	Four times /two minutes and forty-four

4 Multimodal Analysis of the Visual Images in Digital Dunhuang

Multimodal visual image is used extensively in the three clips. The analysis of multimodal visual image in this thesis adopts Rowley-Jolivet's (2002)^[17] framework for analyzing visual image modality in conference presentations, which is summarized in Table 2. Following Table 3 shows the relation of Multimodal visual image in key clips from the perspective of the Framework of Multimodal Visual Image from Rowley-Jolivet (2002).

Table 2. Framework of Multimodal Visual Image from Rowley-Jolivet (2002)

Types of Multimodal Visual Image	Means of presentation	Main Functions
Word	Characters	Text organization
Number	Equation, formula and table	Expression of abstract ideas
Image	Curve graph, schematic diagram and map	Expression of abstract ideas
Symbol	Picture and image	Text organization

Table 3. Relation of Multimodal Visual Image in Key Clips

Types of Multimodal Visual Image	Means of presentation	Main Functions	Relation of Multimodal Visual Image
Word: Verbal modality; Annotation	Characters	Text organization	Salience, priority, augmentation, independence
Number	0	0	0
Image: Image Modality	Curve graph (Figure of Buddha)	Attracting audiences	Overlap, repetition
Symbol	0	0	0
Symbol+ Image: Animated Modality; Non-verbal Modality	Webpage+ mouse icon	Vivid narration in the digital Dunhuang	Intensive, coordination, combination
Symbol+ Word+ Image: Layout Modality	Webpage+ digital Dunhuang+ Buddha	Text organization/ Attracting audiences/ Vivid narration	Explanation, expansion

4.1 Results and Discussion

Modality Sequence.

Firstly, the modalities are independent of each other. In the introduction clip of Digital Dunhuang, the sequence of modalities presented is typographic modality, non-verbal modality, image modality, and animation modality. For instance, in the introduction clip, the typographic and non-verbal modalities play a significant role in the Digital Dunhuang virtual project. From the typographic modality, we can see the bold characters "Digital Dunhuang"; next, the four characters “数字敦煌” (Digital Dunhuang) are arranged in an interwoven pattern, with the Chinese character "数" directly above the

red graphic logo of a "three-story building"; finally, "Digital DunHuang" is divided into three sections, namely, "Resource Library," "Material Library," and "Digital Cave of Scriptures," with each information bar having its corresponding English translation. Furthermore, the use of non-verbal modality in the introduction of Digital Dunhuang is exemplified by the entire page presenting a golden pattern accompanied by warm lighting, which gives a warm feeling.

Modality Type.

Secondly, the mixed type of modalities in images. Statistics demonstrate that the image modality in the introduction clip of Digital Dunhuang was used 8 times (Above Table 1). This multimodal layout achieves a relationship of enhancement and repetition between the modalities, which not only achieves the purpose of understanding the Dunhuang Grottoes from multiple aspects but also helps the audience to witness the long-standing Chinese culture in an immersive way. It can be seen that the Digital Dunhuang cultural virtual project places great emphasis on the use of various modalities of visual images. At the same time, because the focus of this clip is not on elaborating on numbers or trends, but on frequently displaying the existing images of the Dunhuang Grottoes to prepare for the communication of Dunhuang culture. Therefore, there is a lack of digital and symbolic visual image modalities, while the symbolic + textual modality is widely used. In addition, the symbolic image modality tends to be closely linked with the verbal modality to elaborate on the history of Dunhuang culture.

Modality Relation.

Thirdly, there is a harmonious coexistence among modalities. For instance, the Dunhuang mural module integrates three types of modalities: images, text, and numbers. The Dunhuang murals are represented by the image modality; the culture depicted in the murals is explained through the textual modality, and the cave numbering is conveyed by the numerical modality. This multimodal coexistence of image modality combined with textual and numerical modalities has significant cultural communication importance for interpreting and supplementing the culture of Dunhuang.

5 Multimodality and Cultural Communication in Digital Dunhuang

The Digital Dunhuang Virtual Cultural Project is an important means of cultural preservation and a significant form of cultural communication. It possesses both informative and explanatory capabilities, thus it must cater to the audience's psychology, enabling them to understand the cultural information of Dunhuang, and thereby achieve the goal of spreading Dunhuang culture. The use of multimodal visual images in Digital Dunhuang requires an approach from the theory of multimodal interaction, reflecting, displaying, and interpreting the modal forms and their interdependent and independent relationships from various perspectives.

Additionally, in the highlight clip of Digital Dunhuang, the distinctive features of

the functions promote the innovation of the Digital Dunhuang cultural virtual project such as the storage and sharing of photos, animated videos, and computer games in the mural special topics and classic murals. Therefore, they can bring the benefits to the audience, conveniently and effectively achieving the effect of cultural communication. Furthermore, in the display clip of Digital Dunhuang, the repetition of key information allows the audience to form a longer-term memory, such as the superimposition of the annotation modality and image modality of Dunhuang culture, and the use of the leading function of the non-verbal modality markers in Digital Dunhuang through the comprehensive use of various modalities and multiple emphasis, thereby imprinting the content of Digital Dunhuang culture into the minds of the audience and achieving cultural communication effect.

6 Conclusion

Today, the Digital Dunhuang virtual project is akin to a virtual fantasy world. The multimodal presentation within is much like a virtual communication tool. In the introduction clip, highlight clip, and display clip of Digital Dunhuang, the combination of visual image multimodalities is used flexibly and effectively to disseminate Dunhuang cultural information to the public, influencing them, stimulating their curiosity and desire for knowledge, while constructing and spreading Dunhuang culture, and establishing cultural confidence in China. From the highlight clip of Dunhuang culture, its modality selection and modality functions each have their own characteristics and are consistent with the purpose of the cultural activities at this clip. Different modalities work in synergy to jointly realize the meaning and purpose that Digital Dunhuang culture intends to convey. The visual image multimodality strategies employed in the display process of Digital Dunhuang are ultimately aimed at enhancing the visual effect of spreading Dunhuang culture, meeting the audience's desire for knowledge and exploration, laying a solid foundation for the communication of Dunhuang culture, and thus achieving the goal of spreading Digital Dunhuang culture beyond the national borders and to the world.

References

1. Wang, H., & Yang, C. S. 2019. The Development of Interpretation Technology in the Era of Artificial Intelligence: Concepts, Impacts, and Trends [J]. *Chinese Translation*, (6): 69-79.
2. Meng, D., & Zhou, J. X. 2019. Opportunities and Paths for the Communication of Excellent Traditional Chinese Culture by Artificial Intelligence from the Perspective of Social Participation [J]. *Fujian Forum (Humanities and Social Sciences Edition)*, (6): 41-47.
3. Hu, K. B. 2022. Language Data Science and Applied Discipline: Characteristics, Domains, and Methods [J]. *Foreign Languages World*, (3): 37-44.
4. Jia, J. Z., Cui, X. Y., & Zhang, G. X. 2024. The Impact of Artificial Intelligence Technology on Knowledge Organization: A Perspective of Knowledge Graphs [J]. *Library Forum*, (6): 41-47.

5. D'Silva, M.U. & Atay, A. 2020. *Intercultural Communication, Identity, and Social Movements in the Digital Age* [M]. Routledge: New York, London: xii +190.
6. Kress, G. 2003. *Literacy in the New Media Age* [M]. London: Routledge.
7. Norris, S. 2009. Modal density and modal configurations: Multimodal actions. In C. Jewitt (Ed.). *The Routledge handbook of multimodal analysis*[M]. London: Routledge.
8. Gu, Y. G. 2006. "Role-Based Modeling Language (AML) I: Dynamic Behavior Modeling," in The Office of the Informatization Leading Group of the Chinese Academy of Social Sciences / Computer Network Information Center (Ed.), *Proceedings of the 20th CODATA International Conference*.
9. Li, Z. Z. 2003. A Social Semiotics Analysis of Multimodal Discourse [J]. *Foreign Language Research*, (5): 1-8+80.
10. Kress, G., & van Leeuwen, T. 1996. *Reading Images: The Grammar of Visual Design* [C]. London: Routledge.
11. Hu, Z. L. 2011b. My Reading of "Current Status and Development Trends of Systemic Functional Linguistics Research" [Z]. Speech at the 12th National Symposium on Systemic Functional Linguistics, Nanjing.
12. Zhu, Y. S. 2007. The Theoretical Foundations and Research Methods of Multimodal Discourse Analysis [J]. *Foreign Language Journal*, (5): 82-86.
13. Zhang, D. L. 2009. Exploring a Comprehensive Theoretical Framework for Multimodal Discourse Analysis [J]. *China Foreign Languages*, (1): 24-30.
14. Hu, Z. L. 2007a. Multimodality in Social Semiotics Research[J]. *Language Teaching and Research* (1): 1-10.
15. Hu, Z. G. 2011a. The Dominant Mode in Multimodal Sketches [J]. *Foreign Language Teaching*, (4): 1-5.
16. Hu, Z. G. 2007c. *From Literacy to Multiple Intelligences* [C] // Beijing Forum (2007) Harmony and Common Prosperity of Civilizations—Diverse Development Patterns of Human Civilization: "Language Identity and Change in the Conflict and Integration of Multiple Civilizations" Foreign Language Sub-Forum Proceedings and Abstracts, (4): 84-101.
17. Rowley-Jolivet, E. 2002. Science in the making: scientific conference presentations and the construction of facts. In E. Ventola *et al.*(eds.), *The language of conferencing*[M]. Peter Lang: Frankfurt am Main.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

