



## **A Multimodal Discourse Analysis Of The Tourism Promotional Video "Xi'an China."**

**Jiao Yanwei,<sup>1\*</sup>**  
**Hanita Hassan<sup>2</sup>**

<sup>1\*,2</sup> Universiti Teknologi Malaysia, Malaysia, E-mail: \*jiao@graduate.utm.my, hanita@utm.my

**\*Corresponding Author:-** Jiao Yanwei

\*Universiti Teknologi Malaysia, Malaysia, E-mail: jiao@graduate.utm.my

### **Abstract**

Tourism promotional videos are a popular way to inform potential tourists about tourist destinations, especially given that people can now watch a variety of videos on various online platforms. Multimodal discourse analysis has in recent years emerged as an effective tool for analyzing different mediums of communication, including videos and images. The aim of this study was to investigate the ideational and representational meaning of the tourism promotional video *Xi'an China* and explore the relationship between the verbal and visual elements from the perspective of multimodal discourse analysis. The study made use of a qualitative analysis research design, and specifically employed: a transitivity analysis on the verbal data based on Systemic Functional Linguistics; an exploration of the representational meaning of the visual data based on Visual Grammar; and Royce's Intersemiotic Complementarity to analyze the relationship between the verbal texts and visual images of the *Xi'an China* promotional video. The data obtained from these processes were analyzed using NVIVO 12. The study found that the video's verbal modality conveyed information related to the history, culture, landscape, trade, industry and development of Xi'an; the visual modality represented the natural beauty, entertainment, events happening, and changes in Xi'an from the past to present time; thirdly, the most frequently used semantic relations between verbal and visual modes were found to be repetition and synonymy. The cooperation of the two modes helped to construct the overall meaning of the *Xi'an China* TPV, thereby publicizing the tourist appeal of Xi'an.

**Keywords:** multimodal discourse analysis; tourism promotional video (TPV); verbal and visual elements; Xi'an China.

### **Introduction**

According to the China Statistical Abstract issued by the National Bureau of Statistics of China (2020), the number of people who travel to China from overseas rose from 13 million to 14.12 million between 2008 to 2018, revealing the growing interest in China's tourist attractions internationally. Ever since China opened its doors to the world in 1978, the tourism industry has brought significant investments into China. Since promotional videos are becoming a popular way of providing information to potential tourists with the emergence of video-sharing websites, potential visitors to the country can watch a wide variety of tourism promotional videos created by the national government, local government, and travel agencies (Aoki & Fujimoto, 2021). Many people are more inclined to get information about tourist attractions from promotional videos because, as Hou (2017) reported, promotional videos are very effective, as visual codes are organized in such a way that appealing visual components are highlighted.

As the capital of Shaanxi Province, and the ancient capital of thirteen dynasties, Xi'an is the birthplace of Chinese civilization and the Chinese nation. In 2018, Xi'an was listed in the "Top 10 Most Popular Tourist Destinations in China" based on statistics on the *Top China Travel* website, which attracts tourists from around the world. In December 2017, the Tourism Administration of Shaanxi Province released the tourism promotional video (TPV) *Xi'an China*. With its picturesque images, attentive soundtracks, ultra-long time-lapse photography, and

moving commentaries, the TPV impressed the audience greatly (Wang & Feng, 2021). On YouTube, the TPV has received 106268 views as of October 2022, which shows its success in attracting the public's attention from around the world. Combining both verbal and visual elements, the TPV *Xi'an China* vividly displays the history, culture, trade, natural scenery and modernity of Xi'an.

Among online tourism promotion advertisements, TPVs are one of the most popular forms of spreading tourism information, depicting appealing images of tourist destinations thereby encouraging potential tourists to select the destinations displayed in the videos. According to (Fong et al., 2017), tourists can vividly visualize destinations through TPVs, which enables people to get the audio-visual information of a destination. The scene of a place viewed in a video often inspires a viewer's willingness to travel to that destination and their intention to participate in the activities depicted there (Losada & Mota, 2019). Therefore, many researchers focus on the multimodal features of tourism videos.

Jocuns (2020) analyzed the mediated actions and modalities used in tourism videos about Thainess, with a particular focus on the discursive stances of Thainess. In their research, Yao & Zhuo (2018) discussed the meaning making process of the TPV Hangzhou from the perspective of Systemic Functional Linguistics and Visual Grammar. Graça et al. (2022) similarly adopted the Systemic Functional Linguistics and Visual Grammar framework to analyze the ideational and representational meaning construction of the promotional video *Wonderful Indonesia*. The findings of their research show that mental processes dominate when making ideational meaning while action processes, reactional processes and symbolic processes are widely used in representational meaning construction.

According to Fong et al (2017), TPVs use different forms of language, a variety of images, and emotive background music to spread awareness effectively, which increases the audience's receptiveness to advertising information. The verbal texts of tourism promotional videos convey specific information to the audience; visual images replicate the meaning conveyed by verbal texts more vividly, and combined with visual images, background music stirs up the audience's emotions and resonance. Wang & Han (2020) observed that TPVs have a positive impact on young people with respect to China's image as a travel destination as they help to showcase the eye-catching tourist destinations in the country, particularly the cultural and ecological aspects. By analyzing the verbal and visual modes of TPVs respectively, the multimodal characteristics of TPVs can be fully grasped.

Verbal texts of tourism promotional videos convey specific information to the audience; visual images convey the meanings of verbal texts more vividly (Fong et al, 2017). By analyzing the verbal and visual modes of the TPV *Xi'an China*, its multimodal characteristics can be fully grasped. In terms of verbal language, the ideational function of language in tourism discourse plays a vital role in persuading potential tourists to choose destinations. The representational meaning of visual images in tourism discourse work effectively represents the beauty of destinations and attracts viewers' attention. As such, this study will concentrate on the ideational function of verbal language and the representational meaning of visual images in the TPV *Xi'an China* and shall explore the relationship between the two modes.

## **Aim**

The aim of this study is to investigate the ideational and representational meaning of the tourism promotional video *Xi'an China* and explore the relationship between the verbal and visual elements from the perspective of multimodal discourse analysis.

## **Research Problem**

Despite the importance of TPV's in promoting tourism for China to the outside world, very few studies have been carried out that focus exclusively on critically analyzing these videos, particularly from the perspective of a multimodal discourse analysis. The current study is thus vital, as it helps to close this research gap.

## **Multimodal Discourse Analysis**

In recent years, with the development of discourse analysis, the role of various symbols in discourse, and their relationship in communication has drawn the attention of more scholars. We live in a multimodal society, and the construction of meaning increasingly depends on the integration of various symbolic mediums (Shin, Cimasko, & Yi, 2020). With the emergence of

computers, multimedia and media technology, multimodal discourse has become prominent in peoples' life, which combines image, sound, text, color, space and action. In this sense, discourses nowadays are no longer monomodal but multimodal.

Liu (2019) defines multimodal discourse as a form of discourse that combines multiple communication modalities so as to convey information. According to Pratiwy & Wulan (2018), multimodal discourse is a term used to describe a new type of analysis in which meaning is constructed by combining words, typography, color, sound, image, and other semiotic elements. Hu & Luo (2016) discussed the features of multimodal discourse in communicative situations, noting that most meaning making in discourse is realized through non-linguistic factors, such as accompanying language features (loudness, tone, intonation, speed of sound, etc.), physical features (gestures, posture, facial expressions, movements, etc.), and non-physical features (PPT, audio equipment, network, laboratory, environmental factors, etc.). Discourse's multimodality has emerged as a key component of the numerous discourses utilized for communication throughout the world. Meaning construction is realized through different modes in communication and the language use is implemented through different communication modes, which indicates that all discourses are multimodal in a sense (Liu, 2019).

The study of multimodal discourse is essentially concerned with the design, production, and dissemination of multimodal mediums within a social context. Multimodal discourse analysis (MDA), on the other hand, examines how various semiotic modes interact to generate a cohesive text or communication experience. According to Liu (2019), MDA is the examination of a text or communication event through the analysis of some or all of its different semiotic modalities. The concept of MDA is that textual structure is realized visually in many areas of modern literature, both at the "clause" and "discourse" levels, via layouts, colors, and fonts. MDA overcomes the constraint of linguistic discourse analysis and encourages the spontaneous integration of conventional and single language modes of communication with vision, hearing, and gestures. Lim (2021) explains that MDA addresses the ways in which diverse semiotic systems and multimodal mediums complement one another in meaning construction. Though different scholars have defined MDA in various ways, MDA generally deals with issues related to linguistic and other semiotic modes, which helps people acknowledge the multimodal ways of communication.

Multimodal discourse analysis has been widely used in social sciences, arts and humanities, and psychology. In general, Systemic Functional Linguistics proposed by Halliday, and Visual Grammar proposed by Kress and van Leeuwen are the most frequently used theoretical foundations for Multimodal Discourse Analysis. Moreover, when studying the relationship between the verbal and visual modes, Royce's Intersemiotic Complementarity framework is frequently used.

### **Systemic Functional Linguistics**

Halliday's Systemic Functional Linguistics has inspired many scholars worldwide (Almurashi, 2016). Being a widely applicable and universal theory of linguistics, Systemic Functional Linguistics has been expanded to include the co-deployment of different forms of communication (including language) as meaning-making mediums (Xin, 2011). The main theoretical foundation for Multimodal Discourse Analysis (MDA) is the Systemic Functional Linguistics (SFL) created by Halliday.

As Halliday pointed out, systemic functional theory is ideally suited to give theoretical tools for MDA since it is a social semiotic theory in which meaning is understood as context-dependent (Almurashi, 2016). The metafunctional ideas of Halliday serve as the primary basis for systemic functional MDA. Halliday characterized language's meta-functions as ideational function, interpersonal function, and textual function, which enables analysis of how semiotic mediums combine to produce meaning (Ammara, Anjum, & Javed, 2019). According to him, the ideational metafunction serves as a medium for experience representation—our experience of the world around us as well as our experience of the world inside us (the world of our imagination). It has significance in the sense of achieving substance where the interpersonal metafunction serves as a medium for meaning as a mode of action where the speaker or writer does something to the listener or reader via language. The textual metafunction is responsible for ensuring relevance to context, both the previous (and subsequent) text and the situational context.

To be more specific, MDA accepts the view from SFL that language is socially semiotic and has meaning potential, and that other semiotic systems can also be sources of meaning.

Furthermore, the systemic theory is adopted in MDA given that multimodal discourse is also systematic. MDA additionally approves of the metafunction hypothesis of SFL and holds that multimodal discourse also has ideational, interpersonal, and textual functions just as discourse, containing only linguistic symbols. MDA further accepts the register theory of SFL, arguing that there is an inseparable connection between context and the interpretation of the meaning of multimodal discourse (Almurashi, 2016).

The ideational function consists of the experiential function and logical function; the experiential function refers to language's function of expressing people's experiences in the world while the logical function expresses the logic between two or more meaning units (Ammara, Anjum, & Javed, 2019). In this study, only the experiential function is adopted. Since the experiential function is realized through transitivity and voice, this study focuses on the transitivity analysis of the TPV under review. Transitivity refers to "a network of systems whose point of origin is the major clause" (Matthiessen, 2018). Transitivity systems are about "the type of process expressed in the clause with animate and inanimate participants in the process, diverse attributes and circumstances of the process" (Sutrisna, 2020). Transitivity is a meaning system, which categorizes human's activities into different processes and reveals the participant and circumstantial elements in the processes. In general, there are six processes in the transitivity system, namely, behavioral processes, existential processes, verbal processes, and material processes, mental processes, relational processes (Zein et al, 2019).

### **Visual Grammar**

Kress and Van Leeuwen proposed the Visual Grammar (VG) concept deriving from Systemic Functional Linguistics, inspired by Halliday's social semiotic concept (He, 2017). It is the first academic study to provide a systematic and comprehensive analysis of visual semiotics, and has exerted a significant influence on subsequent research on visual modality as a classic work. They believe that all visual semiotic mediums in communication such as the gaze, angle, composition, and frame size can convey three meanings, namely: a representational meaning, an interactive meaning, and a compositional meaning, which correspond to the ideational, interpersonal, and textual meanings in the Systemic Functional Grammar (SFG) respectively. They define representational meaning as the portrayal of the sensory world, interactive meaning as the interaction between the people shown in a visual design and their viewers, and compositional meaning as the arrangement of visual mediums (Ammara, Anjum, & Javed, 2019).

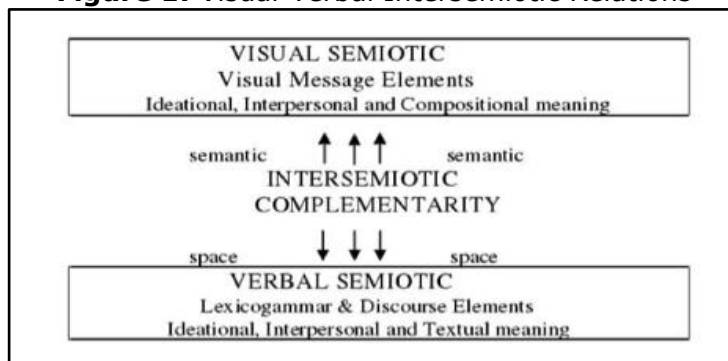
Pictures may be seen as possible mediums for meaning, which expands Halliday's social semiotic view of language. They assert that the three meta-functions of language are applicable to all kinds of representation, including color, music, and typography. According to Machin (2016), image is one of the semiotic mediums available in order to achieve the three meta-functions in terms of the visual mode. With respect to the ideational metafunction, visual modes such as image may communicate ideas, events, and their relationships in a world unencumbered by the verbal representational system. In visual communication activities, there are several ideational options accessible for visual meaning making. In terms of interpersonal metafunction, images and other visual modalities may construct a fictitious social relationship between the image-maker, the observer, and the represented item. From a textual metafunctional standpoint, images and other visual modalities may combine to create unified texts. As a whole, the visual modalities are cohesive, and the unified texts are consistent with their context of production.

After the establishment of Visual Grammar based on SFG, the theoretical basis of studies on visual semiotics became more solid. In addition, Visual Grammar has been widely used in MDA combined with Systemic Functional Grammar to analyze verbal and non-verbal elements in multimodal discourses, which gives rise to various frameworks discussing the relationship between verbal and non-verbal language (Hu & Luo, 2016). The following part will discuss one of the theories, which focuses on the intersemiotic relationship between verbal and visual modes.

### **Intersemiotic Complementarity**

The relationship between images and texts in multimodal discourse is the key to interpreting multimodality, which has become a trend for multimodal discourse analysis in recent years. Based on the relations between images and texts in multimodal discourses, Royce (2015) proposed visual-verbal intersemiotic relations for multimodal discourse analysis, which is shown below:

**Figure 1:** Visual-verbal Intersemiotic Relations



Source: Royce (2015)

Royce believes that existing multimodal theories only focus on the meanings expressed by images and they pay less attention to the complementary relationship between images and texts in multimodal discourses. Therefore, he proposed a theoretical framework of intersemiotic complementarity to explain the complementary relationship between images and texts in multimodal discourses. Intersemiotic complementarity refers to the relationship between verbal modes and visual modes in multimodal discourse that coordinate with and complement each other to construct meanings. Both verbal and visual modes are mediums for expressing meaning, which jointly present readers with a coherent multimodal discourse (Royce, 2015).

The theoretical foundation for intersemiotic complementarity is the Systemic Functional Linguistic view of language as being "social semiotic," which considers language to be functionally, semantically, and contextually semiotic (Hu & Luo, 2016). The analysis of verbal modes in intersemiotic complementarity is based on the three metafunctions of language in Halliday's Systemic Functional Linguistics, while the analysis of visual modes is based on Kress and Van Leeuwen's Visual Grammar. The intersemiotic complementarity analysis of the multimodal discourse is based on the ideational, interpersonal and compositional features of the multimodal discourse.

In terms of intersemiotic ideational features, the visual message elements (VMEs) in the image should be recognized, which includes the participants, ongoing processes/activities, circumstances, etc. VMEs are symbolic mediums for expressing meaning and realizing the ideational function of representing information and experience (Royce, 2015). Then, after the VMEs are identified, the co-occurring text is analyzed to find words in the text that are related (similar or opposite) to the meaning of the visual message elements, and the corresponding vocabulary is obtained. Finally, from the perspective of the lexical cohesion theory (Alyousef, 2016), the visual message elements and the words associated with them in the text are analyzed to determine the intersemiotic complementary relationship between the two. The connection between image and text includes the following aspects: repetition, synonymy, antonymy, hyponymy, meronymy, and collocation (Royce, 2015).

## Methods

### Research Design

This study adopts a qualitative research design, analyzing the verbal texts and images of the *Xi'an China* TPV from the perspective of a multimodal discourse analysis. Qualitative research is beneficial for such a study, as it helps to document the details of events and most critically, identifies meaning. The range and frequency of actions and meaning perspectives are essential for good qualitative research (Papakitsou, 2020). Therefore, this study focuses on the qualitative data from the *Xi'an China* TPV and considers the frequency of certain actions and meaning perspectives.

### Process

For the verbal data in the videos, a transitivity analysis will be conducted based on Systemic Functional Linguistics. Transitivity is a meaning system that divides human actions into processes and identifies the participants and situational components in each phase. In this study, all the relevant processes (material, mental, relational, behavioral, verbal and existential) are identified to find out which processes are frequently used in tourism promotional videos.

For the visual data of the videos, the representational meaning of the frames is explored based on Visual Grammar. Regarding the representational meaning of the videos, narrative representation and conceptual representation are analyzed to examine the frequency of different processes, like action process, reactional action, speech process and mental process, and conversion process for narrative representation and classification process, analytical process and symbolic process for conceptual representation.

In terms of the relationship between the verbal texts and visual images of TPVs in promoting China's tourist destinations, Royce's Intersemiotic Complementarity is adopted to analyze the relationships.

### Sample

Given that the study focuses on analyzing text and visual images rather than collecting data from human participants, no sampling process was carried out. Rather, the study focuses on the semantic intersemiotic complementarity of the verbal texts and visual images of the ten tourism promotional videos.

### Analysis of Data

The current study employs NVIVO 12 to analyze data, as NVIVO enables the researcher to manage both the verbal and visual data. NVIVO 12 is appropriate for this study because of its functions in organizing and categorizing different types of data into groups by coding and calculating the frequency.

The study collected verbal and visual data from the *Xi'an China* TPV released on YouTube which is 8-mins-35-seconds long. The TPV displays every aspect of ancient and modern Xi'an. Therefore, it is selected to explore the cooperation of both verbal and visual elements in construing meaning. The verbal data, i.e., subtitles and captions, were obtained through Nvivo by transcribing the subtitles of the video. At the same time, the visual data are the frames of the video obtained from the software Premier.

### Findings

#### Ideational Function Analysis of the Verbal Mode of the TPV *Xi'an China*

Ideational function is realized through the experiential function and the logical function. The experiential function refers to the expression of language in terms of what people experience in the world around them and what people experience in the world of consciousness and imagination. In other words, it involves the people, the objects, and the circumstance associated with them such as time, place, and other factors. Logical function is the expression of the logical relationship between two or more units of meaning. For the experiential function, it is mainly realized through transitivity and voice, and a transitivity system construes the world of experience into a manageable set of process types. Participant, process, and context are the three main elements of transitivity. The six different processes include material process, mental process, relational process, verbal process, behavioral process, and existential process, which can be used to categorize all the activities and events described by clauses in terms of experiential function.

**Table 1.** The Frequency of Six Processes With in Transitivity System

Process Type	Number of Processes	Percentage
Material Process	14	31.82%
Mental Process	8	18.18%
Relational Process	22	50%
Verbal Process	0	0
Behavioral Process	0	0
Existential Process	0	0
Total	44	100%

This study collected the verbal data from the TPV *Xi'an China*, which contains 44 clauses to determine how each process functions in the TPV. The main form of verbal language in this TPV is the subtitle. After transcription, there were 44 clauses whose content covers history, culture, landscapes, trade and industry. Among the 44 clauses, three main types of process were identified. From the data in Table 1, 31.82% of the clauses in this TPV belong to the material process, 18.18% to mental process, and 50% to relational process. Verbal process, behavioral

process and existential process were not found in this TPV. Given that relational process accounts for the largest proportion among all the processes, it indicates that relational process is dominant in this TPV.

The following parts are the discussions about different process types appearing in this TPV:

**(1) Material Process**

Material process is the process of doing in which there is an Actor, Process and Goal. Material process is a series of events which people do, or events that happen in the world around people or in the world of consciousness and imagination. The clause can either be in active or passive voice. In this TPV, material process accounts for 31.82%, which is used to describe the natural scenery, achievements, and development of Xi'an. The following three are examples that show the material process:

**Example 1**

China Railway Express to Europe, along the Silk Road Economic Belt	links	Europe and Asia continents.
Actor	Process	Goal

**Example 2**

Singling mountains dotted with historic relics	houses	gorgeous natural scenes, humble mountains, modest trees, clouds and mists, fallen leaves, streams and Brooks, flying birds.
Actor	Process	Goal

**Example 3**

Huayin shadow play with Accent from the depth of the soul, the most incredible sound of faith	captivates	you.
Actor	Process	Goal

In Example 1, "China Railway Express to Europe" is the Actor in this material process, "links" is the process, and "Europe and Asia continents" are the Goal. This clause states the importance of Xi'an in connecting the trade and culture of Europe and Asia. Example 2 describes the beautiful natural landscape of Xi'an with "humble mountains, modest trees, clouds and mists, fallen leaves, streams and Brooks, flying birds," which impresses potential tourists most and attracts them to have a visit to Xi'an. Example 3 shows the charm of the traditional Chinese craft *shadow play* and introduces the folk musical performance *Huayin Accent*, which "captivates" potential visitors.

**(2) Mental Process**

Mental process is the process of sensing, including the process of perception, reaction and cognition. In a mental process, there are two participants, one is a Senser and the other is the Phenomenon (Halliday, 1994: 113). The Sensor can be people or other personified things, while the Phenomenon can be people, objects, abstract things, or events. In this TPV, mental processes comprise 18.18% of all the process types. It is used to show people's feelings towards Xi'an and the features of this ancient city.

The following two are examples of mental processes.

**Example 4**

You	can still feel	strength brought by history.
Senser	Process	Phenomenon

**Example 5**

You	would find	yourself enchanted in this city where history and modernity meet.
Senser	Process	Phenomenon

In Example 4, "You" is the Senser, the Process "feel" is the perception, and "strength brought by history" is the Phenomenon. This clause displays the charm of history. In example 5, "You" is the Senser, the Process "would find" is the cognition, and the Phenomenon of this mental

process is a fact. This clause indicates that Xi’an is a city with both historic and modern features. The two examples show that Xi’an is an attractive tourist destination that is worth visiting.

**(3) Relational Process**

The relational process can be interpreted as the process of being, but different from existing in the existential process. It is the relation between two participants involved in a clause. There are three types of relational process, namely intensive, circumstantial and possessive, while there are two modes of relational process, namely attributive and identifying. With three types and two modes, there are six categories of relation process, i.e., intensive attributive clause, intensive identifying clause, circumstantial attributive clause, circumstantial identifying clause, possessive attributive clause, and possessive identifying clause (Sembiring, 2020). In the attributive relational process, the subject is marked as Carrier and the object or predicative is marked as Attribute. In the identifying relational process, the two counterparts in the clause are marked as Identified and Identifier respectively. In this TPV, the relational process accounts for 50%, constituting the largest proportion amongst all. Four categories of relation process appear in this video, namely intensive attributive clause, intensive identifying clause, circumstantial attributive clause, and possessive identifying clause.

The following are examples of the four categories:

**Example 6**

food	is	tastefully beyond the grasp of your taste buds.
Carrier	Process (intensive)	Attribute

**Example 7**

The largest Asian Aviation City and China’s Aerospace City	are situated	here.
Carrier	Process (circumstantial)	Attribute

**Example 8**

Xi’an, before the 10th century,	was	the center of Eastern civilization and the starting point of the silk road.
Identified	Process (intensive)	Identifier
Token		value

**Example 9**

It	is	an integrated transport hub with airlines, railways and roads converging seamlessly.
Identified	Process (intensive)	Identifier
Token		Value

**Example 10**

Persons with higher education	comprises	about one-tenth of Xi’an’s population.
Identified	Process (possessive)	Identifier
Token		Value

Examples 6 and 7 are attributive clauses. In the attributive mode, an entity has some quality ascribed or attributed to it. This quality is labeled as the Attribute, and the entity to which it is ascribed is the Carrier. Example 6 is an intensive attributive clause. In Example 6, “food” is the Carrier, and the quality “tastefully beyond the grasp of your taste buds” is the Attribute. In this clause, the meaning is similarly a member of the class of tasty food, so the process is intensive. Example 7 is a circumstantial attributive clause with “The largest Asian Aviation City and China's Aerospace City” as Carrier and “here” as the Attribute. When the verb in a clause encodes the circumstance of time, place, accompaniment, manner etc. as a relationship between the participants, the process is circumstantial. In Example 7, “are situated” encodes the circumstance of place, so the process of it is circumstantial.

Examples 8, 9 and 10 are identifying clauses. “In identifying mode, one entity is being used to identify another: ‘x is identified by a’ or ‘a serves to define the identity of x’” (Halliday, 1994, p. 121). Furthermore, in any identifying clause, one entity is labeled as Token and the other Value. Examples 8 and 9 are intensive identifying clauses. In Example 8, “Xi’an” is the Identified (Token), and “the center of Eastern civilization and the starting point of the silk road” is the Identifier (Value). While in Example 9, “it” (Xi’an) is the Identified (Token) and “an integrated



transport hub with airlines, railways and roads converging seamlessly” is the Identifier (Value). Example 10 is a possessive identifying clause, with “Persons with higher education” as the Identified (Token) and “about one-tenth of Xi’an’s population” as the Identifier (Value). As the verb “comprises” in this clause indicates the possessive relation between the Identified and the Identifier, the process is possessive.

From the example above, Example 6 introduces the food culture of Xi’an, which is one of the highlights of Xi’an culture. Example 7 shows the charm of Xi’an with its booming industry. Examples 8 and 9 showcase the advantages of Xi’an’s location. Further, Example 10 states that Xi’an is a city with talents. To summarize, the relational process in this video describes the importance of Xi’an both in the past and in modern times, which displays the culture, trade and history of Xi’an as a tourist destination. In this TPV, identifying the relational process weighs more than the attributive one, which makes the content more convincing.

**Results and Discussion**

**Representational Analysis of the Visual Mode of the TPV *Xi’an China***

Just like the ideational function of language in SFL, in the representational process, the “participants” in visual images can also depict the experiential world and represent human’s mental activities. In visual grammar, representational meaning is realized by narrative representation and conceptual representation (Hong & Duan, 2020).

In narrative processes, a vector connects participants while there is no vector in conceptual processes. Vectors represent different processes, such as action process, reactional action, circumstance, speech process and mental process, and conversion process. For conceptual process, no vector linking visual elements together exists, though it represents participants with respect to their more generalized, stable and timeless essence when it comes to class, meaning or structure (Hu & Luo, 2016), and it can be realized through a classificational process, analytical process, and symbolic process.

To analyze the representational meaning of the TPV *Xi’an China*, the video was imported into Adobe Premier and 197 frames were selected for further analysis. Then, the frames were imported into Nvivo for process coding. The table below shows the frequency of different processes:

**Table 2.**The Frequency of Narrative Process and Conceptual Process

Representational Process		Number	Percentage
Narrative Process	Action process	76	38.6%
	Reactional process	27	13.7%
	Circumstance	10	5.1%
Sub-total		113	57.4%
Conceptual Process	Classification process	40	20.3%
	Analytical process	38	19.3%
	Symbolic process	6	3%
Sub-total		84	42.6%
Total		197	100%

From the figures in Table 2, the proportion of the narrative process (57.4%) in the *Xi’an China* TPV outweighs the proportion of conceptual process (42.6%). There are a total of six sub-categories of process in this TPV, namely: action process, reaction process, circumstance, classificational process, analytical process, and symbolic process. For the narrative process, the action process takes up the largest proportion (38.6%), followed by the reactional process (13.7%) and Circumstance (5.1%). For the conceptual process, the classification process accounts for 20.3%, followed by the analytical process (19.3%), and the symbolic process (3%). The following are some examples for each process:

**(1) Narrative Process**

The function of narrative processes is to depict actions and events as they occur, as well as processes of change and momentary spatial arrangements. In this video, the narrative process takes up a larger proportion compared with the conceptual process, indicating that the biggest part of the video represents what is happening in Xi’an, the change of Xi’an as an ancient and modern city, and the settings of events happening in Xi’an.

*Action Process*

In the action process, there are: 'actor,' 'vector,' and 'goal.' The Actor is the participant, which produces the vector, while the vector directs to the Goal. When only an Actor is present in the process, this action is called a non-transactional action process (Hu & Luo, 2016). When there is an 'Actor' and a 'Goal,' the process is a transactional action process. Frames 1, 2, 3, and 4 are transactional action processes. The tourists and local Xi'an people are the Actors, while their activities (like flattening out the dough, boating, striking the bell, playing shadow puppets, and playing stringed and percussion musical instruments) represent their actions, which contain the Vector and the Goal. Frames 5 and 6 are non-transactional action processes. The university students and tourists are the Actors, while walking and jumping represent their actions with no Goals.



Frame 1



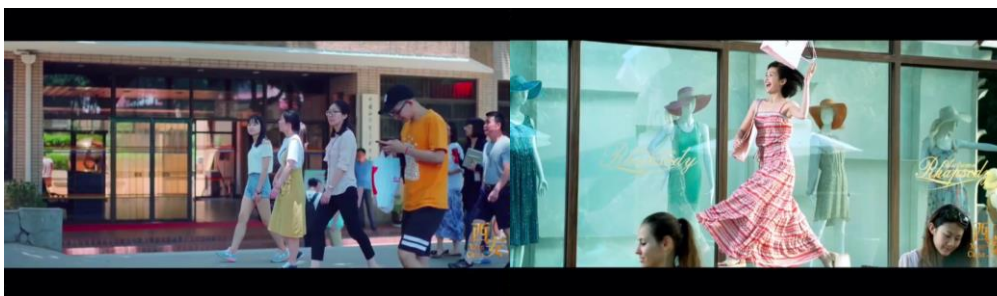
Frame 2



Frame 3



Frame 4



Frame 5



Frame 6

**Reactional Process**

When the vector in the process is formed by an eyeline or glance, the process is reactional in which there are Reactors and Phenomena. Likewise, reactional processes can be transactional and non-transactional. In the transactional process, there are both Reactors and Phenomena, while in the non-transactional process, there is no Phenomena (Ansori & Taopan, 2019). Frames 7, 8, 9, and 10 are a transactional relational process, with the audience, the lady, and the family as the Reactors and the shadow play, Huayin Accent performance, the jewelry, and the Mermaid Show as the Phenomena. Their eyeline and glance is the Vector. While frames 11 and 12 are non-transactional processes, with the little boy and the girl as the Reactors and no Phenomena. The reactional process in this TPV depicts colorful entertainment activities in Xi'an, which give tourists pleasure and unforgettable experiences.



Frame 7

Frame 8



Frame 9

Frame 10



Frame 11

Frame 12

### **Circumstance**

For narrative representation, circumstances are also important. In narrative images, participants are connected not only by vectors but also by other participants. These participants are known as 'Circumstances,' and include Locative Circumstances (Settings), Circumstances of Means and Circumstances of Accompaniment (Royce, 2015). In this TPV, there are two types of Circumstance, namely Locative Circumstances and Circumstances of Means. For instance, Frame 13 is a Locative Circumstance, with a contrast between the man clapping his hands and other audiences. The man in the frame functions as a foreground and the other audiences function as the background. Frame 14 is the Circumstance of Means, with the spatulas in this frame functioning as the tools, which realize the action process of frying pancakes. There is no clear vector between the tools and their users.



Frame 13

Frame 14

**(2) Conceptual Process**

According to Ly & Jung (2015), the conceptual process describes the property and quality of static objects, which is different from the dynamic narrative process. In this video, the conceptual process accounts for 42.6%, which shows the natural beauty, cultural spectacles, the historic heritage and modern civilization of Xi'an.

**Classificational Process**

In the classification process, participants are related to each other through taxonomy, one set of participants being subordinate and the other set superordinate. There are two types of taxonomy, one is covert taxonomy and the other is overt taxonomy. In covert taxonomy, the superordinate can be inferred from similarities and the subordinate are arranged in a symmetrical composition with the same distance, size and orientation. In overt taxonomy, participants are chained, and the intermediate participant will be both the superordinate and subordinate of other participants (Lim, 2019). In this TPV, only covert taxonomies exist. In Frame 15, the terracotta sculpture is the superordinate and the sculptures arranged in the frame are the subordinates. In Frame 16, the Bronze Mirror in Han Dynasty is the superordinate and the bronze mirrors displayed are subordinate. In Frames 17 and 18, the Superordinate is traditional Chinese architecture and the subordinates are the buildings shown in the frames.



Frame 15



Frame 16



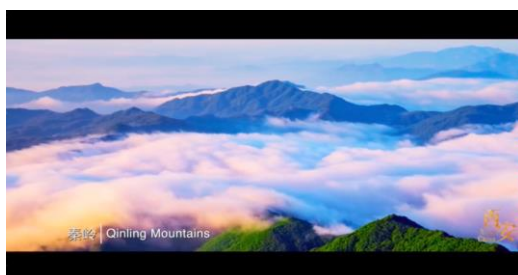
Frame 17



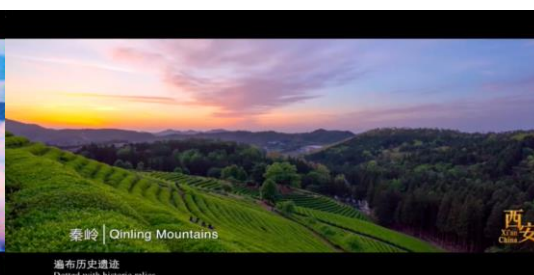
Frame 18

**Analytical Process**

In the analytical process, the relation between participants is that of 'part and whole,' with the Carrier (the whole) and the Possessive Attributes (the parts) (Ly & Jung, 2015). In Frames 19 and 20, Qinling Mountains is the Carrier and the Possessive Attributes are different parts of it. In Frames 21 and 22, Cuihuashan National Geopark is the Carrier and the mountains and the lake are the Possessive Attributes. In the four frames, the Carriers of the Analytical Process are labeled in the captions.



Frame 19



Frame 20



Frame 21

Frame 22

**Symbolic Process**

The symbolic process is concerned with what a participant means or is. When there are carrier and symbolic attributes, the process is called a symbolic attribute. It is called a symbolic suggestive process when there is only one carrier (Ly & Jung, 2015). In this TPV, both processes are included. Frame 23 is a symbolic at tributive process. The Stock Market Board is the Carrier and the red price numbers on it are the symbolic attributes. The price index in red symbolizes the booming Xi’an economy because red numbers in the context of China’s stock market means that the price is going up. Frame 24 is a symbolic suggestive process since there is only one Carrier—the train. China Railway Express to Europe symbolizes the trade connection between China and Europe.



Frame 23

Frame 24

**Ideational Intersemiotic Complementary Analysis of the TPV Xi’an China**

For an ideational intersemiotic complementarity analysis, Visual Message Elements (VEMs) that include Identification, Activity, Circumstance, and Attributes are considered (Royce, 2015). To analyze the relationship between the VEMs and verbal elements in multimodal discourse, Royce proposed six semantic intersemiotic complementary relations, namely: repetition, synonymy, antonymy, meronymy, hyponymy, and collocation. The semantic intersemiotic complementary relations appearing in this TPV are repetition, synonymy, meronymy, and hyponymy. Antonymy and collocation are not included in this TPV. A total of 96 frames containing both verbal and visual elements are selected for analysis. The following table shows the frequency of different semantic intersemiotic complementary relations:

**Table 3** The Frequency of Semantic Relations Between Verbal and Visual Modes

Semantic Relations	Number	Percentage
repetition	36	37.5%
synonymy	26	27.1%
meronymy	20	20.8%
hyponymy	14	14.6%
Total	96	100%

Based on the data shown in Table 3, repetition relation accounts for 37.5%, followed by synonymy (27.1%), meronymy (20.8%), and hyponymy (14.6%), which indicates that repetition and synonymy are the main semantic relations in this TPV. The frequent use of repetition and synonymy shows the complementation of verbal texts and visual images to construe meanings.

The following is an example of an ideational intersemiotic complementarity analysis. In the table, "IC" stands for "Intersemiotic Complementarity, and letters of R, S, M, and H represent the semantic relation of repetition, synonymy, meronymy, and hyponymy respectively.



Frame 25

Frame 26



Frame 27

Frame 28



Frame 29

Frame 30

Table 4

Visual Part	IC	Verbal Part
mountains	R	Entrancing scenery that boasts magnificent <b>mountains</b> and <b>rivers</b> (Frame 25)
clouds	/	
a river	R	
city	R	A livable <b>city</b> for your <b>family</b> and career. (Frame 26)
an engaged couple	H	
the gate of an old building (Frame 27)	M	<b>Old buildings and steles</b> bearing footprints of history tell secrets that are written and beyond words (Frame 27 & Frame 28)
a stele in a park (Frame 28)	R	
the city wall of Xi'an	H	Amid majestic <b>ancient buildings</b> (Frame 29) the <b>hustle and bustle</b> goes on (Frame 30)
cars	S	

Frame 25 depicts the beautiful natural scenery with mountains, clouds, river and trees. The "mountains and river" in the visual part are the repetition of "**mountains** and **rivers**" in the verbal texts. In Frame 26, a young man proposes to a young woman in the city. The "young couple" and the circumstance "city" in the visual part are the repetition of "city" and the hyponymy of "family" in the verbal text. Frame 27 shows the gate of an old building and Frame 28 a stele in a park, which correspond to the verbal texts in both frames. The "gate" is the hyponymy of "old buildings" in the verbal text and the "stele" is the repetition of "steles" in the verbal text. In Frame 29, "the city wall of Xi'an" in the image is the hyponymy of "ancient

buildings" in the verbal text. In Frame 30, the "cars running around the city wall" is the 'synonymy,' as it is synonymous with "hustle and bustle" in the verbal text.

### Conclusion and Recommendations

This study undertook a multimodal discourse analysis of the *Xi'an China* TPV based on Systemic Functional Linguistics, Visual Grammar, and Intersemiotic Complementary Framework. Through the ideational function analysis, representational meaning analysis, and intersemiotic complementarity analysis, the major findings are as follows:

Firstly, the verbal modality of the TPV conveys the information related to the history, culture, landscape, trade, industry and development of Xi'an. Regarding the ideational function of the video, the relational and material processes account for 81.8% of all processes. The relational process is the most frequently used in the verbal mode, which depicts the unique history and culture of Xi'an. The use of the relational process in verbal mode publicizes information about Xi'an to potential tourists objectively. The material process is used to introduce the development and natural beauty of Xi'an, which significantly impresses potential tourists. Mental process is the least used in this video, but it also tells potential visitors how they will feel about this charming city.

Secondly, the visual modality of the TPV represents the natural beauty, the entertainment, the events happening now, and the change of Xi'an from the past to present time. In terms of the representational meaning analysis, the narrative process is used more than the conceptual process. Action processes, reactional processes and circumstance are used in the narrative processes, which vividly shows the life in Xi'an from different aspects and captivate potential visitors. In conceptual process, classificational process, analytical process and symbolic process are involved in this video, which objectively displays different places in Xi'an. From the classificational process, which takes the largest proportion in the conceptual process of this video, tourists can have a feast of cultural relics, folk crafts, and antiques in Xi'an. Through the analytical process, tourists can enjoy the scenes of natural sites and urban landscapes of Xi'an. The symbolic process is used least, but it still plays an important role in showing the development of Xi'an.

Thirdly, based on Royce's Intersemiotic Complementary Framework, the frequency of semantic relations between verbal and visual modes is analyzed. In this video, the most frequently used relations are repetition and synonymy, which shows the complementary relationship between the two modalities. The cooperation of the two modes helps construct the overall meaning of the *Xi'an China* TPV, which publicizes the tourism and culture of Xi'an with expected effect.

### Recommendations

It is recommended that further research be carried out on many more TPV examples in China. This study focused on only one TPV, yet China has produced several TPVs. TPVs have a number of unique features that are important for meaning construction compared with other tourism promotion materials, thus it deserves more research in the future.

In terms of the promotion of Xi'an as a tourist destination in the world, diachronic research on TPVs about Xi'an in different stages is needed. It is hoped that this study will encourage students and researchers to dig deeper into multimodal discourse analysis as an approach to study tourism discourse with, especially dynamic tourism discourse.

Lastly, the study only analyzes the ideational function and representational meaning to examine the multimodal features of the TPV. To understand how verbal and visual modes combine to make meanings, it is recommended that interpersonal and textual functions as well as interactive and compositional meanings also be included in the data analysis part of future studies.

### References

- Almurashi, W. A. (2016). An introduction to Halliday's systemic functional linguistics. *Journal for the study of English Linguistics*, 4(1), 70-80
- Alyousef, H. S. (2016). A multimodal discourse analysis of textual cohesion in tertiary marketing texts written by international undergraduate students. *Semiotics*, 99-122
- Ammara, U. E., Anjum, R. Y., & Javed, M. (2019). A Corpus-Based Halliday's Transitivity Analysis of 'To the Lighthouse'. *Linguistics and Literature Review*
- Ansori, M., & Taopan, L. L. (2019). A multimodal discourse of promotional video Wonderful Indonesia. *English and Literature Journal*, 6(1), 1-18
- Aoki, S. and Fujimoto, T. (2021) 'Examination of Tourism Promotion Using Online Videos', in *Lecture Notes in*

- Networks and Systems. doi: 10.1007/978-3-030-65796-3\_30.
- Fong, Y. L. et al. (2017) 'the Impact of Tourism Advertisement Promotional Videos on Young Adults', *e-Bangi : Journal of Social Sciences and Humanities*, 12(3), pp. 87–102.
- Graça, A. R., Banha, F. and Banha, F. M. (2022) 'Video-Induced Tourism in Central Portugal: Production and Impact of Promotional Videos', *tourism & Management Studies*, 18(1), pp. 41–50. doi: 10.18089/tms.2022.180104.
- Jocuns, A. (2020) 'Aroi Mai? The Multimodality of Thainess in Thai Tourism Videos', *Multimodal Communication*. doi: 10.1515/mc-2019-0011.
- He, X. (2017). Transitivity of kinetic typography: theory and application to a case study of a public service advertisement. *Visual Communication*, 16(2), 165-194
- Hong, X., & Duan, C. (2020). The Construction of Sichuan Image under Multimodal Visual Grammar—Taking the Documentary “Aerial China (Sichuan)” as an Example. *Open Journal of Social Sciences*, 8(12), 108-120
- Hou, Z. Q. (2017). Comparison study of advertising videos of tourism city image-based on the content analysis method. *Eurasia Journal of Mathematics, Science and Technology Education*, 13(10), 6837-6846
- Hu, C., & Luo, M. (2016). A Multimodal Discourse Analysis of Tmall's Double Eleven Advertisement. *English Language Teaching*, 9(8), 156-169
- Lim, F. V. (2021). Investigating intersemiosis: a systemic functional multimodal discourse analysis of the relationship between language and gesture in classroom discourse. *Visual Communication*, 20(1), 34-58
- Liu, S. (2019). A multimodal discourse analysis of the interactive meaning in public service advertisement. *Journal of Advances in Linguistics*, 10, 1523-1534
- Losada, N., & Mota, G. (2019). 'Slow down, your movie is too fast': Slow tourism representations in the promotional videos of the Douro region (Northern Portugal). *Journal of destination marketing & management*, 11, 140-149
- Ly, T. H., & Jung, C. K. (2015). Multimodal discourse: A visual design analysis of two advertising images. *International Journal of Contents*, 11(2), 50-56
- Machin, D. (2016). The need for a social and affordance-driven multimodal critical discourse studies. *Discourse & Society*, 27(3), 322-334
- Matthiessen, C. M. (2018). Transitivity in Systemic Functional Linguistics: Achievements and challenges. *Estudos de transitividade em linguística sistêmico-funcional*. Santa Maria: PPGL Editores, 14-108.
- Papakitsou, V. (2020) 'Qualitative Research: Narrative approach in sciences', *Dialogues in Clinical Neuroscience & Mental Health*, 3(1), pp. 63–70. doi: 10.26386/obrela.v3i1.177.
- Pratiwy, D., & Wulan, S. (2018). Multimodal discourse analysis in Dettol TV advertisement. *KnE Social Sciences*, 207-217.
- Royce, T. D. (2015). Intersemiotic complementarity in legal cartoons: An ideational multimodal analysis. *International Journal for the Semiotics of Law-Revue internationale de Sémiotique juridique*, 28(4), 719-744
- Sembiring, B. (2020). Relational Process Clauses in the Students' Descriptive Essays. *Jurnal Bahasa Inggris Terapan*, 6(1), 25-34
- Shin, D. S., Cimasko, T., & Yi, Y. (2020). Development of metalanguage for multimodal composing: A case study of an L2 writer's design of multimedia texts. *Journal of Second Language Writing*, 47, 100714
- Sutrisna, G. (2020). Processes in Report Text: An Analysis of Systemic Functional Linguistic on Transitivity. *Online Submission*, 11(1), 10-16
- Wang, Y., & Feng, D. (2021). History, modernity, and city branding in China: A multimodal critical discourse analysis of Xi'an's promotional videos on social media. *Social Semiotics*, 1-24
- Wang, Y., & Han, W. (2022). The influence of destination promotion videos on residents' sense of empowerment and support for tourism. *SAGE Open*, 12(3), 21582440221106730
- Xin, Z. (2011) 'New Discourse on Language: Functional Perspectives on Multimodality, Identity, and Affiliation', *System*, 39(4), pp. 580–582. doi: 10.1016/j.system.2011.10.007.
- Yao, Y. and Zhuo, Y. (2018) 'A Multimodal Discourse Analysis of the Promotional Video of Hangzhou', *English Language Teaching*. doi: 10.5539/elt.v11n10p121.
- Zein, T. T., Sinar, T. S., Nurlela, N., & Yusuf, M. (2019, February). Process Types of Transitivity System in English Department Students' Narrative Texts. In *Sixth of International Conference on English Language and Teaching (ICOELT 2018)* (pp. 149-155). Atlantis Press