

# Multimodal Interactivity and Brand Image of Xiaomi Company: Take its Overseas Product Launches as an Example

GUO Fang<sup>[a]</sup>; HE Changlong<sup>[b]</sup>; SONG Xiaoli<sup>[c]</sup>\*

<sup>[a]</sup> Professor, School of Foreign Languages, North China Electric Power University, Beijing, China.

<sup>[b]</sup> School of Foreign Languages, North China Electric Power University, Beijing, China.

<sup>[c]</sup> Associate Professor, School of Foreign Languages, North China Electric Power University, Beijing, China.

\*Corresponding author.

Supported by the North China Electric Power University Teaching Reform Project (No. XM2212362).

Received 12 July 2022; accepted 7 September 2022

Published online 26 October 2022

## Abstract

In recent years, the mobile phone has become an indispensable part of people's life, which is an important tool for people to socialize, shop, relax and pay. With the saturation of domestic mobile phone market, Chinese manufacturers are gradually casting their eyesight abroad and seeking to enter the overseas market. Xiaomi company is one of the best. Nowadays, almost all mobile phone manufacturers will launch new products through the product launching ceremony, which is also a good way of marketing and brand image transmission. The product launch usually contains abundant multimodal resources, such as image, text, music, video and so on. Therefore, the multimodal resources of Xiaomi product launch are worth exploring, which can provide experience and reference for other companies preparing to penetrate the international market. Drawing on the analytical framework of the interactive meaning of visual grammar, this thesis selects five videos of Xiaomi mobile phone product launch from 2016 to 2020 as the main research objects to elucidate how Xiaomi allocate the multimodal resources in its product launch and how multimodality facilitates the construction and transmission of Xiaomi's brand image.

**Key words:** Multimodal interactivity; Visual grammar; Xiaomi phone; Brand image; Product launch

Guo, F., He, C. L., & Song, X. L. (2022). Multimodal Interactivity and Brand Image of Xiaomi Company: Take its Overseas Product Launches as an Example. *Canadian Social Science*, 18(5), 61-68. Available from: <http://www.cscanada.net/index.php/css/article/view/12828>  
DOI: <http://dx.doi.org/10.3968/12828>

## 1. INTRODUCTION

### 1.1 Research Background

With the integration of the world economy and the rapid development of Internet technology, going abroad to seek opportunities and explore new markets has become a new normal state for companies to continue to succeed and flourish. What's more, more international business communication and exploration are ascribed to the "One Belt and One Road" Initiative. The Smartphone, now an integral part of our modern life, is an important device to socialize, shop, entertain, learn and so on. With its powerful features and relatively cheap price, Xiaomi mobile phones have become one of the leaders in the domestic mobile phone industry. In recent years, China's smartphone market, with the slowdown of consumer growth, is shifting to a relatively saturated market.<sup>1</sup> As a result, many smartphone companies started to enter the global market, and Xiaomi company has achieved great success. According to Counterpoint, Xiaomi captured the biggest market share in India<sup>2</sup> and the fourth in Europe<sup>3</sup> in 2019. Whether mobile phones sell well in overseas markets depends not only on the quality but also on the

<sup>1</sup> Counterpoint. Infographic: CY-2019 | Mobile Market Monitor. <https://www.counterpointresearch.com/infographic-cy-2019-mobile-market-monitor/>.

<sup>2</sup> Counterpoint. India Smartphone Quarterly Market Data (2019Q4 – 2021Q1). <https://www.counterpointresearch.com/india-smartphone-share/>.

<sup>3</sup> Counterpoint. Growth Opportunities in the Saturated Global Smartphone Market. <https://www.counterpointresearch.com/growth-opportunities-saturated-global-smartphone-market/>.

marketing strategies. The huge success of Xiaomi can be a good example in case for other Chinese companies.

When a company enters a new market, it will use a variety of strategies to attract prospective customers. This progress often comes with many challenges due to cultural differences and local competitors. One important reason for Xiaomi's success is its product launch, an effective and efficient business strategy for advertising and brand-construction. To gain consumer attention and to promote new products, many companies hold product launches. For example, Apple holds new products launch twice a year with about ten million viewers just on YouTube<sup>4</sup>. Xiaomi put overseas launches live on many different platforms, such as YouTube, Twitch, and its homepage, and the number of audiences on YouTube often exceeds two hundred thousand<sup>5</sup>. These launches usually contain multimodal resources like verbal messages, images, sounds and videos. Every resource, especially in a commercial launch is designed with intentional meaning. Therefore, through Xiaomi's global product launches, we are able to see how Xiaomi constructs its brand image and attracts customers through multimodal discourse analysis. In Xiaomi's product launches, the multimodal resources will be discussed by Kress and van Leeuwen's (1996,2006) visual grammar.

## 1.2 Research Questions

The research questions of this thesis are represented as follows: (1) How is the interactive meaning constructed in these launches? (2) What are the similarities and differences about the distribution of the four components of interactive meaning in these five launches? (3) How does Xiaomi build and transmit brand image in its overseas markets?

This thesis will analyze five Xiaomi's overseas product launches based on the theoretical framework of visual grammar and aims to figure out how Xiaomi company builds and transmits its brand image in its overseas markets. Then this thesis will offer some feasible suggestions for other domestic companies who want to go global.

---

## 2. LITERATURE REVIEW

---

### 2.1 Previous Studies on Multimodal Discourse Analysis

As the pioneer of multimodal discourse analysis, Barthes (1977) discusses how language and images interact with each other to convey meanings. Then Halliday (1978) puts forward the Systematic Functional Grammar theory arguing language has three meta-functions: ideational function, interpersonal function, and textual function. Based on Halliday's SFL, many linguists find out that

other modes can also be "semiotic". They start to extend the system-functional theories to various modes. Knox (2007) analyzes the multimodality of newspaper discourse and discusses the language mode and the image mode. Hu (2007) makes contributions to the construction of multimodal discourse both theoretically and practically, such as multimodal construction of social semiotics and meaning.

Bateman & Schmidt (2012) find out that the combination of different modes makes up the art and aesthetics of movies by a multimodal approach. Djonov & Leeuwen (2013) conduct research on how the layout affects the effectiveness of PowerPoint. Bateman (2014) explains the text-image relations in his database based on different multimodal theories. Nørgaard (2019) points out the multimodal nature of novels and constructs a new framework to analyze multimodal resources in novels.

### 2.2 Previous Studies on Visual Grammar

Based on Halliday's SFL, Kress and van Leeuwen (1996, 2006) believe that these meta-functions not only apply to language but also to all modes of human communication including visual modes. They develop the theoretical framework of visual grammar and expand three meanings (representative meaning, interactive meaning, compositional meaning) corresponding to Halliday's three meta-functions.

Many scholars have conducted the research on multimodal analysis of videos with visual grammar structure. Bateman (2012) applies multimodal discourse analysis on movies. The research shows the importance of different modes of the meaning transfer in movies. Some other researchers mainly focused on analyzing multimodality in Boeing's apologizing video (Liu, 2020), interpersonal meaning of TV advertisement (Li, 2016), interactive meaning in promotional video (Ansori, 2019). For example, Ansori(*ibid*) put forward that the interactivity of promotional video is the key factor for arousing customer's purchasing desire. Customer is more likely to make a purchase when they can interact with the promotional video. However, there are few linguists trying to research product launch, which is full of many modes, and few of them have conducted diachronic study. Therefore, this thesis attempts to use the "fresh new" commercial field for research to further enrich the application scope of MDA, visual grammar in particular.

---

## 3. THEORETICAL FRAMEWORK

---

### 3.1 Contact

Kress and van Leeuwen (1996) deem that contact refers to the image viewer establishes an imaginary interpersonal relationship with people or things in the image through the line of sight, making the viewer to have a certain emotional experience. It can be divided into demanding

---

<sup>4</sup> Apple. Discover More from our Apple Events. <https://www.youtube.com/user/Apple>.

<sup>5</sup> Xiaomi. Mi Launch. <https://www.youtube.com/user/XiaomiChina>.

and offering type. The difference between the two is whether the image participants and the viewer have eye contact. If they have eye contact, it is demanding, if not, it is offering. Demanding images often seek for trust or recognition from the viewers, arouse resonance and produce positive interaction with the viewers; Offering images are more objective, without direct eye contact, providing some objective facts or rational suggestions from the perspective of a third party, which is often neutral.

### 3.2 Distance

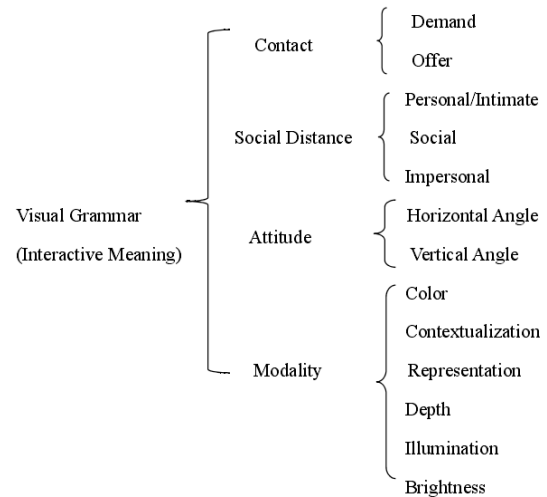
Kress and van Leeuwen (1996) refer to that social distance as related to the distance between the participant and the viewer. The distant view means a long distance, and the close view means a short distance. The visual grammar divides the distance into three categories according to the length of view taking: the personal distance, the social distance, and the impersonal distance. The personal distance shows only participant's head and shoulder, which is often realized by a close-up shot. The social distance shows participant's half body, which is often realized by a medium shot. The impersonal distance includes the participant's whole body and the surrounding environment, which is often realized by a long shot.

### 3.3 Attitude

Kress and van Leeuwen (1996) depict that attitude can be understood as the angle selected in photography or painting, which expresses the subjective or objective attitude towards the image viewers, mainly including two categories: horizontal angle and vertical angle. The former consists of frontal angle and oblique angle. Frontal angle can make viewers feel that they are part of the image world while the oblique angle can present viewers a world that is far away from them, thus having a sense of alienation. As for vertical angle, it consists of high angle, low angle and eye to eye angle. The high angle makes the viewers powerful over the participants in images. On the contrary, the participants look more authoritative, and viewers feel a sense of being controlled when the image is taken from low angle. In eye-to-eye angle, the viewer and participant have an equal relationship.

### 3.4 Modality

Kress and van Leeuwen (1996) propose that modality in visual grammar is related to the degree of viewers' attention about the images. Modality is the social semiotic approach to the question of truth, which reflects the authenticity or credibility of the images. The higher the authenticity and restoration of an image, the higher the modality of the image. Modality can be divided into three kinds: high modality, medium modality and low modality. Color is an important fact that influences modality, such as color saturation, and is the main approach to study modality in this research. The structure of the interactive meaning is represented in Figure 1.



**Figure 1**  
**Interactive Meaning in Visual Grammar (Kress & van Leeuwen, 1996)**

## 4. DATA COLLECTION AND RESULTS

### 4.1 Data Collection

In this thesis, the analyzing objects are five videos of Xiaomi's overseas product launch downloaded from its official account on Youtube: <https://www.youtube.com/c/xiaomi/featured>. These launches were held for five successive years from 2016 to 2020. All the pictures cited in Chapter Four are screen shots cut from these videos. There are 108, 92, 93, 97 and 106 pictures from 2016 to 2020, respectively as shown in Table 1.

**Table 1**  
**Basic Information of Five Product Launches**

Year	Length	Products	Numbers of screen shots
2016	00:57:58	Mi Max	108
2017	01:39:53	Redmi 4	92
2018	01:35:49	Mi 8	93
2019	01:56:31	Mi 9	97
2020	01:58:05	Mi 10	106

### 4.2 Results

According to the theoretical framework in Section 3, all the pictures from these five overseas product launches are analyzed. The data of their distribution will be represented and discussed.

#### 4.2.1 Distribution of Contact

The distribution of contact can be summarized in Table 2.

**Table 2**  
**Distribution of Contact**

Year	Demanding image	Offering image
2016	20	88
2017	18	74
2018	17	76
2019	25	72
2020	33	73



**Figure 2**  
**Distribution of Contact**

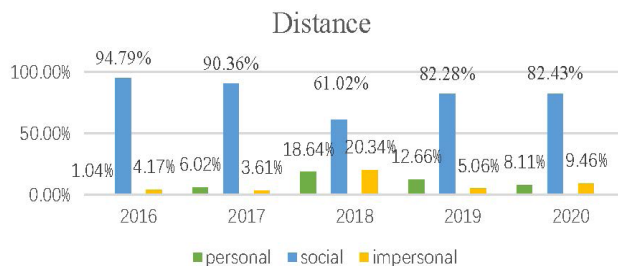
The function of offering image is to provide information. The function of demanding image is to ask information from the viewer to reduce distance mentally. From Figure 2, it can be known that offering image takes the majority since the main purpose of product launch is to introduce products. However, Xiaomi has gradually increased the proportion of offering, from 18.52% to 31.13%, aiming to reduce the sense of distance and allow the audience to participate more in the interaction with the pictures.

**4.2.2 Distribution of Distance**

The distribution of contact can be seen in Table 3. Some pictures are just words and numbers, and are not included in statistics.

**Table 3**  
**Distribution of Distance**

Year	Personal distance	Social distance	Impersonal distance
2016	1	91	4
2017	5	75	3
2018	11	36	12
2019	10	65	4
2020	6	61	7



**Figure 3**  
**Distribution of Distance**

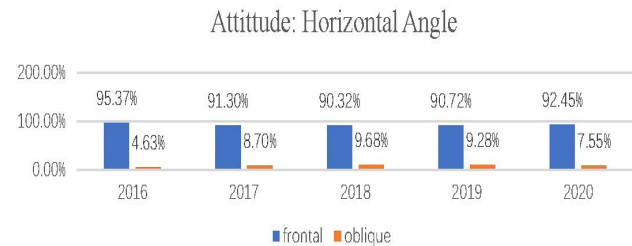
From Figure 3, all these product launches except 2018's uses social distance at most, exceeding 80%. The use of social distance can make the audience feel close with the participants in pictures and give the audience a more comfortable watching experience. The percentage of personal and impersonal distance is relatively low, almost less than 10%. The main purpose of using personal distance is to attract the audience and put themselves into the world of pictures. The use of impersonal distance allows the audience to see the data clearly and receive complete information.

**4.2.3 Distribution of Attitude**

The distribution of attitude can be seen in Table 4.

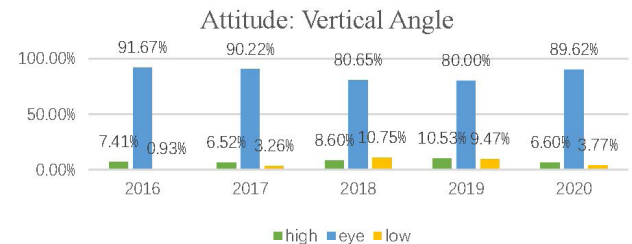
**Table 4**  
**Distribution of Attitude**

Year	Horizontal Angle		Vertical Angle		
	Frontal Angle	Oblique Angle	High Angle	Eye-level Angle	Low Angle
2016	103	5	8	99	1
2017	84	8	6	83	3
2018	84	9	8	75	10
2019	88	9	10	76	9
2020	98	8	7	95	4



**Figure 4**  
**Distribution of Horizontal Angle**

Figure 4 shows that more than 90% of the pictures are directly oriented to the audience, which makes it easier for the audience to obtain information from the pictures. Oblique angle is sometimes used to make the audience less involved with the image and provides an observer's perspective.



**Figure 5**  
**Distribution of Vertical Angle**

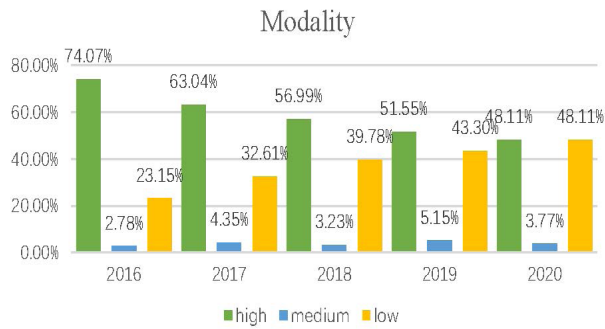
From Figure 5, the results show that more than 80% of the images are on the same level as the viewer's line of sight, which can make the viewer gain information at a comfortable position. The high angle and low angle are only used a few times, less than 20%, indicating that the audience does not need to move their heads or bodies. Xiaomi tries to make the viewers to get information with ease and comfort.

**4.2.4 Distribution of Modality**

The distribution of modality is listed in Table 5.

**Table 5**  
**Distribution of Modality**

Year	High-saturation Color	Medium Color	Low-saturation Color
2016	80	3	25
2017	58	4	30
2018	53	3	37
2019	50	5	42
2020	51	4	51



**Figure 6**  
**Distribution of Modality**

Figure 6 shows that the majority of the pictures are bright and colorful. Such images can reflect a high degree of authenticity and deeply impress the audience, which make them easier to enter the world of pictures. The percentage of high-saturation colors gradually decreases while low-saturation increases. The reason is that Xiaomi's has developed a design style of simplicity, as shown in Figure 6. The product launches a gradually turn into a simplified style, using a lot of low-saturation colors like black and white to showcase a more and more mature design style of the products and the taste Xiaomi company guiding the public to pursue.

## 5 ANALYSIS

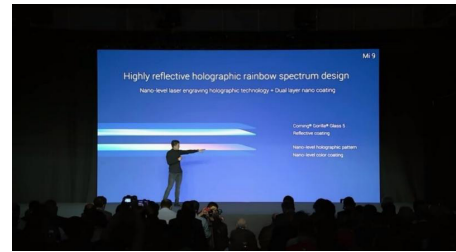
### 5.1 Contact

Figure 7 is about Xiaomi fans, from all over the world, gathering on the Great Wall. Nearly all participants are looking directly at the audience, and their eyes and the audience eyes form eye contact. Therefore, this is a demanding image. The participants all wear clothes of Xiaomi, raise hands high and seem to offer invitations to the audience to join them. The audience will feel a sense of "Mi community" and want to join them automatically. Such an image can show Xiaomi's belief—just for fans.



**Figure 7**  
**Xiaomi Fans' Party**

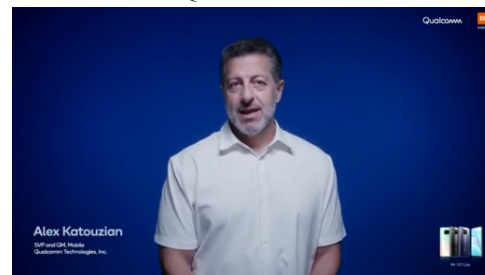
Figure 8 is chosen from the introduction of the phone's rear cover. The speaker's eyes are focusing on the screen with hands pointing to introducing words in the PPT. There is no eye contact between the speaker and the audience, and so Figure 8 belongs to offering pictures. The audience will follow the speaker's sight and focus on information on the screen. The purpose of introduction products can be achieved.



**Figure 8**  
**Introduction of Phone Design**

### 5.2 Social Distance

The man in Figure 9 is Alex Katouzian who is the SVP and GM of Mobile Qualcomm Technologies company. He is introducing his great collaboration with Xiaomi. Since his upper body can be seen clearly by the audience, image 9 belongs to personal distance. The audience can feel like they are talking with some friends. Therefore, the speaker and the audience can establish a close relationship. The audience will feel the speaker, who comes from the leading company of wireless technology, is talking personally to them about how excellent Xiaomi is. As a result, it is easy for the audience to think Xiaomi is at the same level like Qualcomm.



**Figure 9**  
**Interview**

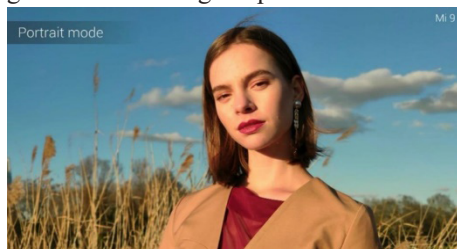


**Figure 10**  
**Selfie with Xiaomi Phone**

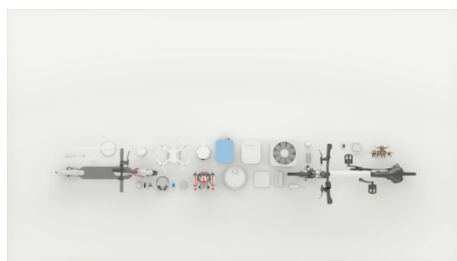
In Figure 10, the three main participants are taking selfie. Only their heads and shoulders can be observed easily, and so this figure belongs to personal distance. The viewer can see their happy smiling faces clearly and be impressed more deeply. Because of close shot, the distance between the participants and the viewer is also close. The viewer will feel like he or she are taking selfie with friends. The beautiful light and shade in Figure 10 can effectively demonstrate the great photography performance of Xiaomi phone.

### 5.3 Attitude

Figure 11 is selected from the introduction of portrait mode, one photography feature. The participant is in the middle of Figure 11 and her eye sight is basically at the same level with the viewer's eye sight. As a result, this figure can be seen as frontal and eye-level angle. The audience will feel comfortable while looking at the picture because they do not have to move head or body to see Figure 11. Therefore, they can focus on the image lastly and carefully. The terrific portrait mode function will be represented directly and effectively, thus completing the goal of introducing the product.



**Figure 11**  
**Portrait Mode**



**Figure 12**  
**Array of Products**

Figure 12 gives the audience a look at Xiaomi's products and is photographed from a high angle. Normally, the viewer will look down at the objects in the picture from a downward angle. This kind of angle gives power to the viewer and makes the things in the picture be within reach and at the command of the viewer. (Kress and van Leeuwen, 1996) Therefore, the audience will have a deep impression of Xiaomi's array of products and is more likely to make a purchase.

### 5.4 Modality

Figure 13 mainly consists of words and numbers. The major colors are low saturation colors like black and white, with little high saturation color: orange. The words and numbers are all white and the background is black. Such a way is convenient and comfortable for the audience to gain information. The orange color, which is also the main color of Xiaomi's logo, is quite conspicuous in the low saturation background. The audience will have a deep impression that the overseas market is also important to Xiaomi and Xiaomi is an international company.



**Figure 13**  
**Proportion of Revenue**



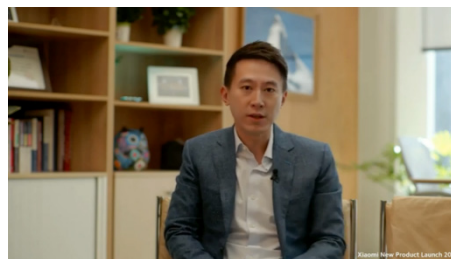
**Figure 14**  
**Dxomark's Comment**

In Figure 14 there are mainly black background and white words, which are low-saturation color. The viewer can see the information clearly by the use of such colors. The contrast of black and white can make the information quite obvious, and so the viewer can pay attention to the information immediately.

## 5.5 The Building-up of Xiaomi's Brand Image

### 5.5.1 Demonstration of core values

Figure 15 is chosen from the interview of Shou Zi Chew, who is the president of International Xiaomi. It can be seen the half body of the speaker, and so it is social distance.



**Figure 15**  
**Interview of Shou Zi Chew**

In addition, the participant and the audience have eye contact, which make it a demanding image. Therefore, the audience will feel they are brought into the image and become the interviewer, taking face to face with the participant. The audience will feel comfortable and have a deep impression on Xiaomi's core value: making quality technology available for everyone.

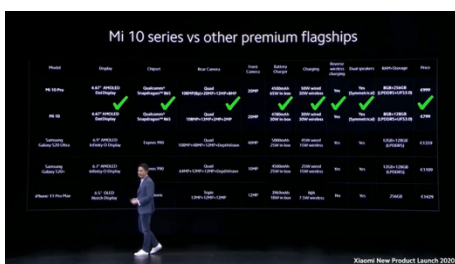
Xiaomi mentions its core value almost in every overseas product launch, and some other examples are given in Figure 16. This can improve brand image a lot.



**Figure 16**  
**Xiaomi's Core Value**

### 5.5.2 Brand Positioning

Image 17 is chosen from the comparison between the Xiaomi 10 series and other flagship competitors. The language clearly shows that Xiaomi 10 series have many better features than other competitors, such as camera, processor, price, etc. As for modality, the part of Xiaomi has bright white and green color, which are high-saturation color whereas the rest competitors only have low-saturation color, dusky white. As a result, the audience will put more attention to Xiaomi's part and know that the Mi10 series are the best phones now. At the same time, the competitors are famous and established phone companies like Apple and Samsung, so that the audience will have the impression that Xiaomi is at the same level with or even better than the leading companies, bringing Xiaomi a very good brand image. By comparing with other leading companies in phone industry, Xiaomi shows its strong scientific and technological strength and relatively affordable prices.



**Figure 17**  
**Mi 10 Series VS Other Flagships**

All in all, by interacting with the audience, Xiaomi's product launches effectively introduce the product information and constructs a good brand image: close to the people, simplified design, high-tech and responsible company.

## 6. CONCLUSION

### 6.1 Major Findings of the Research

In this thesis, different modes and their interactions of Xiaomi product launches are analyzed on the basis of visual grammar focusing on the interactive meaning. Here are some main findings after data analysis:

Firstly, the four components of the interactive meaning are employed well by Xiaomi company. As for contact, the demanding image usually aims to attract the attention of the audience and makes them actively participate in the launch. In addition, the offering image aims to provide information of the products. As for social distance, close shot image and medium shot image are used frequently and alternately to form a close or comfortable social distance with the audience, making them feel at ease and intimate with Xiaomi product and Xiaomi company. As for attitude, eye-level angle can make the audience feel equal with the speaker and high-level angle is used to offer a sense of ownership to the audience. As for modality, high-saturation colors are often used to highlight key information and impress the audience.

Secondly, Xiaomi uses many modes like image mode and language mode together to build brand image by demonstrating core values and showing brand positioning. These modes can strengthen each other's effect to build brand image and make the audience willing to buy the products.

Thirdly, Xiaomi has built an excellent brand image: taking customers seriously, providing good product quality and approachable price, making innovation for everyone, owning strong technical strength, etc. The brand values have also been well-transmitted through the manipulation of multimodal visual resources to their overseas prospective customers who are caring more and more about product experience, which is a very crucial element for its success in the global markets.

### 6.2 Limitations of the Research

Although some efforts are made to analyze these five launches, there are still limitations in this thesis. First of all, this thesis only analyzes the language mode, image mode and color mode. In Xiaomi product launches, there are many other modes, such as music, gestures of the speakers and so on. All modes can be a good resource for research. Then, limited by the length of the article, the examples given in the thesis just occupy a small part of the launch. More examples need to be added to support the findings. Third, this thesis only conducts qualitative research mainly focusing on content analysis, which lacks quantitative research based on multimodal analytical software.

### 6.3 Suggestions of Further Research

In the future, the multimodal discourse analysis can be conducted by specific software. The complexity

of multimodal analysis involves annotation, analysis, and interpretation of semiotic models from multimodal resources. If the field is to move beyond general observations to gain empirical insights from detailed semiotic analysis, we need a series of theoretical and practical tools (O'Halloran, 2012). In addition, some relevant surveys or interviews can be made to study the effectiveness of Xiaomi product launches to form a triangular test for the research to further improve the persuasiveness of the analytical results.

---

## REFERENCES

---

- Ansori, M. (2019). A multimodal discourse of promotional video wonderful Indonesia. *English and Literature Journal*.
- Barthes, R. (1977). *Rhetoric of the image*. London: Fontana.
- Bateman, J. A. & K. H. Schmidt. (2012). *Multimodal film analysis*. London: Routledge.
- Bateman, J. A. (2014). *Text and image: A critical introduction to the visual-verbal divide*. London and New York: Routledge.
- Djonov, E., & Van Leeuwen, T. (2013). Between the grid and composition: Layout in PowerPoint's design and use. *Semiotica*.
- Halliday, M. A. K. (1978). *Language as social semiotic: The social interpretation of language and meaning*. London: Arnold.
- Hu, Z. L. (2007). Multimodalization in social semiotics. *Language Teaching and Linguistic Studies*, (1), 1-10.
- Knox, J. (2007). Visual-verbal communication on online newspaper home pages. *Visual Communication*, 6, 19-53.
- Kress, G. (2006). *Literacy in the New Media Age*. London: Routledge.
- Kress, G., & Van Leeuwen, T. (1996). *Reading images: The grammar of visual design*. London: Rutledge.
- Li, D. Y. (2016). Multimodal discourse analysis of the interpersonal meaning of TV advertisements. *International Journal of Social Science and Humanity*, 6(12).
- Liu, S. L. (2020). Multimodal critical discourse analysis of Boeing's original CEO apology video (in Chinese). *Overseas English*, (22), 15-18.
- Nørgaard, N. (2019). *Multimodal stylistics of the novel: More than words*. London and New York. Routledge.
- O'Halloran. (2012). Interactive software for multimodal analysis. *Visual Communication*.