

## Discursive Strategies of a Political Cartoon Video

Xiqin Liu<sup>a,\*</sup>, Guangxue Dai, Shuozi Wu<sup>b,\*</sup>, Wenwen Zheng, Ying Qu

School of Foreign Languages, South China University of Technology, Guangzhou, China

<sup>a</sup>flxqliu@scut.edu.cn, <sup>b</sup>2214355755@qq.com

\*Corresponding authors

**Keywords:** Multimodal Critical Discourse Analysis, Political Cartoon Video, Positive Self, Recontextualization

**Abstract:** This article makes a multimodal critical discourse analysis of a cartoon video *How Leaders are Made* as a new Chinese political discourse in social media. The video clips about Chinese and Western presidential election systems were compared to identify variations in discursive strategy use. The effectiveness of the cartoon can be explained by identifying it as a blended mental space incorporating elements of the textual, musical and visual worlds. Two strategies for positive self-representation are identified: imbalanced comparison, and rhetoric-based recontextualization. The video discourse fulfils implicit and rational self-representation as an upgrade of the communist discursive tradition.

### 1. Introduction

Animated political cartoons deserve more scholarly attention because they can, with rich multimodal devices, provide insights into key socio-political issues, including identities, ideologies and power relations. They are often used for ridicule, satire, or negative representation<sup>[1:7]</sup>. However, a recent cartoon video produced in China, *How Leaders are Made*, fulfils the role of representing the positive self of Chinese government or leaders.

This hot video was uploaded to China's largest online video site, Youku, in 2013, and gathered over 1,000,000 views in the first two days. It was also running on YouTube ([https://www.youtube.com/watch?v=M734o\\_17H\\_A](https://www.youtube.com/watch?v=M734o_17H_A)), attracting the attention of western outlets. It shows that the secret of Chinese Miracle (incredible economic success) lies in the careful way of training leaders. As a hallmark video featuring cartoon characters representing Chinese national leaders for the first time, it deserves in-depth critical analysis.

A general trend in cartoons is towards iconicity and indexicality, in that one or more resources systematically “stand in” for others<sup>[2:35]</sup>. The visual features indicating movement produce different attitudinal stances in cartoons. A study of modality thus needs to take into consideration how semiotic resources other than language, such as movement, gaze and depiction, contribute to the expression of attitudinal and evaluative meanings by increasing the range of possibilities. The use of caricature in cartoons is a good illustration of the ways in which visual and other semiotic forms can be shaped and deformed in particular ways to achieve interactional ends.

Previous studies on discursive representation of positive self and negative other mainly focus on linguistic analysis. Scholars explored language ideologies and attitudes embodied in the lyrics of songs. But in Critical Discourse Analysis (CDA), few researchers, especially those comparing self and other identities, pay adequate attention to the connections and interactions among the linguistic, visual and sound resources in a multimodal text, such as an animated cartoon video.

The multimodal turn in sociolinguistics suggests the need to consider how language interact with musical and visual features<sup>[3]</sup>. Considerable research has been devoted to music videos from a cultural studies and media studies perspective contributing to an understanding of the multilayered meanings of music videos from the viewpoint of gender representation, race and ethnicity (ibid.). But cultural studies tend to lack a systematic focus on meaning making that can account for the multiplicity of ways that language interacts with visual and musical modes (ibid.). A combination of

multimodal and critical discourse analytical approaches arguably provides a most comprehensive framework or blueprint for conducting a systematic analysis of music videos that considers how different semiotic resources interact and connect to macrolevel discourses<sup>[4:6]</sup>. Furthermore, most literature is devoted to the videos in Western countries, rather than those produced in China.

With increasing attention on discursive power, the “proper telling of China’s stories” has recently become a key slogan in China, and social media outlets emerge for articulating China’s communist discourse to domestic and international audiences. China vs West political settings and discourses are often considered contrastive extremes<sup>[5]</sup>. Comparative discursive studies on the linguistic level<sup>[6]</sup> provide insightful findings, mostly revealing the low transitivity, high modality<sup>[5]</sup> and preoccupation with political ideology representation<sup>[7]</sup> in communist discourse. But little research has investigated their interactions multimodally to see if there is any movement between eastern and western discourses. This study examines whether and how the semiotic resources in the video combine to represent a positive national image.

## 2. Theoretical Framework and Analytical Method

This study was conducted under the theoretical framework of Multimodal Critical Discourse Analysis<sup>[8]</sup> as well as Machin’s<sup>[4]</sup> paradigm of music analysis. The two are grounded on the multimodal analysis of social semioticians<sup>[9]</sup> drawing on the systemic functional theory of Halliday<sup>[10]</sup>. Multimodal Critical Discourse Analysis aims to identify and reveal communicative choices through a careful process of describing the language and the visual. Texts or discourses use linguistic and visual strategies that appear normal or neutral on the surface, but which may actually be ideological and seek to shape the representation of events and persons for particular ends<sup>[8:9]</sup>. Machin<sup>[4]</sup> views music as discourse, and also adopts a socio-semiotic perspective to analyze semiotic choices in sound and music, and the meanings of arrangement, sound qualities and rhythm.

A method of multimodal comparison was employed. Analyzing a political discourse characterized by comparison as a strategy for persuasion or argumentation, the article makes socio-semiotic comparison in each of the three modes (verbal, visual and auditory) for meaning-making to identify similarities and variations.

The verbal, visual and auditory modes of this video were examined systematically. Machin and Mayr<sup>[8]</sup> provide analytical tools for exploring the verbal and visual, and Machin<sup>[4]</sup> for investigating the auditory.

## 3. Multimodal Critical Discourse Analysis

### 3.1. Verbal Analysis

The video falls into four parts and the caption consists of 41 sentences (see Appendix). Part 1 (Sentences 1-3) of the video tells ways to realize the dream of becoming a president vary from country to country. Part 2 (Sentences 4-14), with 187 words, briefly illustrates the US and the UK electoral systems. Part 3 (Sentences 15-38), with 463 words, depicts the Chinese electoral system in greater detail. Part 4 (Sentences 39-41) summarizes the theme, all roads lead to national leadership, echoing to Part 1. Comparison between Parts 2 and 3 is made in rhetoric, reference, transitivity, concealment, and modality, following Machin and Mayr<sup>[8]</sup>.

Regarding referential strategies, the second and third persons are abundantly used in Part 2 and Part 3 respectively. That may indicate a change in social distance from near to far. Quantifiers denoting time length, country size and money are used for comparison. In brief, statistics gives the impression of objective research and scientific credibility<sup>[11]</sup>. As for transitivity, Part 3 displays lower agency than Part 2. Nine sentences in Part 3 (Sentences 22-24, 26-27, 36-39) illustrate presidential candidates or officials as passive recipients rather than active agents. Concerning concealment, more nominalizations and abstractions are used in Part 3 for concealing the agent. Last, Part 2 and Part 3 both feature high modality, but difference exists. Part 2 mostly uses words expressing polarity or a high degree (e.g. anyone, wherever, no one, definitely; super-complex,

excellent) to convince people but the exaggeration may conceal truth. Part 3 mostly uses modals (e.g. have to, need to, could, must, would), asserting the narrator's certainty and confidence. The striking use of modals corresponds with Chen's<sup>[6]</sup>findings.

The video employs implicit and explicit comparisons linguistically to represent positive self against negative other (Table 1). Sharp differences lie between Parts 2 and 3 in rhetoric, reference, transitivity, and concealment. As a result, Part 3 displays richer ideational metafunction with greater details for authority, but Part 2 exhibits stronger interpersonal metafunction for vitalization.

Table 1 Verbal qualities and strategies for comparison.

	China (Part 3)	US & UK (Part 2)
1.Rhetoric	metaphor, metonymy, parody	metaphor, metonymy
2.Reference	3rd person, honorifics; more quantifiers/toponyms	2nd person; a quantifier for money; few temporal/spatial references
3.Transitivity	lower	higher
4.Concealment	more nominalizations/abstractions	less nominalizations
5.Modality	high, modal	high, polarity
6.Metafunction	ideational	interpersonal

### 3.2. Visual Analysis

Visual comparison reveals sharp differences between Parts 2 and 3 in these aspects: 1) rhetoric; 2) setting, dress, posture and body movement; 3) information value, salience and color (see Table 2).

Since the video theme is presidential election, the clips of final winning by Barack Obama, David Cameron and Xi Jinping can be most representative of the visual discourse. To explore discourse dynamism and achieve analytical comprehensiveness, screenshots of the last three scenes of each winning are taken as samples for comparison.

Table 2 Caricature qualities and discursive strategies for comparison.

	China (Xi)	US & UK (Obama & Cameron)
1. Setting	dark, little variation	bright, much variation
2. Dress	more formal	casual, individual
3. Posture & body movement	standing with legs closed and body side to viewer, bouncing	standing with limbs spread and body front to viewer, fist moving outward, pounding/somersault
4. Information value	up and bottom	left and right (US); center and margin (UK)
5. Salience	lower degree, invariability	higher degree, variability
6. Color	limited palette; low saturation; luminosity; low modulation	high differentiation; high saturation
1. Rhetoric	metaphor, metonymy	metaphor, metonymy
2. Transitivity	lower	higher
3. Modality	lower	higher
Metafunction	textual	interpersonal

On the whole, high transitivity and modality can be found in the video clips about US and UK which seem to be more emotional and thus display stronger interpersonal metafunction than those about China. This visual contrast reinforces the differences between the two electoral systems, one valuing the inner qualities of the candidates, and the other highlighting the “outer” abilities like lobbying.

There is abundant and systematic use of visual metaphors and metonymies. The three presidents, Obama, Cameron and Xi, are compared to a boxer, an acrobat and a kungfu master respectively. The mouth of Obama (min.00:42) is visually dramatized together with the use of a trumpet or horn

(min.00:29, 05:01) to show the importance of lobbying. This metonymy is related to the informal idiom, “blow your own trumpet/horn” with negative connotations.

Textual metafunction is realized differently. The visual discourse about China displays stronger textual metafunction. The step-climbing scenario appears four times (starting from min.02:20, 02:52, 03:16 and 04:26). The heavy smoke (min. 04:13-04:25) symbolizes serious pollution, in dark colors harmonizing with those of Xi’s election clips, and creating visual cohesion. The smoke is soon dispersed with explosion effects, connoting an optimistic future. This may imply the government’s open-mindedness and determination although the main reason for pollution is not disclosed.

### 3.3. Auditory Analysis

The English narrator speaks with an American accent as if the story were from the mouth (and perspective) of a US citizen as a witness and bystander, corresponding with the use of the third person “he” to denote President Xi in Part 3. A male voice in the middle register is heard, with associations of brightness and energy. Among the four pieces of background music, sharp differences exist as they evoke varying moods and feelings.

The first melody is high-pitched pop music, lively and slightly comical. With connotations of joyfulness, playfulness and brightness, it has medium transitivity and modality. The second melody is an intensified version of the first melody, with higher transitivity and modality. It has a faster and even pulsating rhythm and displays comical qualities more clearly. Fast tempos and bouncy dotted rhythms are used to suggest something playful or silly<sup>[4:167]</sup>, and convey excitement and action. Perhaps the tension of the presidential campaign is implicitly depicted with slight irony because there are extremes of tonal range, often indicating comedy characters. It is arguable that a clownery metaphor is used in the first two melodies.

The third melody (min. 01:30-02:42) is a Chinese march for celebration, a happy and solemn occasion with high transitivity and modality. It marks the most important transition in the whole video since three modes are all segmented here (min. 01: 30, Figure 1). As the only melody creating hifi soundscape, it implies the coming of a large group of people in good order, and a march metaphor is used. This auditory recontextualization of Chinese election connotes collectivism, strength, order, and consistency instead of chaos.

The last melody is relaxing pop music of slow rhythm, featuring low transitivity and modality. Arguably, it is a peace metaphor. Softer instruments are used with a sense of femininity<sup>[4:166]</sup>, related to gentleness, inwardness and close social distance. Sweeping sounds represent peaceful wide-ranging landscapes<sup>[4:99]</sup> of China. Low pitch suggests closeness and relaxation. This calm music displays evenness, free of raspiness and tension, with metaphorical association with political stability. It may imply China’s long-term development plans illustrating Chinese rationality and patience about the nation’s revival.

In sum, the last two melodies are quite different from each other and from the first two, dissimilar emotions and connotations are expressed. The four melodies help realize strong interpersonal metafunctions (Table 3).

Table 3 Music qualities and discursive strategies for comparison.

	China		US & UK	
	3 <sup>rd</sup> melody	4 <sup>th</sup> melody	1 <sup>st</sup> melody	2 <sup>nd</sup> melody
1. Genre	march	pop music	pop music	pop music
2. Rhythm	strong	slow	fast	pulsating
3. Pitch	high	medium	high	high, wide range
4. Soundscape	hifi	lowfi	lowfi	lowfi
5. Mood	solemn	relaxed	happy	excited, nervous
1. Rhetoric	march metaphor	peace metaphor	clownery metaphor	clownery metaphor
2. Transitivity	high	low	medium	high
3. Modality	high	low	medium	high
Metafunction	interpersonal		interpersonal	

## 4. Discussion

Two interrelated strategies are found for positive self-representation, imbalanced comparison and rhetoric-based recontextualization.

Imbalanced comparison is the guideline running through the whole video. There is a big gap in time and information between Parts 2 and 3. Much more information is provided about the main social actor, President Xi with the Chinese electoral system to communicate a clear message: Chinese meritocratic politics have a rigorous selection process and reward success. Take the quantifiers as an example, abundant statistics show the benefits of the Chinese electoral system, but no data is given about the results that UK/US systems produce. This imbalance can be identified from variation in the discursive strategies of different modal discourses (Tables 1-3) between positive self and negative other.

Analysis of linguistic choices and audio-visual representations allows us to reveal the subtle details of the recontextualization process creating the “scripts” justifying the Chinese electoral system. The process is realized mainly by the predominant use of metaphor and metonymy.

At least seven metaphor clusters were found indicating distinct cultural schemas. The first and dominant one is the sports metaphor cluster: kungfu master, boxer and acrobat. The second, third and fourth metaphor clusters are verbal-visual: idol, road and step. The other three are auditory: march, peace and clownery. The sports and idol metaphor clusters depict leader identity, and the others describe the election process. Also, metonymy clusters indicate: 1) national identity: flag, map, money, and place; 2) election process: speech (mouth, trumpet).

These metaphors and metonymies contribute to the creation of a blended mental space<sup>[12]</sup> which results in recontextualization of the political discourse world. Visual and auditory devices endow the verbal with further meaning. These discursive modes combine to build a cohesive conceptual frame underlying the mental spaces triggered by the video.

In brief, the metaphor and metonymy clusters interact and unite to convey attitudinal meaning and ideology in three main categories: leader identity, national identity and election process. These categories correspond with the three major components of the transaction: participant (leader), circumstance (nation), and material process (election). Figure 1 summarizes how multimodal recontextualization is realized by metaphor and metonymy clusters through conceptual blending.

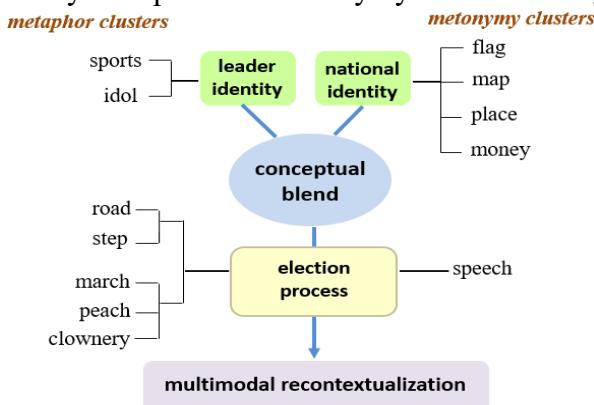


Figure 1 Rhetoric-based recontextualization.

This video basically follows the linguistic tradition of communist discourse while exhibiting new features. It provides evidence against boundaries between communist and democratic/populist discourses, as a possible result of globalization and marketization. The hybridity of this communist video discourse supports previous findings based on linguistic texts<sup>[7]</sup>.

## 5. Conclusion

The implicit and rational self-representation in the communist multimodal discourse could be regarded as an upgrade of its discursive tradition. The effectiveness of the cartoon can be explained by identifying it as a blended mental space incorporating elements of the textual, musical and visual

worlds. This study informs some robust ways that CDA can embrace meaning making in multimodal communication. The new way of positive self-representation should not be considered a unique policy change, but rather as part and parcel of a wide paradigm shift concerning the national identity construction. The research is not free from limitations. If more semiotic resources could be examined, including typography and pitch, the findings would be more persuasive.

## Acknowledgements

This work was supported by the Key Research Project of Philosophy and Social Sciences, Chinese Ministry of Education [17JZD039] and Fundamental Research Funds for the Central Universities of China (South China University of Technology) [ZDPY202214]. Shuo-zi WU, Wen-wen ZHENG and Ying QU share co-third authorship.

## References

- [1] Forceville, C. J., & Urios-Aparisi, E. (Eds.). (2009). *Multimodal metaphor*. Berlin: Mouton de Gruyter.
- [2] Baldry, A., & Thibault, P. J. (2006). *Multimodal transcription and text analysis*. London and Oakville, CT: Equinox.
- [3] Helland, K. I. (2018). Mona aka sad girl. *Discourse, Context & Media*, 23, 25-40.
- [4] Machin, D. (2010). *Analysing popular music*. London: Sage.
- [5] Smirnova, A. (2015). Aligning language to ideology. *Journal of Language and Politics*, 14(2), 205-232.
- [6] Chen, L. (2004). Evaluation in media texts. *Language in Society*, 33(5), 673-702.
- [7] Wang, J. (2017). Representing Chinese nationalism/patriotism through President Xi Jinping's 'Chinese dream' discourse. *Journal of Language and Politics*, 16(6), 830-848.
- [8] Machin, D., & Mayr, A. (2012). *How to do critical discourse analysis*. London: Sage.
- [9] Kress, G., & van Leeuwen, T. (2006). *Reading images*. London: Routledge.
- [10] Halliday, Michael A. K. (1985). *An Introduction to Functional Grammar*. London: Edward Arnold.
- [11] Van Dijk, Tuen A. (1991). *Racism and the Press*. London: Routledge.
- [12] Fauconnier, G., & Turner, M. (2002). *The way we think*. New York: Basic Books.

## Appendix

The video *How Leaders are Made* and its transcription are accessible upon request.