

# Multimodal Discourse Analysis of the Interpersonal Meaning of TV Advertisements

Dongyan Li

**Abstract**—This author attempts to explore the interpersonal meaning of TV advertisements with the *hilltop of coca cola jingle* selected as the studied case using both multimodal theory and systemic functional linguistics with a view to seeing how various semiotics or modes such as speech, color, movement, image and sound work together to create and enhance the interpersonal meaning. Hopefully, we can make some contributions to the studies on social semiotics as well as television advertisements and become more aware of the fact that language is just the tip of an iceberg among social semiotics.

**Index Terms**—Interpersonal meaning, multimodal discourse analysis, social semiotics, systemic functional linguistics.

## I. INTRODUCTION

Multimodal discourse approach adopts the view that diverse semiotic modes (e.g. language, images, music, sound, animation, and so on) are combined together to multiply the meaning in multimodal texts or communicative events. It has been widely used to a lot of fields ranging from classroom teaching environment, working settings to architecture, sculpture and theatres and so on.

In fact, advertising discourse is also another good instantiation of the application of multimodal discourse theory to practices as a practical genre endowed with its own unique characteristics. Although much research has been carried out as for its linguistic features like its rhetoric and cohesion, coherence, etc. both in the past and at present, the practical application of multimodal discourse theory to the analysis of TV advertisement still remains to be further tapped.

Advertising discourse is ultimately oriented towards the interaction with people, so this thesis attempts to explore the interpersonal meaning of TV advertisements with the *hilltop of coca cola jingle* song as the case resorting to multimodal discourse theory which derives from systemic functional linguistics with a view to seeing how various semiotics or modes such as speech, color, movement, image and sound work together to create and enhance the interpersonal meaning.

## II. STUDIES ON INTERPERSONAL MEANING FROM LANGUAGE TO OTHER SEMIOTICS

### A. *The Interpersonal Meaning in Language*

Initially, from the perspective of the social semiotics,

interpersonal meaning [1], viz. meaning as an exchange in the clause of language as one of the three meta-functions of language, is primarily referred as to categorizing the interrelations between the participants and the attitude of the speakers towards communicative acts or utterances which contains subsystems like mood, modality and key plays an indispensable part in functional systemic linguistics. The clause of language as an interactive event involves speaker, or writer, and audience. The speaker in the act of speaking adopts for himself a particular speech role, and in so doing assigns to the listener a complementary role which he wishes him to adopt in his turn. For example, in asking a question, a speaker is taking on the role of seeking information and requiring the listener to take on the role of supplying the information demanded.

### B. *The Interpersonal Meaning in Other Social Semiotics*

Later on, interpersonal meaning in language is extended into interactive meaning in image which establishes a kind of (imaginary) relationship between the producer and the viewer of the image [2].

On the basis of the investigations of interpersonal meaning by Halliday[1], Kress and Van Leeuwen[2], subsequent followers go into more and more detail to exploit interpersonal meaning from a multimodal perspective ranging from image, movement/action to sound, music and speech and the like [3]-[7]. As for the specific aspects of the interpersonal meaning in filmic genres such as in TV advertising discourse, interpersonal meaning is represented and conveyed through the synchronization of multiple meaning-making resources such as Kinesic, Proxemic, Rhythm, Gaze, Duration of the Image, Perspective of the image, etc..

Further discussion and multimodal analysis of interpersonal meaning are included as the temporal unfolding of the text goes on in TV advertising discourse in particular of the *hilltop of coca cola* advertisement as the case.

## III. METHODOLOGICAL FRAMEWORK

The research is essentially qualitative and the general principle is drawing on those representative researches on multimodal discourse analysis in particular of the outcomes concerning speech, image, sound, color and action/movement with the aim to pursue how interpersonal meaning is made in dynamic video texts. To be pretty much clearer, analytical procedures of this thesis are based on the following ones.

### A. *Research Questions*

It suffices to say the ultimate goal of advertising is to

Manuscript received February 13, 2016; revised April 22, 2016.  
D. Y. Li is with the Heilongjiang Bayi Agricultural University, Daqing 163319 China (e-mail: redsun\_2002@163.com).

persuade audience to take actions on the advertised products and/or services where interpersonal meaning of the advertisement plays a vital role. The author therefore is going to depart from analyzing the interpersonal meaning, viz. the interaction between the advertiser/addresser and audience/addressee in TV ads through scrutinizing the conflation of various modes concomitant in the dynamic unfolding of the digitalized TV text. To achieve the purpose of the study and make it feasible, the scope of the paper is confined to answer the following questions:

General question:

How is the interpersonal meaning represented and communicated in the hilltop of coca cola jingle via the co-articulation of speech, music and sound, image, and action/movement?

Specific questions:

- 1) How is the interpersonal meaning of image represented and communicated in TV/video advertising?
- 2) How is the interpersonal meaning of action/movement represented and communicated in TV/video advertising?
- 3) How is the interpersonal meaning of speech, music and sound represented and communicated in TV/video advertising?

#### B. Coca Cola As the Studied Case

The thesis is mainly qualitative-oriented with the hilltop of coca cola television advertisement chosen as the studied case, the video of which comes from the CD titled Classical TVC Advertising Appreciation by Xuelin Publishing House. Such a choice is made out of the ensuing reasons. In the first place, coca cola is the most renowned drinking brand all over the world; it's said that "coca cola", "God" and "her" are acclaimed as the most popular lexis during the last century. Above all, the hilltop advertisement among a number of other coca cola ads has proved to be well-known as one of the most successful and classical case ever since it was produced in 1971 [8]. In the end, despite advanced development of information technology and easy access to real-life data and the fact that the author may make use of online video materials to retrieve some classical and representative soft drinking advertisements, then download them from the internet, the practical application of multimodal discourse theories to do a detailed analysis of one classical case in fact is sufficient here.

The hilltop of coca cola advertisement is features its global pitch for peace, love and unity. In the setting, more than thirty veritable young, fresh-faced people from the globe, clad in their national costumes and clutching bottles of Coke, and directed in sign language, as the image of coca cola sang the song together on the hillside near Rome, very moving, impressing people with peace, love and world unity. It lasts some 54 seconds and the line of the song is as follows:

*I'd like to buy the world a home and furnish it with love.*

*Grow apple trees and honey bees and snow white turtle doves .*

*I'd like to teach the world to sing with perfect harmony.*

*I'd like to buy the world a Coke and keep it company.*

*That's the real thing.*

*What the world wants today.*

*It's the real thing, what the world wants today.*

#### C. General Picture of the Hilltop of Coca Cola Jingle

The *hilltop* occurs on a hillside near Rome; in the setting over 30 young people from different nations wearing their national costumes sang together, clutching bottles of coca cola. The moving song taking place in the grand scene is really a great hit. It lasts approximately 54 seconds with 8 shots made altogether. Shot 1 including 14 frames and shot 8 containing 10 frames are obviously the two longest. Shot 2 covers 6 frames with shot 3 and shot 5 comprising 7 frames. Shot 4 and shot 6 are both composed of 4 frames with 3 frames constituting shot 7. The dynamic changes are reflected from shot to shot but when the analysis is going on, some frames at times will be clipped from the video in order to get a clear comprehension of what is talked about for the sake of convenience. Now we are going to analyze how multiple modes, namely, image, kinesic action and sound, speech and music, are co-deployed to articulate the interpersonal meaning in this jingle with each shot taken as the basic unit.

### IV. MULTIMODAL DISCOURSE ANALYSIS

#### A. Interpersonal Meaning in Image

The choices in the mobile visual image made across the meaning-making semiotic resources here relate to the Color (CR), the Distance of the Image and the Perspective of the Image. The semiotic resource of color consists of a system with five elementary colors (black, white, yellow, red and blue) from which all other colors could be mixed. Seen clearly from the following three frames clipped out of the video and the annotation tier on Color, the color dark red of bottles of coke, of slogan and of brand name maintains throughout the largest part of the scene the dominant one in this hilltop jingle against the background of blue sky and green hill. People use color to actually try to energize or calm down people, or more broadly, to act on others[9]. In other words, color can be 'interpersonal', and could 'excite', 'inspire sentiments' and so on [10]. Likewise, dark red in the hilltop which is salient could excite people and visually appeal to the audience and be identified in such a way that the viewer can immediately grasp the significance of such choices with the aims at last achieved, namely, getting attention and arousing desire [11]. Also, the background colors of green and blue stand for peace and serenity with black and white young people singing together represent harmony and unity, which could be naturally connected with the spirit of the brand Coke going after a peaceful, loveable and unified world by the audience who watch it and will subconsciously have a good feel for Coke. It is obvious that such delicate choices from the paradigmatic system of color in this jingle can arouse target audience's attention and their interest in the advertised product, Coke here, and the image of the coca cola company.

In addition to Color, two other crucial elements in Image are Perspective and Distance. There are two further

subsystems in light of the former, viz. Horizontal and Vertical. According to [2], the horizontal axis usually reflects a degree of involvement. Base on the annotation on tier Perspective, during most of the time the viewer looks at the hilltop jingle from a frontal viewpoint which gives him/her the feeling of being involved in a part of the depicted world. Only one oblique perspective is shot due to the mobile position of the camera which pans leftward and the stationary position of the throng on the hilltop. This oblique angle doesn't detach the audience from the shot world in fact but it on the other way around invites audience to pay more attention to the prominence of the bottles of coke from a different viewpoint which can be judged from the middle frame among the following 3. The vertical angle, in contrast, usually expresses power, status and solidarity [2]. There are 7 median vertical shots and only one high shot, that is, the last shot in the hilltop jingle as shown in frame 3 below which is shot from high in a helicopter and leads audience to catch an eye view of the whole splendid hilltop scene. As most of the jingle is seen neither from above nor below, there is no power difference construed between the viewer and the participants; in other words, they are on the same level.

The distance of the image indicates social distance between the viewer and the represented participants [2], the hilltop jingle is composed of three very close shots, e.g. frame 4 and frame 5, and medium close shots like frame 1 respectively, one close shot, one medium long shot like frame 2 and one very long shot at the end of the scene in accordance with the annotation on tier Distance. That is, the participants are portrayed from very close shot which occurs in the beginning to attract viewer's attention gradually to close shot, then to medium close shot, medium long shot and to the very long shot at the end of the jingle scene which is exactly called aerial shot owing to the fact that it is shot from a high angle in a helicopter. This progression of the mobile frames running in shots interpersonally orients the viewer towards the image and furthermore contributes largely to the intimate and socially closer relations constructed between the participants in the hilltop world and the viewer who could be brought into the depicted world easily.

#### *B. Interpersonal Meaning in Kinesic Action or Movement*

Kinesic action or movement is a meaning-making semiotic resource by means of which participants in interactions relate to one another [4], which is, in other words, about the means of constructing and expressing interpersonal relations through action or movement. Three types of action are distinguished by Martinec and presenting action (communicating meanings regard the immediate context of the here-and-now) as one of those three is mainly concerned with the movement afforded in the hilltop jingle. It mostly consists of movements of whole bodies and of facial muscles. The interpersonal resources of presenting action could further be described as engagement, modality and affect. Engagement regards interpersonal meanings communicated by interacting persons' body distance and angle. Modality is about the willingness or unwillingness, and tentativeness or assuredness with which an act is performed. Affect regards display of emotions, which takes place for the most part in the face.

In light of engagement, modality and affect given above, movement may describe or designate some situation and evaluate this situation by adopting a particular interpersonal orientation towards it, e.g. pleasure, disapproval, and so on. Let's look at how the specific movement is exercised and interpersonally conveyed in the coca cola jingle. Apparently we can see that Gesture, Gaze and Facial Expression in kinesic movement/action are highlighted in the hilltop dynamic text based on the following 3 annotation tiers together with frame 4 and frame 5 extracted from the video. Bodily movement of the participants is interpersonally directed frontally towards the viewer which keeps congruous with the horizontal perspective of the image and naturally engages the viewer and creates a concerted feeling during the chorus sung by the young people. What is more important, the same gesture is made up of forearms on hold at a right angle slightly above the waists of the participants with hands clutching bottles of coke horizontal to the ground (e.g. frame 4) and lasts for some 28 seconds during the total seconds. According to Poynton's notion of amplification of movement [3], a given movement or part thereof can be articulated with decreased or increased speed, or by repetition of the same movement, or by some form of embellishment of the basic movement so as to invest it with heightened subjective commitment or intensity. the consistent and continuous action of gesture which is transcribed as GE: I on the annotation tier on Gesture below belongs to 'repetition of the same movement' and so this specific action signifies that the intense attitudinal stance adopted by the addresser to the advertised Coke is assuredness of the advertised product and strong willingness to express their positive values as well. At the same time, smiling (e.g. frame 5) as annotated result displays below constitutes a key one in facial expression though the total time of smiling is much shorter compared with that of gesture owing to its being a song jingle. In this case, it signals participant's commitment and confidence in what they are doing to coca cola as well as his or her solidarity with the viewers, creating affinity by smiling between the depicted participants in the dynamic text world and us viewers who watch it in reality.

Gaze maintains another vital meaning-making semiotic mode in kinesic action. In reference to Baldry and Thibault [3], the annotation on Gaze is transcribed as GE: F; GE: I and GE: C; GE: V. The former indicates that gaze in distance and in orientation is respectively far and indeterminate due to off-screen; the latter suggests that gaze in distance an in orientation is close and oriented to the viewer. Judging from the annotated result below, the participants in the jingle gaze at the viewer at a close distance about 15 seconds establishing a direct eye contact with him/her and bridging the psychological distance between each other. Although the participants gaze far away off-screen more than 20 seconds, it doesn't mean the aversion from the viewer here in fact, because the setting is on the hilltop and the mobile camera shoots from high above; on the contrary, it appeals more to the audience thanks to the associative link with the depicted world provoked by the far and off-screen gaze.

#### *C. Interpersonal Meaning in Speech, Music and Sound*

Speech, music and sound are brought together on the

assumption that they all have characteristics in common providing a basis for talking about them and transcribing them in a unified way rather than as entirely separate phenomena [3]. That is, speech, music and sound are highly synchronized as the dynamic text unfolds in time and the annotated tiers on the Lines of the jingle song, Sonic Interaction together with Rhythm Group shown below reveal this. Sound can be used to represent our environment and to represent the actions and interactions of people [7]. These sound acts themselves are realized by a configuration of choices from all the sound resources available in the given context, for instance, by certain rhythmic choices (e.g. a medium tempo, accented syllables); by a choice of 'social distance' (soft, hence close) and so on.

Now comes the interactive meaning configured in the speech, music and sound of this fifty-four-second coca cola jingle song demarcated through the concordant survey of the three semiotic modes together. Where speech is concerned, based on the annotation on tier Line of the song with accented and lengthened syllables, we know that some accented and lengthened rhythmic choices marked as in round brackets (\*), (\*\*) and (!! ) such as (\*\*Coke), (\*\*real), (\*\*company), thing (!!), wor (!! )ld, today (!!), cola (!! ) sung with remarkable emphasis and procrastination respectively in a recurrent way during the whole sequence communicate intensely to the addressee or the target consumers the emphasized import of the advertised coca cola—it's the real thing that the world wants today and a perfect company. The lines obviously offer goods & services rather than demand them. The mood choices in the flow-in-flux text are restricted to declaratives and do not make use of imperatives. No mood adjuncts such as probably expressing uncertainty are used and in this way the speech shoulders its responsibility for the truth value of the message. A strong identification with and affirmation of the brand could be made by the target audience and such an empathy as love, peace and harmony may be transferred from the singing participants in the depicted world to the audience in reality and render the vicarious feeling among them. That is the salient or prominent interpersonal meaning of this jingle song intending to convey.

As the jingle song is a leader-group chorus in deed and the interactivity can also be represented and communicated through the choices from sonic interactions and rhythm group. In shot 1, for example, the soundtrack starts with a median tempo and a soft voice sung with no accompaniment by a female soloist who is the initiator delivering I'd like to buy the world a home and furnish it with love, which are correspondingly annotated on Sonic Interaction displayed below as SE: I (initiator in sequential sequence); RGT: M (median tempo in rhythm group); RGV: P (soft volume in rhythm group). In collaboration with the very close up of the visual track, it functions to attract the attentions of the audience. Gradually to the accompaniment of the electronic guitar the female leader singer and the choral group produce the same piece of sound grow apple trees and honey bees and snow white turtle doves together as a unison still at a median tempo and in a soft voice but the dominant sound is that of the leader in contrast to that of the group, which occurs both in a simultaneous and sequential way and is represented by SI: UU (simultaneous unision unblended) and SE: SRF1S

(sequential segregated fully stated). As the song progresses, the tempo of the jingle song is basically at a median level, and the volume at the first half of the song is soft and ascends gradually from shot 3 to the loudest from shot 6 till the end of the song. Altogether, the lines I'd like to buy the world a Coke and keep it company (sung twice); It's the real thing (three times); A company. (twice); What the world want today (three times) constitute the repetitive sonic interactions and rhythmic melody in congruity with the tempo and the volume.

The continuous simultaneous speech, music and sound of more than two speakers are maximally oriented toward interactivity, which are geared towards the 'rapport' function of speech and involve equality, cooperation, collaboration and the joint production of meanings. The same range exists true when speech and music merge in song [7]. The same is also true of the hilltop of coca cola jingle because more than 30 people from various nations are on the hilltop singing in unison, which conveys the equality and collaboration with the world. Thus the rapport is established between the singers and the audience and a maximum interactivity is accordingly performed.

#### *D. The Integrated Multimodal Analysis*

For the sake of convenience, interpersonal meaning constructed through the multiple modes viz. image, kinesic action/movement, speech, music and sound is analyzed as above separately. The integration of those modes is nevertheless supposed to be simultaneously taken into account as the dynamic text unfolds in time.

Take shot 6 and shot 7 as examples to interpret the simultaneous systemic choices across the meaning-making coordinating modes. The two shots (as frame 4 and frame 5 display) constitute a sequence in time as the line I'd like to buy the world a Coke and keep it company. It's the real thing is sung with a median tempo together in which the lexis 'buy', 'Coke' and 'real' are prominently accentuated alongside 'and' 'company', 'thing' are lengthened in very loud voice. Both of the shots are very close in distance, frontal in horizontal angle, median in vertical perspective with shot 6 featuring the red bottle of Coke held in hands and shot 7 featuring a smiling face of the leading singer singing 'It's the real thing.' All of those choices together leave the audience a strong visual and aural impact which can easily grasp the attention of the audience at once, arouse the identification with the advertised Coke and finally take actions on the purchase.

The interpersonal meaning of the hilltop of coca cola jingle represented and communicated by the systemic choices from the multiple modes like image, kinesic action/movement, speech, music and sound has been discussed and analyzed as above. In fact, all the salient modes working hand in hand renders this jingle song so impressive and memorable that the brand, slogan and image of Coke could be more identifiable and more influential throughout the world via the publicity of its global pitch for love, peace and unity. Definitely, it is the choices made across semiotic resources in their interaction with but not in isolation from other resources that makes the interpersonal meaning between the participants and the target audience. That's the key point worthy of note here.

## V. CONCLUSION

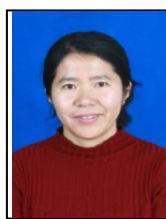
The thesis has made an attempt to probe into the way the interpersonal meaning is represented and communicated through the multiple modes such as image (color, perspective, distance), movement (gaze, gesture, facial expression) as well as speech, music and sound (accented syllables, sonic interaction, rhythm group) with the hilltop of coca cola jingle as the studied case in television adverting discourse through the utilization of related multimodal discourse theories in terms of image, movement, speech, music and sound. In conclusion, the concomitant multiple semiotic modes in the dynamic unfolding of the digitalized text are co-working to generate and communicate the interpersonal meaning that is enacted between the depicted participants and the viewer or audience. By doing this research, the author really hopes it may make some contributions to the further studies on dynamic television advertising discourse as well as to the application of multimodal discourse theory to practices at home.

Due to the limited time and the academic ability of the author, there are some limitations to this research. The first is that the author knows not very much with respect to knowledge involving music that the multimodal analysis on this part seems sort of weak. The second is that the author is more or less restricted to the use of visual art, which may affect the perfect construal of the visual impact of TV advertisements on the audiences. Here the weaknesses are pointed out hoping that similar difficulties could be overcome in the future.

## REFERENCES

- [1] M. A. K. Halliday, *An Introduction to Functional Grammar*, London: Edward Arnold, 1994.

- [2] G. Kress and T. Van Leeuwen, *Reading Images: The Grammar of Visual Design*, London & New York: Routledge, 1996.
- [3] A. Baldry and P. Thibault, *Multimodal Transcription and Text Analysis: A Multimedia Toolkit and Coursebook*, London Oakville: Equinox, 2006.
- [4] R. Martinec, "Interpersonal resources in action," *Semiotica*, vol. 135, pp. 117-145, 2001.
- [5] K. O' Halloran, "Visual semiosis in film," in *Multimodal Discourse Analysis: Systemic Functional Perspectives*, K. O' Halloran, Ed. London and New York: Continuum, 2004, pp. 109-130.
- [6] M. O' Toole, *The Language of Displayed Art*, London: Leicester University Press, 2004.
- [7] T. Van Leeuwen, *Speech, Music, Sound*, London: Macmillan, 1999.
- [8] B. Kanner, *The 100 Best TV Commercials*, New York: Times Business, 1999.
- [9] G. Kress and T. Van Leeuwen, *Multimodal Discourse: The Modes and Media of Contemporary Communication*, London: Arnold, 2001.
- [10] G. Kress, "Color as a semiotic mode: notes for a grammar of color," *Visual Communication*, vol. 1, pp. 343-368.
- [11] G.-w. Huang, *Discourse Analysis of the Theory and Practice of Advertising*, Shanghai: Shanghai foreign language education press, 2001.



**Dongyan Li** was born in Binxian county, China in 1978. Her educational background is listed as follows: a M.A. in applied linguistics from Beijing University of Aeronautics and Astronautics, Beijing, China, in 2009. The degrees should be listed with type of degree in what field, which institution, city, state or country, and year degree was earned. The author's major field of study should be lower-cased.

She has worked as a Chinese university English teacher for 13 years. She is also a member of Daqing Foreign Language Teaching Association and of Heilongjiang Business English Society.

Education Faculty, 2002-2016 pedagogic instructor and lecturer. The College of Social Sciences and Humanities, Hei Longjiang Bayi Agricultural University, Daqing, China. She is interested in English education especially the pedagogical issues, discourse analysis and teacher development. She published many articles in these issues, in Chinese and English.