

A Study on Verbal and Image Relations in Multimodal Texts From the Perspective of Intersemiotic Complementarity

HU Yong^[a,*]; QIU Qing^[b]

^[a]Foreign Languages College, Jiangxi Normal University, Nanchang, China.

^[b]School of Intercultural Studies, Jiangxi Normal University, Nanchang, China.

*Corresponding author.

Supported by the Critical Discourse Analysis on Multimodal Text in the Context of New Media (YY162018), a Social Humanities Project in Universities of Jiangxi Province in 2016; and A Study on Verbal and Visual Complementarity in Multimodal Narrative Discourse (YY17210), a Social Humanities Project in Universities of Jiangxi Province in 2017.

Received 12 July 2020; accepted 13 September 2020
Published online 26 October 2020

Abstract

Multimodal texts play a key role in the context of new media with the acceleration of modern technology of communication. Instead of merely keeping an eye on the verbal semiotic, recent developments in multimodal studies place a particular focus on the multi-semiotic texts involving some other non-verbal semiotics. The primary focus of this paper is placed on the textual meaning which is realised in both visual and verbal semiotics within children's picture books. In this paper, the textual features of visual components are first identified throughout the database, and it then moves to the analysis of compositional intersemiotic complementarity between verbal and visual semiotic systems to create messages in multimodal texts.

Key words: Multimodal texts; Intersemiotic complementarity; Verbal and image relations

Hu, Y., & Qiu, Q. (2020). A Study on Verbal and Image Relations in Multimodal Texts From the Perspective of Intersemiotic Complementarity. *Canadian Social Science*, 16(10), 50-56. Available from: <http://www.cscanada.net/index.php/css/article/view/11938>
DOI: <http://dx.doi.org/10.3968/11938>

Along with the acceleration of modern technology of communication, we are engaged in social interactions

which are composed of more than one mode. Instead of merely keeping an eye on the verbal semiotic, recent developments in multimodal studies place a particular focus on the multi-semiotic texts involving some other non-verbal semiotics, such as on visual images. Children's picture books serve as a typical type of multimodal texts which are key to the development for children. And thus the primary focus in this paper is placed on the textual meaning which is realised in both visual and verbal semiotics within children's picture books. In this paper, the textual features of visual components are first identified throughout the database, and it then moves to the analysis of compositional intersemiotic complementarity between verbal and visual semiotic systems to create coherent messages, using the picture book *Reunion* as an exemplar.

1. THE TEXTUAL METAFUNCTION IN MULTIMODAL TEXTS

Systemic functional linguistics proposes three metafunctions: ideational, interpersonal and textual metafunctions (Halliday & Matthiessen 2004). In comparison to the ideational and interpersonal metafunctions, the textual metafunction enables ideational and interpersonal meanings to be organised as a meaningful whole and thus creates "periodic, wave-like patterns of discourse" (Halliday 1985: 8). According to Halliday (1994/2000: 334), the following features make up the textual component in the grammar of English:

(A) Structural: (1) thematic structure: Theme and Rheme and (2) information structure and focus: Given and New

(B) Cohesive: (1) reference; (2) ellipsis and substitution; (3) conjunction and (4) lexical cohesion

The textual metafunction is the enabling metafunction that provides the resources for presenting ideational and interpersonal meanings as a flow of information in a particular text, and gives texture to a piece of discourse.

Hence clauses in this sense can be perceived as messages. From the perspective of structural texture, Theme usually refers to the starting point of a clause which “locates and orients the clause within its context”, whereas the remainder of that clause consists of the Rheme (Halliday & Matthiessen 2004: 64). There is a direct semantic relationship between thematic structure (Theme + Rheme) and information structure (Given + New). Although they are closely related, Theme + Rheme and Given + New should not be seen as the same thing. The Theme is what I (the speaker) determine to take as the point of departure and thus the Theme + Rheme structure is governed by the clause order. Whereas the Given is what you (as the listener) already know about or have accessible to you (Halliday & Matthiessen 2004). Compared with the Theme + Rheme representing speaker-oriented structure, Given + New is listener-oriented structure.

Both theme and information are realised as structural configurations of textual functions, though they are different in some domains. Theme is a system of clauses that is realised by the sequence in which the elements within that clause are placed. Information is not equal to Theme, and it is not a system of clauses, but has its own realisation in the form of tonic prominence. Although with differences between them, Theme and Information “operate at the level of the clause and are parallel and interrelated systems”, and therefore constitute the internal resources for texturing the clause as a message (Chang, 2004, p.115). Besides the internal relationship, it is necessary to take into account the external relationship between one clause or clause complex and another. Cohesion as non-structural resources for discourse “occurs where the INTERPRETATION of some element in the discourse is dependent on that of another. The one PRESUPPOSES the other, in the sense that it cannot be effectively decoded except by recourse to it” (Halliday & Hasan 1976, p. 4). Cohesion in English, according to Halliday and Hasan (1976), can be categorised into two major types: grammatical cohesion, including reference, ellipsis, conjunction, and lexical organisation; and lexical cohesion including repetition synonymy/antonymy, hyponymy/meronymy and collocation.

2. THE LOGICO-SEMANTIC RELATIONS ON IMAGE-TEXT RELATIONS

Martinec and Salway (2005) bring forth a detailed explanation of the generalized system of text-image relations based on a variety of different genres of multisemiotic discourse in which texts and images co-occur and can be viewed as being related. Their system of image-text relations derives mainly from Barthes’ studies (e.g. Barthes 1977) on the relations between text and image (which are in terms of anchorage and relay) and from the SFL architecture of clause relations such

as dependency (parataxis or hypotaxis), expansion and projection (e.g Halliday & Matthiessen, 2004). The system of text-image relations complement those proposed in the semantic model and should be useful for distinguishing various types of relations between text and image in different genres of multimodal discourse, especially in “(genuinely) new and old media” (Martinec & Salway, 2005, p. 337).

The generalised network of the semantics of text-image relations proposed by Martinec and Salway has two primary subsystems: Status and Logico-semantic relations. Similar with the relationship between linguistic clauses in a clause complex, the system of Status is primarily concerned with whether an image can be independent, complementary, or subordinate with the text. Thus the equal and unequal status serve as two basic subsystems within the system of Status. As to the unequal relation between texts and images, one of them modifies the other and hence the modifying element is perceived to be dependent on the modified one. Equal status between texts and images can be further divided into independent relation and complementary relation. On the one hand, a text and an image are considered independent and the relation between them is viewed as equal when “they are joined on an equal footing and there are no signs of one modifying the other” (Martinec and Salway, 2005, p.343). On the other hand, a text and an image are considered as complementary when they are joined equally and one modifies the other in the process of making meaning.

Drawing on the ideas of Halliday’s (e.g. Halliday 1994/2000; Halliday and Matthiessen 2004/2008) grammatical categories of projection and expansion, Martinec and Salway propose the Logico-semantic network in which the relation between text and image may be categorised as elaboration, extension, or enhancement, or locution or idea. Elaboration between images and texts can be further divided into two kinds of relations: exposition and exemplification. Within the system of elaboration, the exposition refers to the relation in which an image and a text are of the same level of generality, whereas the exemplification suggests that the image and the text are of different levels of generality. To be specific, exemplification as a level of delicacy within the system of elaboration serves as the entry point for two more choices: [text more general] or [image more general]. Martinec and Salway identify the circumstantial relations of time, place and reason/purpose and add that for an image to be considered enhancing a text (or vice versa), it should be in relation to its ideational content. Despite expansion subsystem within Logic-semantic system, there is another logic-semantic relation between texts and images: projection. Projection can be subdivided into two particular kinds, depending on whether an exact wording is quoted (adopting Halliday’s technical term “locution”) or an approximate meaning is represented (adopting Halliday’s technical term “idea”). Generally speaking,

projection as a logic-semantic relation seems to appear in the co-occurrence of image and text as such: in comic strips that locutions are conventionally enclosed in speech bubbles and ideas in thought bubbles; or in combinations of diagrams and text, for instance, the combinations

of diagrams and explanatory text found in textbooks, scientific publications and economic magazines. Martinec and Salway's network for image-text relation is presented in Figure 1.

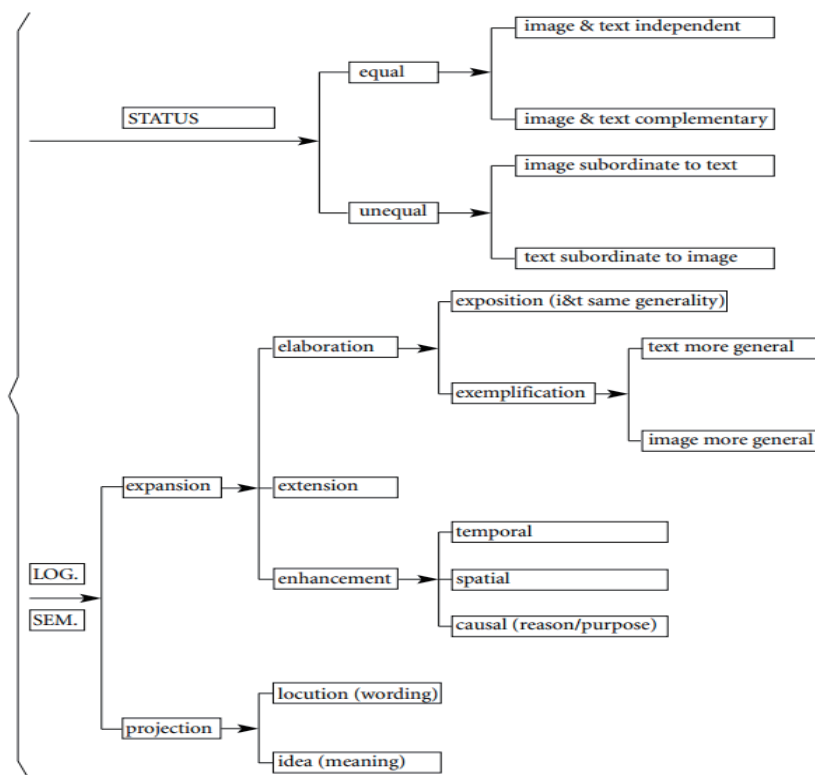


Figure 1
The network of combined Status and Logico-semantic relations between text and image (Martinec & Salway, 2005, p.358)

Although Martinec and Salway's network provides insightful analytical tool for explaining intermodal relations, there are some inconsistencies in the Status system.

3. THE INTERMODAL COMPLEMENTARITY FRAMEWORK

Unlike some other SFL based approaches to the analysis of multisemiotic discourse, which center upon the text-forming relation between different semiotic resources within a multisemiotic discourse, such as the analysis focusing on intersemiotic cohesion between verbal texts and images (Royce, 2007), on logico-semantic relations between words and visuals (Martinec & Salway, 2005), Painter and Martin (Painter & Martin, 2011; Painter et al., 2013) endeavour to propose an integrated framework to explore the intermodal complementarity across image and verbiage within a visual narrative, with great attention to the construction of intermodal relation by examining the various meaning realised by each semiotic within the multisemiotic text. As a cornerstone of the

theoretical construct of their framework, the availability of "descriptions for both linguistic and imagic semiotic systems modelled as sets or 'systems' of meaning choices with specifiable realisations" is key to the analysis of multisemiotic texts (Painter & Martin, 2011, p.133). For the modelling of linguistic semiotic systems, they draw upon the insight of metafunctional diversity in SFL (Halliday, 1994/2000; Halliday & Matthiessen, 2004) and appraisal resources (Martin & White, 2005; Martin & Rose, 2007); while for building visual semiotic systems, they develop a systematic framework for analysing visual narrative on the basis on Kress and van Leeuwen's (1996) visual grammar (for the details of the framework of visual narrative, see Section 3.3). With these choices of semiotic systems and realisations, Painter et al. conduct a comparative study on visual narratives by checking the diverse affordances for meanings of both verbal and visual semiotics (e.g. Painter & Martin, 2011; Painter et al., 2013). Based on the concept of "instantiation", Painter and Martin (2011) put forwards a comprehensive analytical framework of intermodal complementarity to account for the meaning making between verbal and

visual semiotics and the relation between the totality of meaning choices of the multimodal discourse and the specificity of the instantiated text (both verbal and visual texts) within it. In order to examine the interplay of verbal text and visual text in a visual narrative, Painter et al. map out the complementarities between meaning systems across language and image within three metafunctions (Painter & Martin, 2011, Painter et al., 2013).

The complementary systems of meaning across visual and verbal modalities may serve as a schematic presentation, suggesting the general feature of linguistic (and visual) meaning potential and the general forms expressed by the specific realisations.

4. TEXTUAL AND VISUAL ANALYSIS OF THE PICTURE BOOK

The textual metafunction of language is concerned with the linguistic resources for creating coherent texts, enabling the interpersonal and ideational meanings to be configured as a meaningful whole with relevance to the contexts in which they are produced and interpreted. In SFL, the thematic structure that gives the clause its character as a message can be generally regarded as a structure comprising thematic element and rhematic element (Halliday & Matthiessen, 2004). Theme is considered to be “the starting point of the message”, referring to “what the clause is going to be about” (Halliday 1985, p.39). Apart from the Theme of the clause, the rest of the message is labelled as Rheme, which usually follows the initial element in the clause.

By utilising different thematic structure, the writer/speaker intends to project a same message from divergent points of view. Theme in this sense serves as a meaningful choice that “sets the scene for the clause itself and positions it in the relation to the unfolding text” (Halliday & Matthiessen, 2004, p.66). Within these concepts, we are moving to the verbal analysis of *Reunion*, with specific reference to the typology of themes that predominates in the verbiage, aiming to see how this picture book story is structurally organised. As shown in Table 4-1, both simple and multiple themes have been identified, with simple type of theme appearing more frequently (88.8%).

Table 1
Simple and multiple themes in *Reunion*

Theme	Number	Percentage
Simple theme	79	88.8
Multiple theme	10	11.2
Total	89	100%

Evidence of this fact is shown in most of the clauses within this picture book where the thematic slots of them tend to be realised by the type of sole ideational elements

that makes reference to the two main protagonists within the book, I (the little girl protagonist) (26.7%) and the father (24%).

For instance, the first clause of the picture book story introduces the father as one of the main protagonists as he is located in the thematic position of the clause and continues in such slot in the following clauses. The rhematic parts of the two clauses inform the child reader of his occupations and hard work outside. A number of similar uses of thematic structure have been found in the picture book story. The thematic positions of these clauses are taken up by two major types. The first type in which thematic position is occupied by the girl protagonist (including the pronoun “我” “I” and the nickname “毛毛” “Maomao”) occurs in the story to represent the activities taken by her on the holiday of the Spring Festival, conveying some interesting Chinese customs to the child reader. In this sense, the status of the girl protagonist within the verbal text is highlighted and the reader’s attention is attracted by the foregrounding of the protagonist to share the feelings and moods of her.

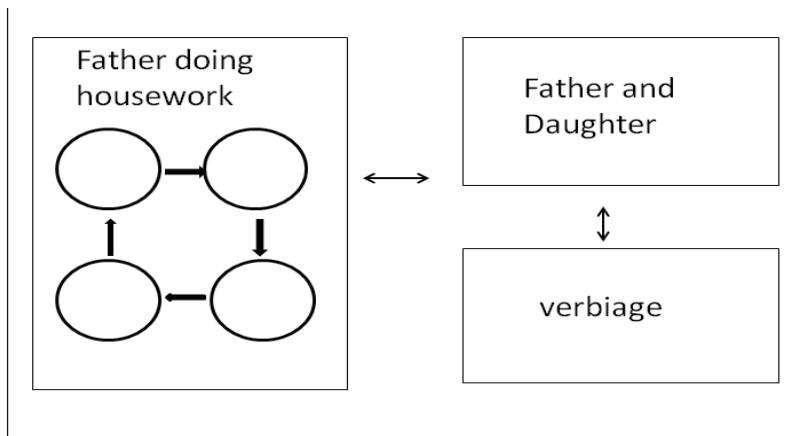
The high utilisation of simple theme in language plays an essential role in bringing the child reader’s attention to the major protagonists as well as their activities in which they jointly participate during the precious time of reunion. Coupled with the high frequency of simple choices, there are frequently-adopted options of unmarked themes identified in the verbal text where the verbal component of themes is able to fulfill the syntactic function of the subject of that clause. Consequently, the unmarked realisations of themes in language complement the purpose and strategy of facilitating the child reader’s interpretation of the narrative plot and his/her understanding of the empathising stance.

Standing on the shoes of the girl protagonist and her father, we are likely to experience two differentiated kinds of emotions: happiness for the annual reunion and sadness for the departure.

With regard to the visual analysis of the bi-modal texts within the picture book, especially the choices of FOCUS in the picture book story, this section explores in detail the visual elements placed in frames to manage the reader’s attention. The focus groups are defined as some kinds of eye-fuls constituting a pulse of message. Within the story, the choices made from the system of FOCUS in effect emphasises experiential meaning and highlight several focus groups which deserve repeated looks. A typical example as such can be found in the eighth double-page spread, where the choices of focus groups have been schematically analysed in Figure 2. In this visual composition, there appears a [centrifocal: plorised: orthogonal: horizontal] relationship between the image on the verso page and the visual composition composed of an image and a verbal text on the right. On the verso page, the various versions of the father in conducting

different kinds of housework form an [iterating: aligned] manifestation, emphasising that the laborious father involved in a series of activity sequences is brought to the reader's notice as a principal focal point. The ideational meaning is reinforced by the ongoing process of various activity sequences taken by the depicted character in a successive relation. On the recto page, there is obviously a choice of [plurised: orthogonal: vertical] connecting the image and verbiage, each of which occupies the same amount of visual space and thus affords equal semantic weight. By utilising such choice of focus groups, the visual composition stress two primary focal points in a

coherent way in which the verbiage supports the image by elaborating on it. In the upper image, the ideational meaning is strengthened by the deictic vectors between the represented participants. In this case, the deictic vectors are formed by the orientation and gaze interaction between the father and the girl protagonist, which serves as an important means in realising interpersonal meanings of proximity. A social contact has been established by the face-to-face placement of the two depicted characters, and a close intimacy can be indicated from the close social distance and the embracing posture between them, as well as by their facial expressions of care and delight.



Note: $\square \leftrightarrow \square$ [and other orientations]= polarised focus groups
 $\circ \rightarrow \circ$ [and other orientations]= iterating focus groups

Figure 2
 Focus groups represented in the eighth spread in *Reunion*

5. THE INTERPLAY OF WORDS AND IMAGES AT THE TEXTUAL LEVEL

With the ideational and interpersonal meaning completed in previous chapter, this section conducts an integrated analysis on the different meanings committed by verbal and visual semiotics within the tale with respect to

ideational, interpersonal and textual meaning. To facilitate the analysis, the last spread of the picture book has been drawn on to explore the different commitment of both two modes. It selects Illustration 1 (the final spread of the book) as the exemplar to investigate the commitment of meaning related to verbal and visual semiotic resources.



Illustration 1 from *Reunion* (2008)

Concerning the ideational meaning committed in the visual image, we can observe an action depicted (a little girl putting a coin onto the hand of a man), as indicated by the vectors informed by her arm and the coin, and the middle-aged man looking down at her (with invisible gaze vector towards the little girl, indicating a perception process). As far as the **character** manifestation is concerned, we are provided a complete depiction of the man but only the profile of the little girl. Through their depicted attributes (e.g. dress, size, hair, skin colour), we can also infer a variety of factors related to them such as the age, race, or even social class and social relationship. The represented participants within the image are depicted in a very close social distance (e.g. the girl being held tightly by the arms of the man), suggesting a kind of family role as “daughter and father”. Although the reappearance of the girl protagonist at this moment is portrayed with profile, the child reader is able to make an identification of the character through the “marked attributes” of her (e.g. red coat and pink-grey hat). In view of the ideational meanings instantiated by the linguistic text, some of actions (e.g. putting the coin, nodding, hugging) are described verbally and a verbal action is taken by the girl protagonist to his father, with the projection of the direct speech. However, there are no cues to deduce that the depicted characters are illustrated visually as talking to each other (e.g. mouth closed, without speech bubble). The manifestation of character is not construed verbally at all except from the identification of the familiar role as daughter and father.

Looking at the visual composition through an interpersonal lens, the reader is offered an image where a little girl being held tightly by her father. As for the visual focalisation, the depicted characters have no eye contact with us and hence we are kept outside of the story world. Nevertheless, we are positioned to see the profile of the girl protagonist with the father’s eyes, since the father is looking straight at his daughter. A choice of mediated viewing has been encoded into the image so that we are placed temporarily in the stance of the father to make a vicarious contact with him. We are then encouraged to get involved in the important moment of the story to share with the father a mood of sadness and being reluctant to part with his daughter. Unlike the commonly accepted choice of minimal character depiction in picture books, the represented characters are depicted in a generic style, indicating an empathic role in evoking attitudes in the child reader. Different from the minimalist depiction keeping the reader as detached observer outside the story, the generic style in the image invites the reader to see themselves in the focalising character’s view, creating alignment and affinity between the reader and the represented characters. When it comes to the involvement, combined with the generic style of character depiction which creates an empathetic pathos, we are kept at a

close social distance with the depicted characters by a close shot to see clearly their facial expression. In terms of the affiliation between the depicted characters in the image, the father looking down at his daughter symbolises the former’s power or authority. Furthermore, the two represented characters are in a close proximity to each other, and the father’s posture in the image (embracing the girl protagonist in arms) ensures an actual physical touch between the two, reinforcing the interpersonal affinity between them as well as evoking a mood of sadness in the child reader.

Turning to the visual expression of **feeling**, we are likely to share an emotion of sadness in view of the facial expressions and their bodily postures. This feeling is underlined by the ambience of the image where the represented characters are suited in a both cool and familiar ambience through using of a full palette of dark hues despite that the girl protagonist is wearing a red cloth (the red colour in this case serving only as the maintenance of character depiction). The choice of vibrant and cool ambience thus contributes to building a depressed and sad emotion for the father in parting with his daughter. The verbal text by comparison commits some differing meanings in relation to the interpersonal aspect. The interpersonal meaning systems in play are those of proximity between the two represented characters, realised via choices in naming (爸爸, Dad) and verbal projection. Through the projected direct speech and a series of feedbacks given by the father (e.g. promising by nodding his head), a more experiential reader is able to infer the feeling of sadness and upset.

Lastly, with regard to the textual meaning encoded in the image, we need to consider **prominence** in terms of framing, focus group and intermodal integration. The double-page spread is regarded as a kind of “micro” frame where the verbal text is framed off from the image by being placed on a separate page. A choice of complementary layout is adopted in the spread where the verbal text and visual image reside in its own space with equal semantic weight. In this sense, we are introduced to the picture book story by reading separately the visual and verbal semiosis in the spread. In view of framing device used in the spread, an option of [unbound: decontextualised: individuated] in framing encourages the reader to enter the story world. The decontextualisation choice has a strong effect in highlighting the affect or behaviour of the depicted characters, facilitating some of the interpersonal meanings by shifting our attention to the depicted characters. Within the unbound image, a removal of circumstantial setting narrows the distance between the depicted story world and the reader’s world, so that the represented characters are less constrained by the unnecessary circumstantial details. As for the focus group, the represented characters are arranged in a polarised relation, implying that both of them are

viewed as principal focal point of ideational meaning. In addition, interpersonal meanings are expressed by the deictic vectors between the two characters formed by the face to face orientation and gaze interaction, indicating the close proximity and affiliation between them. Thus the choice of focus group within the image encourages our immediate apprehension of the two characters as a pair of focal points rather than focusing merely on any one of them. Thus the visual textual choices contribute to organising the ideational and interpersonal realisations and all three metafunctions are co-patterned in the visual text. However, as for the verbal text at textual level, there is no such commitment of meaning to help to build primary focal points.

CONCLUSION

This paper conducts an intermodal analysis of textual meanings in visual and verbal texts within children's picture books. It adopts SFL as the linguistic analytical tool to analyse the verbal meanings and Intermodal Analysis as the visual analytical tool to probe into the visual meanings. Through the analysis of the differing commitment of meaning between visual and verbal text, it is noticeable that there is more ideational meaning committed in the visual text than the verbal one in respect to the depiction of the characters. However, at the textual dimension, there is more meaning committed by the visual semiotic than the verbal one, without complementary meaning choices realised in the verbiage.

REFERENCES

- Barthes, R. (1977). Rhetoric of the image. In R. Barthes (Ed.), *Image–music–text*. London: Fontana.
- Chang, C. G. (2004). *English idioms and interpersonal meanings*. Guangzhou: Sun Yat-sen University Press.
- Halliday, M. A. K. (1985). *An introduction to functional grammar*. London: Arnold.
- Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. London: Longman.
- Halliday, M. A. K., & Matthiessen, C. M. I. M. (2004). *An introduction to functional grammar* (3rd ed.). London: Arnold.
- Kress, G., & van Leeuwen, T. (1996). *Reading images: The grammar of visual design*. London: Routledge.
- Martinec, R., & Salway, A. (2005). A system for image–text relations in new (and old) media. *Visual Communication*, 4(3).
- Painter, C., & Martin, J. R. (2011). Intermodal complementarity: modelling affordances across image and verbiage in children's picture books. In G. W. Huang (Ed.), *Research in functional linguistics and discourse analysis (Vol.3)*. Beijing: Higher Education Press.
- Painter, C., Martin, J. R., & Unsworth, L. (2013). *Reading visual narratives: image analysis of children's picture books*. London: Equinox.
- Royce, T. (2007). Intersemiotic Complementarity: A framework for multimodal discourse analysis. In T. Royce and W. Bowcher (Eds.), *New directions in the analysis of multimodal discourse*. Mahwah, New Jersey: Lawrence Erlbaum Associates.