A multimodal analysis of picture books for children: A Systemic Functional approach is part of a book series entitled Discussions in functional approaches to language designed to provide a forum for scholars to address theoretical issues of functional approaches to language.

The book’s range extends beyond the exploration of the creation of reality, interaction and textuality in language. Not only is there “an exploration of the roles that language and images play, both by themselves and through their interaction” (Jia & Zhang 2014: 342), but there is also an analysis of nine picture books intended for children at different stages of cognitive development to exemplify the three metafunctions in Systemic Functional Grammar (henceforth SFG) and their counterparts in Visual Social Semiotics.

There are nine chapters. After an introductory chapter and literature review, Chapters 3 to 8 make up the body of the book and can be grouped into two parts. The first part, Chapters 3 to 5, explains and exemplifies the three streams of meaning proposed in SFG and Visual Social Semiotics, respectively. The second part, Chapters 6 to 8, analyzes the nine picture books from the three theoretical perspectives in relation to age. Chapter 9 offers a conclusion.

The introductory chapter lays out the aims, scope and structure of the book and offers a detailed description of the data of the study. It states that the book attempts to reveal how the verbal and visual modes in picture books for children create reality, interaction and textuality. By way of justification, the author claims to have identified certain gaps in the research by reviewing classic studies of picture books for children. For instance, “the intersemiosis between verbal and visual aspects has been neglected”; “the meaningful exegetical possibilities afforded by a purely linguistic analysis” have been disregarded; and the age of the targeted children has not been reckoned with (pp. 1–2). The aims of the book are threefold. Above all, it seeks to explore the choices available to the writers and illustrators of picture books for conveying the three streams of meaning through verbal and visual modes. Secondly, it aims to contribute to the classification of picture books.
based on the age group of the child reader. Thirdly, it attempts to contribute to the critical analysis of SFG and Visual Social Semiotics. After a brief outline of the book’s structure, the author compares “the topics, literary structure and format of the tales” of picture books intended for children of three stages of cognitive development: the sensory-motor, pre-operational and concrete operations stages (p. 8). The chapter ends with a description of the research data, which consist of nine picture books for children from 0 to 9 years old. For a better understanding of the data, the plots of each of them are expounded.

Chapter 2, entitled “Systemic Functional Grammar and Visual Social Semiotics”, gives a survey of relevant theories, laying a solid foundation for the construction of the theoretical frameworks of the ensuing three chapters. A stated aim is to justify why Halliday’s SFG and Kress and van Leeuwen’s Visual Social Semiotics are better adapted to the verbal and visual analysis of picture books than other functional and multimodal accounts. The author compares formal and functional approaches to language, arguing that formal approaches are mainly concerned with “formal aspects of a phonological, morphological, and syntactic nature”, while functional approaches are more interested in how language functions in human communication and interaction (p. 20). Functional approaches are classified into three broad groups: generative, moderate and extreme functionalism. It is worth noting that SFG, the chosen approach, falls within the moderate group. SFG is then discussed in relation to other functional accounts of language such as Functional Discourse Grammar and Role and Reference Grammar. The author argues that SFG is especially well-adapted to the study of language in picture books by virtue of its “attention to natural data and text structure”, its “commitment to socio-cultural aspects”, its orientation to metafunctions, and its “use of systems and system networks” (pp. 28–36). The choice of Kress & van Leeuwen’s (2006) Visual Social Semiotics as the model of analysis is justified with reference to the other two main approaches to multimodality, namely Multimodal Discourse Analysis (O’Halloran 2004, 2005; O’Toole 2011) and Multimodal Interaction Analysis (Jewitt 2009; Norris 2004). The main reason for adopting Visual Social Semiotics lies in its correspondence with SFG. Some of the basic assumptions of Visual Social Semiotics are then laid out in more detail: (i) communication tends to involve several modes, and therefore is multimodal in nature; (ii) modes are a joint product generated from the synergy of social, historical, and cultural factors; (iii) each mode has its own meaning potentials, as determined by the specific semiotic resources it carries; (iv) the meaning created by the intersemiosis of several modes is “beyond the information transmitted by each mode in isolation” (p. 52). Finally, the chapter discusses different interpretations of “mode” within some important multimodal approaches, arguing that modes for making meaning
comprise mainly “written language, spoken language, visuals, music, sound and gestures” (p. 58).

The ensuing three chapters present the theoretical frameworks fundamental to the study by examining the three streams of meaning in turn. Each is structured in the same way, starting with an overview of the theories under discussion, proceeding with a detailed explanation of how words and images work both individually and in combination to create meaning, and ending with an exemplification of relevant theories.

Chapter 3 looks in particular at ideational meaning from the perspective of SFG and representational meaning from that of Visual Social Semiotics. Ideational meaning is concerned with the representation of reality at the lexico-grammatical stratum and conveyed through choices in the transitivity system which “construes the world of experience into a manageable set of process types” (Halliday & Matthiessen 2004: 170). Having introduced ideational meaning in SFG, the chapter turns to its counterpart in the visual mode – representational meaning, which can be classified into narrative and conceptual patterns based on the presence/absence of vectors (Kress & van Leeuwen 2006). In the case of narrative representation, two types of process are identified – action and reaction, both of which have a transactional or non-transactional character. Conceptual representation is associated with analytical, classificational, and symbolic processes. It is worth noting that, although speech and mental processes are regarded as a third pattern of representation in visual grammar, they are not much discussed by the author, maybe because there are few examples in the data. The chapter then focuses on the interplay of words and images at the representational level with reference to research undertaken in this field, and the author expresses a preference for the three types of inter-modal relationships (ideational concurrence, ideational complementarity and connection) proposed by Unsworth (2006). Then, the author discusses some problems that may arise when analyzing picture books, such as how to identify process types that are controversial, how to count the number of processes, and how to categorize transactional reaction process. The chapter ends with an analysis of how the verbal and visual modes in the picture book The very hungry caterpillar contribute to representing reality. By identifying the verbal choices made by the writer, it is shown that material and relational processes are best suited to construe the world of experience, and processes in language are in most cases intransitive. By studying the visual choices made by the illustrator, it is demonstrated that there is a predominance of action images and a preference for transactional processes. By analyzing the interplay of words and images, it can be seen that ideational concurrence and complementarity take a dominant role.

Chapter 4, which tackles interpersonal metafunction and interactive meanings, falls into two parts. The first explains interpersonal and interactive meanings,
through which readers can gain a clear idea of their frameworks. Interpersonal meaning is argued to relate not only to interaction, but also to evaluation. Interaction is interpreted at two strata: at the lexico-grammatical stratum, it is mainly realized by the mood structure (indicative or imperative); at the semantic stratum, it is expressed through speech acts (statement, offer, question and directive). Evaluation comprises the systems of polarity and modality. Since language is not the only mode to create interaction and evaluation, much space is also devoted to elaborating the interactive meaning conveyed by images. Interaction in images, according to Kress & van Leeuwen (2006), is established between three types of participants, namely image producers, image readers and the represented participants. Four types of system related to interactive meaning are distinguished, namely image act and gaze, social distance, horizontal angle, and vertical angle. In addition to the interaction created by images, the author discusses their evaluative aspect, or rather modality, which is not merely concerned with its resemblance to reality, but also associated with the cultural standards of what is real or unreal within a certain social group (Kress & van Leeuwen 2006). In order to measure the degree of modality, four types of coding orientation are differentiated – naturalistic, scientific, abstract and sensory – all of which can be described by modality markers like color (saturation, modulation and differentiation), background, detail, depth, illumination and brightness. The second part of the chapter exemplifies interpersonal and interactive meanings in the picture book Where the wild things are. An in-depth analysis of the book is preceded by a discussion of two major problems. One concerns the identification and counting of mood structures in language; the other is associated with the application of modality markers and genre-related issues in images. Having resolved these controversial issues, the author identifies the verbal and visual choices made by the writer and the illustrator of the picture book, showing that choices of mood in language are mainly restricted to the declarative type, which indicates that the verbal texts are more likely to give information than to encourage interaction. Regarding images, there is a preference for offers over commands, middle-shots over close-ups and long shots, oblique over frontal angles, and eye-level over high and low angles, all of which contribute to shortening the distance between the child reader and the characters in the tales. The intersemiosis of words and images suggests an attitudinal dissonance and demonstrates no reinforcement of address. Finally, there is a brief review and explanation of the verbal and visual choices made in the selected picture book at interpersonal/interactive level.

Chapter 5 addresses the textual and compositional aspects of SFG and Visual Social Semiotics. First, in the discussion of the textual metafunction, considerable importance is attached to different types of theme (unmarked and marked, simple and multiple) and thematic progression pattern (constant, linear and derived).
The visual counterpart of the textual metafunction, the compositional meaning, also attracts due attention. When introducing visual composition, three principles are elaborated in detail, namely information value, framing, and salience. Having explained textual and compositional meanings, the chapter exemplifies them by examining the picture book *Gorilla*. First, however, two major aspects that may give rise to misinterpretations are commented on, i.e. the identification of theme in complex clauses, embedded clauses, and ellipsed clauses; and the information value of left and right in visual grammar. The chapter then conducts an analysis of *Gorilla*, covering textual analysis, compositional analysis, and the analysis of the interplay of images and words at textual level. It is shown that simple themes greatly outnumber multiple themes, and unmarked themes greatly outnumber marked ones. The constant pattern of thematic progression plays a dominant role, suggesting that the reiteration of given information makes it easier for the child reader to follow the plots. The analysis of visual composition demonstrates that the gorilla and her father are often placed at the center of the illustrations to highlight their dominant roles in the tale. In addition to information value, framing and salience are also explored. It is shown that frames are used in the majority of illustrations to suggest an initial conflict between the represented participants and the young child, while there is no frame or dividing line in the last illustration, which symbolizes the resolution of the conflict and the construction of balance. In terms of salience, three characters – Hannah, her father and the gorilla – are given more prominence than other visual elements. The analysis of *Gorilla* ends with a discussion of the interplay of images and words at textual level, concluding that the verbal and visual features of the picture book complement each other in expressing contrastive meanings of companionship and isolation.

As an application of the theoretical frameworks introduced in the previous three chapters, Chapters 6 to 8 explore how the verbal and visual choices made by the writers and illustrators of the nine picture books represent reality, generate interaction, and create textuality.

Chapter 6 analyzes the verbal and visual elements in the selected picture books from a representational perspective. It starts by studying how the narrative world is represented by the verbal mode based on the analysis of process types, their transitive/intransitive property, and circumstances. The analysis suggests that material processes are very typical of picture books; this is ascribed to the fact that they enable writers to describe a series of actions of the main characters involved in the tale. Meanwhile, there is not much quantitative difference between transitive and intransitive relationships, both of which are crucial to the representation of reality. Furthermore, an analysis of circumstance finds that expressions of place and time occur rarely in picture books intended for younger readers, while the use of circumstantial elements increases considerably both in quantity and variety.
in picture books for older children. The chapter then explores to what extent “the circumstantial participant-process configuration” is reflected in the visual mode (p. 152). As the analysis of narrative processes suggests, embedded images in which action and reaction processes combine to convey the representation of reality account for the largest proportion, and transactional processes are favored over non-transactional ones in establishing personal relationships. Compared with narrative processes, conceptual ones occur far less frequently in picture books. Turning to the settings of images, their contribution to the representation of location is greater than that of circumstances in language. Having made clear how verbal and visual elements contribute to the representation of reality, the chapter moves to an account of the interaction between words and images with reference to the intersemiosis relationships introduced in Chapter 3. The author hypothesizes that images in picture books intended for younger children will tend to be an echo of language, while those in picture books targeted at older children are no longer subservient to language. However, this is found not to be so; no matter which age group of children a picture book is written for, its verbal and visual elements are always essentially complementary. The chapter concludes with a discussion of how age factors exert influence upon the use of verbal and visual strategies to create the representation of reality. The discussion mainly reveals that verbally, material processes run through picture books intended for children at three different stages of cognitive development, while some differences concerning relational and mental processes are identified in relation to the age of the target group; visually, embedded images involving action and reaction processes rank first in number in all nine books independently of the age factor.

Chapter 7 explores the verbal and visual strategies that are employed to generate interaction among different characters in the tales, and that between characters and the intended readers. A thorough investigation of the interaction created through language by identifying the mood structures reveals that the mood choices in the selected picture books are mainly restricted to declarative structures. The predominance of declarative mood is shown to be partly because they offer rather than demand information, and partly because the use of imperative or interrogative mood tends to interrupt the continuity of the plots of the tales. In addition, modal adjuncts are found in very limited number; modality is typically expressed by modal auxiliaries. The chapter then moves to an exploration of engagement generated by visual elements, showing that, as is the case with language, offers outnumber demands due to the lack of eye contact between the represented participants and the child reader. Moreover, the analysis of social distance suggests a predominance of middle-distance shots over close-ups and long distance shots, with the result that the characters are “depicted within the reach of the viewers’ gaze” (p. 205). Furthermore, it is found that the majority of horizontal angles are oblique,
and vertical angles are in most cases eye-level ones through which the child reader may identify with the characters illustrated. As for modality, it tends to “combine indicators of high and low modality” in selected picture books (p. 199). The chapter also includes a reflection on how the verbal and visual modes work together to generate interaction by referring to the relevant theoretical premises discussed in Chapter 4. Generally speaking, the visual mode contributes more than the verbal mode to the creation of interaction in picture books. For instance, the lack of engagement in the verbal mode, which is mainly caused by the wide use of declarative mood structures, is compensated for to some degree by visual elements. Finally the author discusses how the age factor plays a role in determining the verbal and visual choices that are used to convey interpersonal meaning. In the verbal mode, the choice of declarative mood is not so influenced by the age factor as that of interrogative and imperative moods, but there are more expressions of modality as the age of the target reader increases. In the visual mode, the predominance of offers over commands and oblique angles over frontal ones have little to do with the age factor, but significant differences in social distance, vertical angle and modality are identified in tales for different age groups. The degree of modality, for instance, is higher “the greater the age of the children for whom the tales are intended” (p. 214). With regard to the relationship between the intersemiosis of images and words and the age factor, the major finding is that the attitudinal assonance constructed by the verbal and visual modes in picture books intended for younger children turns into attitudinal dissonance in picture books for older children.

Chapter 8 studies the contribution of words and images to forming coherent texts. First, it focuses on the organization of information in the verbal mode. An overview of thematic choices in the data confirms that simple themes are more notable than multiple themes, and unmarked ones than marked ones, so as to facilitate “the young child’s understanding of the plot” (p. 219). With respect to thematic progression, the constant pattern occurs with such a high frequency that it is easy for the child reader to grasp the thread of the plots. Moreover, it is very common for theme and topic to overlap in picture books, which helps to activate “the entities about which information is given” from “the beginning of the clause” (p. 220). Next, compositional features are explored in order to see how images contribute to creating textuality. It is shown that the given-new distribution of information and the placement of elements in the foreground are more likely to be adopted in picture books for younger readers, and that frames tend to occur more frequently in tales intended for older children. Then the chapter looks at how the verbal and visual modes work together to create the overall coherence of the books, concluding that these two modes complement each other so as to make the tales more readable for children. Finally, there is a discussion of the extent to which the age factor determines thematic organization and visual composition.
The analysis of the verbal choices in relation to the age factor demonstrates that picture books for children at the three stages of cognitive development all favor simple and unmarked themes. However, the influence of the age factor is shown by the fact that theme/topic correspondence is more evident in tales created for children at the sensory-motor and pre-operational stages, and the constant thematic progression pattern occurs with a lower frequency in tales for children at the sensory-motor stage than in those for older children. Turning to the relationship between compositional organization and the age factor, it is found that the given-new distribution of information is more common in tales intended for the youngest children than in those for older children, that the use of framing increases as the children’s age increases, and that the elements in the background are given more prominence and detail in tales for older children.

Chapter 9 summarizes the major findings of the study. Each of its three sections reviews the verbal and visual features of picture books intended for children at a specific stage of cognitive development. In reexamining the verbal and visual modes in picture books for children, the author points out that on the one hand, they have something in common, such as the preference for material processes, declarative mood, offers, and oblique angles; on the other hand, they are different in terms of framing, information distribution and so on. In this sense, the chapter is more like a comparison of words and images in picture books intended for children at three different stages of cognitive development than a mere summary.

To sum up, A multimodal analysis of picture books for children: A Systemic Functional approach is successful in that it not only identifies the verbal and visual choices available to the writers and illustrators of the picture books under discussion, but also attaches great importance to the influence the age factor exerts upon the representation of reality, the generation of interaction, and the creation of textuality. The book’s value is twofold: theoretically, it takes “the most relevant systemic-functional and visual social semiotic theories a step further from previous studies” by giving due prominence to the intersemiosis of words and images; analytically, it systematically applies its theoretical fruits to the multimodal analysis of picture books – a genre of children’s tales – in relation to the age factor (p. 1). However, there are some flaws. One of them is that the research is based on only nine picture books, a mere three per age group, which is an inadequate basis for the broad conclusions it draws. Another is that Chapters 6 to 8, where the main analysis is to be found, are so densely structured that they are sometimes difficult to follow. Nevertheless, the book undoubtedly succeeds in offering a comprehensive theoretical framework for those engaged in the study of both language and images. It is not only a must-read for students, lecturers and researchers who are working in the field of multimodal discourse analysis, but also highly recommended for those interested in picture books.
Acknowledgements

I am grateful to Dr. Mark Buck and Prof. Hilde Hasselgård for helpful comments on earlier versions of this book review. I also wish to express my gratitude for support from the Fundamental Research Funds for the Central Universities (2017QW03).

References