

Article

Interpersonal Meaning: Verbal Text–Image Relations in Multimodal Science Texts for Young Children

Maria Koutsikou ¹, Vasilias Christidou ^{2,*}, Maria Papadopoulou ² and Fotini Bonoti ¹

¹ Department of Early Childhood Education, University of Thessaly, 382 21 Volos, Greece; mkoutsikou@uth.gr (M.K.); fbonoti@uth.gr (F.B.)

² School of Early Childhood Education, Faculty of Education, Aristotle University of Thessaloniki, 541 24 Thessaloniki, Greece; mpapadopoulou@nured.auth.gr

* Correspondence: vchristidou@nured.auth.gr

Abstract: Verbal text and images constitute the principal semiotic modes interacting to produce interpersonal meanings in multimodal science texts for young children. These meanings relate to pedagogical perceptions about children’s learning. This study examined verbal text–image relations regarding the interpersonal meaning dimensions of address (the way the reader is addressed), social distance (the kind of the relationship between the reader and represented participants), and involvement (the extent to which the reader is engaged with what is represented) in multimodal text excerpts from science-related books for preschool children. The sample consisted of 300 randomly selected units of analysis. For each unit, the verbal and the visual content was analyzed along each dimension, and the relevant verbal text–image relation was determined. Results indicated that regarding address and involvement, relations of convergence appeared significantly more frequently than relations of complementarity and divergence. Concerning social distance, relations of complementarity and divergence were observed more frequently than relations of convergence. Results are discussed in the context of the Systemic Functional Grammar and the Grammar of Visual Design, in the light of the socio-cognitive perspective on science teaching and learning. Implications for the selection, design, and use of multimodal science texts for young children are also discussed.

Keywords: address; interpersonal meaning; involvement; multimodal texts; science; social distance; socio-cognitive perspective; verbal text–image relations; young children



Citation: Koutsikou, M.; Christidou, V.; Papadopoulou, M.; Bonoti, F. Interpersonal Meaning: Verbal Text–Image Relations in Multimodal Science Texts for Young Children. *Educ. Sci.* **2021**, *11*, 245. <https://doi.org/10.3390/educsci11050245>

Academic Editors:
Konstantinos Ravanis and
James Albright

Received: 6 April 2021
Accepted: 14 May 2021
Published: 19 May 2021

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

‘Multimodality’ has been widely used to denote the contribution of different semiotic modes in the construction of meaning [1]. Teaching and learning in science are considered as multimodal processes, where a multitude of semiotic modes (e.g., language, visual representations, gestures, and body language) contribute to the presentation and communication of scientific meanings, while inter-semiotic interaction produces new meanings during instruction [2–10]. This multiplicity and complexity of meanings produced by the interplay of various modes suggests that in the context of a ‘pedagogy of Multiliteracies’ [11] students, from an early age [12,13], need to be able to analyze, interpret, critically understand, and use different representational modes of meaning-making, as well as the ways these modes interact [11,14–16].

Teaching materials play a key role in science learning [17,18]. More particularly, during early childhood, children’s interaction with science-related texts is an integral part of their first learning experiences. Among different kinds of such materials, informational books are widely used by preschool teachers as instructional aides to introduce science concepts and phenomena in the classroom [19,20]. Science teaching material for young children is vastly multimodal, i.e., it involves the synergy of different modes of representation [21,22], with verbal text and image being prevalent [23].

Analyses of multimodal teaching materials often draw on two fundamental theoretical schemes: The Systemic Functional Grammar and the Grammar of Visual Design. The Systemic Functional Grammar [24] is a grammatical model that belongs in a broad social semiotic approach to language (i.e., Systemic Functional Linguistics). According to this model, language is a system, namely a set of available choices for making meaning, and language functions (or “metafunctions”) relate to different levels of meanings [24]. The Grammar of Visual Design [25] is a social semiotic approach to visual communication and describes the visual resources available for constructing meaning through images. It was based on the semiotic principles of Systemic Functional Grammar and supports that the metafunctions of language apply also to images [25].

Therefore, according to the Systemic Functional Grammar [24] and the Grammar of Visual Design [25] in every text the verbal and the visual mode can produce meaning in three different levels: the representational (actions, participants involved in them, and overall context), the compositional (composition of verbal and visual elements on the page) and interpersonal (relations between writer and reader, and reader–represented knowledge).

The interpersonal meaning is particularly significant pedagogically, since it relates to a crucial interaction taking place in educational environments, namely the interaction between the learner and the teaching material [26]. More specifically, the selections made in the construction of teaching material regarding interpersonal meaning shape the role assigned to the reader and how s/he is positioned vis-à-vis the presented knowledge [3,27–29]. The interpersonal meaning promoted by an educational material delineates the terms of the child’s interaction with it and her/his knowledge construction, while it reflects the pedagogical positions that—deliberately or not—underly its design. Therefore, the semiotic selections regarding the interpersonal meaning disclose specific views about the child’s role in the learning process as well as the nature of knowledge and its construction, consequently affecting the quality and effectiveness of the learning experience resulting from the use of the material [30–35].

Furthermore, interpersonal meaning becomes even more important in the light of the socio-cognitive perspective on science learning, according to which science teaching materials are mediating tools and the young readers’ interaction with them determines the learning process [36,37]. The interpersonal meaning is particularly significant for preschool education, during which children acquire their first experiences with science topics, which are largely based on interactions with multimodal materials [20].

The aim of the present study is to investigate verbal text–image relations regarding the interpersonal meaning in excerpts from multimodal science-related books for preschool children (2.5–6 years old). The study adopts a socio-cognitive perspective [36,37] on teaching and learning science. More particularly, it draws upon key assumptions of the socio-cognitive model for designing science teaching material. In the following sections these assumptions and their association with the dimensions of interpersonal meaning will be presented, followed by a presentation of the interpersonal meaning dimensions and the relations between verbal text and image in regards to interpersonal meaning. Then, the rationale, research question and hypotheses of the study are exposed.

2. Interpersonal Meaning and Key Socio-Cognitive Assumptions

The interpersonal meaning is a particularly significant aspect of science teaching and learning in the socio-cognitive framework, according to which knowledge is constructed and transformed by the child in the context of social and physical interaction with others (e.g., peers, teachers) and culturally mediated tools and artifacts [37–40]. Science learning occurs through children’s interactions with other members of the school community and with teaching material [17,36,41–43]. These interactions determine children’s construction of scientific knowledge [44]. In the context of the socio-cognitive perspective, then, the terms of communication, the kinds and forms of the aforementioned interactions become of critical importance for the learning process [45]. Moreover, particular emphasis is given to the role of teaching materials as mediating tools and to the way they present knowledge.

Within this framework, special attention is paid to the function of teaching materials in shaping the role of the learner and her/his relation with the represented knowledge. Therefore, an analysis of teaching materials' interpersonal characteristics can shed light on the pedagogical assumptions underlying their design [30,34,40,46].

According to the socio-cognitive framework [36,37], children do not acquire knowledge passively. Instead, they are considered as active agents who construct knowledge in the context of social interactions and teaching mediation, based on their previous science-related experiences and conceptions [36,37,47–49]. Under the socio-cognitive lens, knowledge is not seen as objective or independent of the learner, but closely related to the child's life, while knowledge construction, i.e., learning, presupposes her/his action, participation, and active engagement [39,41,48,50].

Therefore, the design of science teaching material which aligns with the socio-cognitive perspective (i) addresses young readers by assigning them an active role in their own learning, (ii) promotes the development of an intimate social relationship between the reader and the represented knowledge, and (iii) engages the reader with what is presented [32,33,51–53]. In the context of the present study, the degree to which these characteristics are reflected in the verbal and visual semiotic selections of a text refers to *address*, *social distance*, and *involvement* respectively, namely three fundamental dimensions of the interpersonal meaning [24,25]. These dimensions are analyzed in the following paragraphs.

3. Interpersonal Meaning Dimensions in Verbal Text and Image

As already mentioned, interpersonal meaning is verbally and visually constructed and promoted in multimodal science teaching material by means of three dimensions, namely (a) *address*, (b) *social distance*, and (c) *involvement* (Figure 1). Power relations between reader and represented participants constituting the fourth dimension of interpersonal meaning and realized verbally by the use of evaluative language and visually by the vertical angle of the image [25], was not examined in the present study. This selection derives from the nature of the current study and the particular characteristics of the material. Specifically, informational science books for preschool children typically involve a very short verbal text consisting of simple clauses and evaluative words or vertical angles other than eye-level shots are mostly absent. More particularly,

- *Address* refers to the way a text addresses the reader. Address is verbally realized by the type of clause used (i.e., imperative, for example “Show the planets’ motion around the sun”, interrogative, e.g., “Have you ever seen a lightning?” or declarative, e.g., “We have five senses: We can see, hear, smell, taste and touch”) and the person of the verb in a clause [24]. Visually, address is articulated by means of the presence or absence of the represented participants’ gaze towards the reader [25,54]. The term “represented participants” refers to the verbally and visually represented living entities, people, or animals, participating in the actions presented in the verbal text and images [25].
- *Social distance* reflects the kind of social relationship between the reader and the represented participants promoted by a text. Verbally, social distance is expressed through the voice of the verb (active or passive) and the type of relationship between clauses (parataxis/absence of subordinate clauses or hypotaxis) [24]. In the Greek language there are four choices regarding the voice of the verb: active, i.e., the subject of the verb performs an action (e.g., “The sun heats the earth”), passive, i.e., the subject of the verb receives an action (e.g., “The earth is heated by the sun”), middle, i.e., the subject of the verb both performs and receives an action, or neutral, i.e., the subject of the verb neither performs nor receives an action, but is in a state (e.g., “In winter, brown bears hibernate”). Regarding the type of relationship between clauses, the choices are parataxis (e.g., “In autumn the leaves of some plants change color and fall”), absence of subordinate clauses (e.g., “During hibernation bears do not need to eat”), or hypotaxis (e.g., “When water droplets get too heavy in the cloud, they fall to earth as rain”). In images, social distance is realized by the size of frame, referring to

whether the participant's body is depicted in full, or partly. In particular, size of frame relates to three visual choice alternatives: close, medium, or long shot [25].

- *Involvement* is associated with the degree to which the reader is invited to engage with what is represented. Involvement is verbally articulated through the person of the possessive pronouns (first, second or third) [24] and visually expressed by means of the horizontal angle of the image (frontal or oblique) [25].

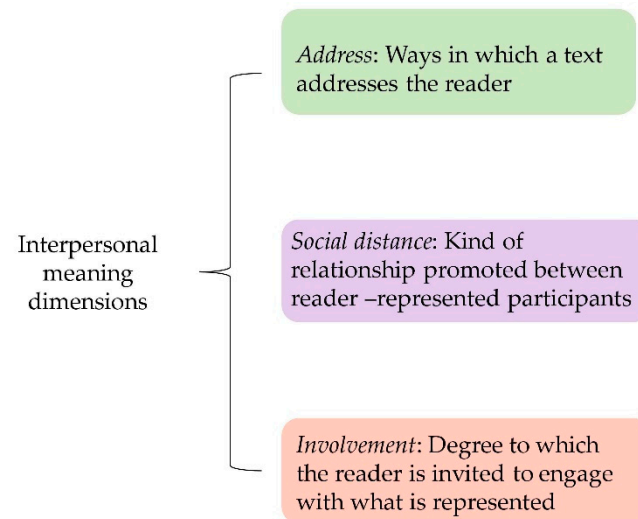


Figure 1. Dimensions of interpersonal meaning [24,25].

4. Relation Between Verbal Text and Image Regarding Interpersonal Meaning

The overall interpersonal meaning conveyed by a multimodal text is not formed by the distinct contribution of verbal text and image, but from their interrelation [55–57]. For each dimension of interpersonal meaning (address, social distance, and involvement), the verbal and the visual mode in a text can be characterized as [58]

1. *Convergent*, when they produce similar meanings (e.g., when both verbal text and image denote low address, i.e., simply present information to the reader instead of calling her/him to undertake some action);
2. *Complementary*, if one of the semiotic modes contributes additional meanings to those promoted by the other (for instance, verbal text denoting familiarity with the reader, i.e., signifying small social distance combined with an image signifying moderate social distance between the reader and the represented entities); or
3. *Divergent*, when the two modes promote contradictory meanings (e.g., when the verbal text promotes strong reader involvement with what is represented, while the image discourages reader involvement).

Exploring the relations between verbal text and image in science-related multimodal texts for children is important for science education, since it has been suggested that young readers have a difficulty in appropriately associating verbally and visually conveyed information during their early encounters with such texts, and this affects their understanding of the presented knowledge [59–65]. Previous studies [66–69] have indicated the significance of prompting similar meanings through verbal text and image in multimodal texts addressing children, since this similarity of meanings facilitates understanding. Meaning convergence between the two modes facilitates the cognitive process of creating associations between verbal and visual representations, which is a requirement for comprehending multimodal texts [70]. The convergence relation is considered to be the optimal for enhancing children's understanding of various types of text, while more complex relations between the verbal and the visual mode, like complementarity, may cause difficulties to young readers [67,69]. In addition, divergence relations between word and image tend to

pose further challenges to children and lead to confusion and misinterpretations of the represented knowledge [71].

It is suggested that the interpersonal meanings promoted by verbal text and image in multimodal science teaching materials should align with the socio-cognitive principles on science learning. In particular, science teaching materials are expected to assign an active role to the child in the process of knowledge construction, to promote an intimate relation between the child and the represented knowledge, and to encourage the child to engage with it [32,33,51–53]. Therefore, science teaching materials designed according to the socio-cognitive perspective on learning, promote both verbally and visually (see Figure 2):

- High address, encouraging the reader to perform some action regarding the represented knowledge, therefore acknowledging the readers' active role in their own learning;
- Small social distance between the reader and what is represented, thus implying intimacy and familiarity;
- Strong involvement of young readers, by presenting knowledge as belonging in their personal world and experiences and inviting them to engage with it.

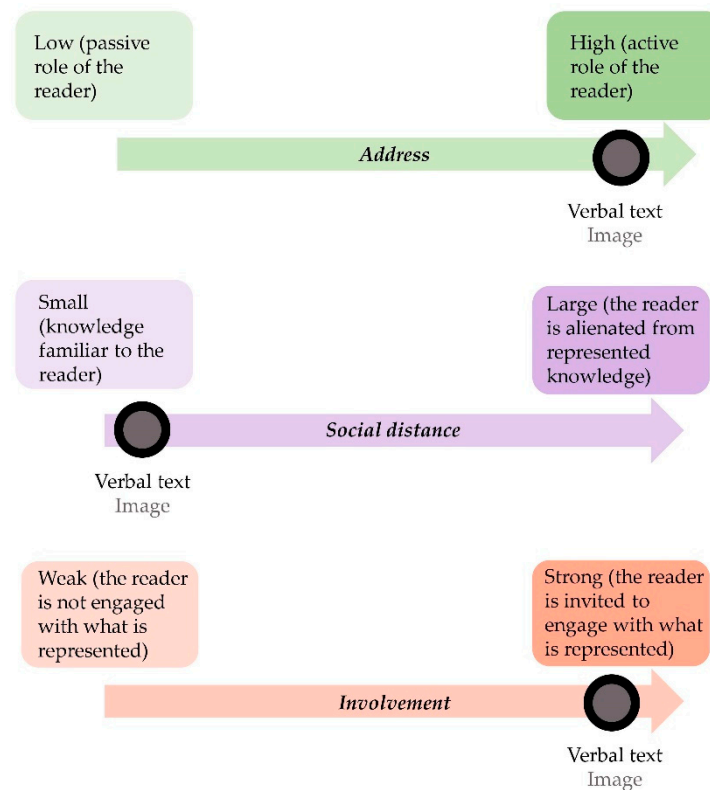


Figure 2. Verbal and visual semiotic selections and their correspondence to key socio-cognitive assumptions.

5. The Present Study

As already mentioned, the interpersonal meaning promoted by teaching materials is significant for science education, since it reflects the pedagogical views about the learner's role in the construction of knowledge, thus affecting the quality and effectiveness of the learning experience supported by the use of these materials. This meaning is particularly important for the preschool years when children develop their first systematic science learning experiences, which extensively rely on their interaction with multimodal texts. This kind of materials largely involves informational books covering topics from biology and physics [20,72]. Since meaning in these multimodal books is not constructed through the distinct contribution of each mode, but through their synergy and interaction [11,22,73–76], understanding how these two modes interrelate is a precondition for

their comprehension [77,78]. Several studies [79–85] have focused on the interpersonal meaning in children’s picture books. It is showed that verbal text and images promote low address, serving the presentation of information to the young readers in such books [79–81]. Furthermore, some studies indicated that images in picture books promote an intimate relationship between readers and represented participants [82,84] and invite readers to engage with what is presented [82], while other studies found that a remote relationship is promoted [79,80,85] and the reader is not engaged [79] visually. However, research on the dimensions of interpersonal meaning promoted by verbal text and/or images in science materials [27,29,33,51,52,86] and more particularly those addressing young children [53] is limited.

Exploring the relation between the verbal and the visual mode in science teaching materials is important, especially in the case of those addressed to young children. It has been suggested that children find it difficult to appropriately coordinate verbal text and image in such multimodal texts, which affects their comprehension of the presented knowledge [62,64,68,87]. In fact, it has been indicated that different kinds of relation between verbal text and image affect children’s comprehension in varying degrees [67,69]. However, research on verbal text–image relations in multimodal science texts for children has mainly explored the representational or the compositional meaning [88–93]. Furthermore, previous research on the interpersonal meaning in multimodal science texts has examined the distinct contribution of the verbal and the visual mode in meaning construction instead of their interrelation [27,29,33,51,52,86]. Studies on verbal text–image relations in terms of the interpersonal meaning in science teaching material for preschool children are practically absent from the international literature, despite the pedagogical impact of this meaning on young children’s science learning.

In an attempt to fill this gap, the present study aims at addressing the following research question: What are the relations between verbal text and image promoted by multimodal texts about science in related books for preschool children regarding the interpersonal meaning dimensions of address, social distance, and involvement?

Adopting the socio-cognitive perspective on science learning [36,37,39,41,47,48], it was expected that the material under study, both verbally and visually, would ascribe children with an active role as readers/learners, by instilling an intimate relation between them and the presented knowledge, and by encouraging their engagement with what is represented. More particularly, based on our hypotheses outlined below, it was expected that the verbal and the visual mode in excerpts from multimodal science books for young children would promote:

Hypothesis 1 (H1). *Convergent interpersonal meanings instead of complementary or divergent ones.*

Hypothesis 2 (H2). *High address.*

Hypothesis 3 (H3). *Small social distance between the child and what is presented.*

Hypothesis 4 (H4). *Strong reader involvement with the presented knowledge.*

It is noted that in the current study the verbal and the visual choices related to interpersonal meaning are analyzed and discussed according to the Systemic Functional Grammar [24] and the Grammar of Visual Design [25] and examined through the lens of the socio-cognitive perspective. Therefore, the current study neither examines the book creators’ intention regarding the semiotic choices analyzed, nor considers the semiotic choices as a result of the intention of the books’ creators.

6. Materials and Methods

6.1. Sample

Since there are no valid and official records of the constantly expanding publishing production, the creation of a complete list of all population units from which the sample

could be randomly selected was not feasible. For this reason, the following sampling procedure was followed. Firstly, in order to determine the appropriate material for the aims of the study, books related to science for preschool children available in the Greek market and written or translated in Greek were systematically searched on the Internet using book-selling websites. Search results were limited to those meeting the following criteria:

- Informational books, namely books aiming to familiarize young children with concepts, phenomena, processes etc. [94]. Science-related books addressing preschool children but belonging in other genres (such as activity books, fictional books) were excluded for consistency, because of the different purposes, content type, and writing style in these book categories.
- Books published within the last decade from the sampling date (2018). This criterion was applied in order for the sample to include recent books, therefore available for purchasing, but also revealing the current trends and probably echoing contemporary pedagogical perceptions on learning, and more particularly the socio-cognitive perspective.

This process has allowed the construction of a list of Greek publishing companies which publish such books. A written request was sent to all these publishers to supply a copy of their books that met the abovementioned criteria for our research purposes. Seven Greek publishing companies responded to this request, offering a total of 55 books, from which 30 fully complied with the criteria. According to the descriptions on their publishers' catalogues, these specific books are aimed at preschool children, aged 2.5–6 years [95,96]. These 30 books were divided into units of analysis. Each unit of analysis consisted of an image and the verbal text accompanying it. Since the estimation of address for the visual mode is based on the direction of the participant's gaze (towards the reader or elsewhere) [25], units that did not include images of living entities as participants (i.e., people, animals, or anthropomorphic entities illustrated with eyes) were excluded from the sampling process. Through this process, the 30 books were divided into a total of 670 units. A list comprising all these units was created. A sequential number was assigned to each unit in the list. Using simple random sampling, a total of 300 units of analysis were selected, which formed the sample of this study. In specific, using the random number generator of SPSS statistics program, 300 random numbers were selected from the 670 numbers of the list. The units of analysis with sequential numbers corresponding to these randomly selected numbers were included in the sample. The sample size of 300 units of analysis was estimated as adequate considering the number of interpersonal meaning dimensions examined in the study, the number of levels of each dimension and the number of verbal text–image categories.

6.2. The Framework of Analysis

For the purposes of the present study a framework of analysis was constructed, including the realization means of each dimension of interpersonal meaning, the different levels of each dimension, and the specific choices indicating each level of dimension for the verbal and the visual mode. This framework was based on the Systemic Functional Grammar [24] and the Grammar of Visual Design [25] regarding the verbal text and images, respectively. Furthermore, regarding verbal text–image relations the framework was based on the three basic categories of relations: convergence, complementarity, and divergence [58]. For each dimension of interpersonal meaning (i.e., address, social distance, and involvement) and for each of the two semiotic modes (SM), the framework of analysis included: the realization means, the levels of each dimension, and the rules for classifying the content into each level. Furthermore, it included the categories of verbal text–image relation and the classification rules for each category. For the construction of the analysis framework, the criteria of exhaustiveness and mutual exclusiveness were taken into account, so that the content of each unit of analysis could be classified into one of the levels of the dimension and could not fall into more than one levels, respectively. The framework of analysis is presented in Figure 3 and is described below in detail.

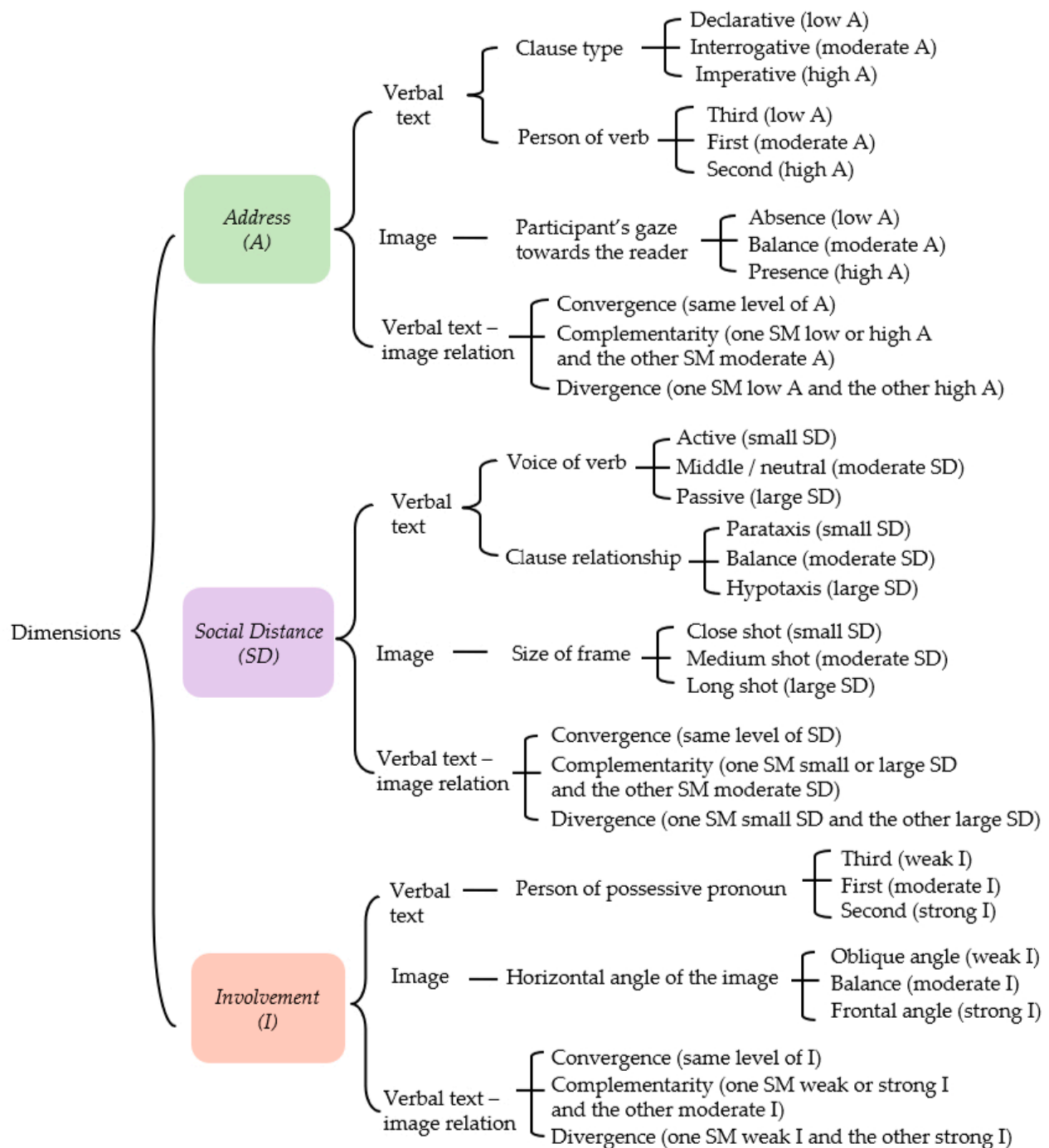


Figure 3. The framework of analysis.

6.3. Analysis Procedure

For each unit of analysis, the verbal and the visual content was classified with regard to each dimension of interpersonal meaning (i.e., address, social distance, and involvement) at one of the levels of each dimension according to the classification rules and then the category of verbal text–image relation was defined. Consistency between two coders, who coded independently 30 units of the total sample, was assessed using Cohen's Kappa which indicated high ($\kappa > 0.80$) interrater reliability [97,98] for the dimension of address, social distance, and involvement ($\kappa = 0.88$, $\kappa = 0.84$ and $\kappa = 0.89$, respectively). Inconsistencies between the two coders were resolved by a third coder as proposed by several researchers [97–99].

6.3.1. Classification of Verbal and Visual Content in Relation to Address

Regarding the verbal text, address was assessed based on the combination of clause type and the person of the verb. For example, address was classified as “low” when there was a combination of a declarative clause and third person (e.g., “Giraffes live in the savannas of Africa” [100] p. 8), as “moderate” when an interrogative clause and a verb in the first person were combined (e.g., “Do I have the same shape as my shadow?”), and as “high” when an imperative clause and second person were combined (e.g., “Avoid looking directly at the sun!” [101] p. 11).

Visually, address was determined by assessing the depicted participants’ gaze. The visual content was classified as “low address” when the participants did not gaze towards the reader and as “high address” when the participants gazed towards the reader. The content was categorized as “moderate address” when there was an equal number of participants gazing and not gazing towards the reader, signifying balance between high and low address.

Classification rules for the verbal (clause type, person of verb and their specific combinations) and visual (participants’ gaze) realization means for address are summarized in Table 1.

Table 1. Levels of address, realization means, and classification rules.

Address	Verbal Text			Image
	Clause Type	Person of Verb	Combination	Participants’ Gaze
Low (L)	Declarative	Third	L+L or L+M	Participant(s) not gazing towards the reader
Moderate (M)	Interrogative	First	M+M or H+L	Equal number of participants gazing and not gazing towards the reader
High (H)	Imperative	Second	H+H or H+M	Participant(s) gazing towards the reader

For reasons of exhaustiveness, when a unit of analysis involved different types of clauses, address was classified as “low” when there was an equal number of declarative and interrogative clauses, as “moderate” when there was an equal number of declarative and imperative clauses, and as “high” when there was equal number of imperative and interrogative clauses. Similarly, concerning the person of the verb, the content was classified as “low address” when there was equal number of verbs in the first and the third person, as “moderate” when there was an equal number of verbs in the third and the second person, and as “high” when there was an equal number of verbs in the first and the second person.

6.3.2. Classification of Verbal and Visual Content in Relation to Social Distance

Regarding the verbal text, social distance was estimated based on the combination of the voice of the verb and the type of relationship between clauses. For example, social distance was categorized as “small” when active voice of the verbs and parataxis were combined (e.g., “Squirrels climb tree trunks and jump from tree to tree”), as “moderate” when there was a combination of middle/neutral voice of the verbs and balance between parataxis and hypotaxis (e.g., “When brown bears hibernate, they stay in their dens and rest”), and as “large” when there was a combination of passive voice and hypotaxis (e.g., “If the earth was not heated by the sun, it would be frozen”).

Visually, social distance was determined by the size of frame and was estimated as “small” when participants were depicted in a close shot (e.g., when the head and the shoulders, or only the head of a participant was depicted). The visual content was categorized as “moderate social distance” when participants were depicted in a medium shot (i.e., when a participant’s body was ‘cut’ to the chest, waist, or knees, or the full body was depicted occupying more than 50% of the image space), and as “large” when participants were depicted in a long shot (the full body was depicted, occupying less than 50% of the image space).

Table 2 presents the classification rules of the verbal (voice of verb, relationship between clauses and their combinations) and the visual (size of frame) content.

Table 2. Levels of social distance, realization means, and classification rules.

Social Distance	Verbal Text			Image
	Voice of Verb	Clause Relationship	Combination	Size of Frame
Small (S)	Active	Parataxis or absence of dependent clauses	S+S or S+M	Close shot
Moderate (M)	Middle/neutral	Balance between parataxis or absence of dependent clauses and hypotaxis	M+M or S+L	Medium shot
Large (L)	Passive	Hypotaxis	L+L or L+M	Long shot

For reasons of exhaustiveness, when a unit of analysis included more than one verbs of different voices, social distance was classified as “small” when there was an equal number of verbs in active and middle/neutral voice, as “moderate” when there was an equal number of verbs in active and passive voice, and as “large” when there was an equal number of verbs in passive and middle/neutral voice.

6.3.3. Classification of Verbal and Visual Content in Relation to Involvement

Regarding verbal text, involvement was estimated by the person of the possessive pronouns used. Thus, the verbal content was classified as “weak involvement” when the pronoun used was in the third person (e.g., “her”/“his”, as in the following sentence: “A butterfly lays her eggs on a plant leaf” ([102] p. 31). The absence of possessive pronouns in the verbal text also indicated “weak involvement”. The content was classified as “moderate involvement” when the pronoun used was in the first person (e.g., “my/our”, as in the sentence: “The sun is at the center of our solar system” ([103] pp. 18–19) and as “strong involvement” for the use of a second-person possessive pronoun (e.g., “your”, as in the following sentence: “Your teeth cut the food into small pieces” [104] p. 20).

Visually, involvement was determined by the horizontal angle of the image. The content was categorized as “weak involvement” when the participants were illustrated in an oblique angle and as “strong involvement” when the participants were illustrated in a frontal angle. The visual content was categorized as “moderate involvement” when there was an equal number of participants in oblique and frontal angles, signifying balance between weak and strong involvement (Table 3).

Table 3. Levels of involvement, realization means, and classification rules.

Involvement	Verbal Text	Image
	Person of a Possessive Pronoun	Horizontal Angle of the Image
Weak (W)	Third person	Participant(s) depicted in an oblique angle
Moderate (M)	First person	Equal number of participants in an oblique and frontal angle
Strong (S)	Second person	Participant(s) depicted in a frontal angle

For reasons of exhaustiveness, when the verbal content in a unit of analysis involved more than one possessive pronouns, involvement was classified as “weak” when there was an equal number of possessive pronouns in the first and the third person, as “moderate” when there was an equal number of pronouns in the third and the second person, and as “strong” when there was an equal number of possessive pronouns in the first and the second person.

6.3.4. Classification of Verbal Text–Image Relation for Each Dimension of Interpersonal Meaning

The verbal text–image relation in a unit of analysis was categorized as “convergence” when the two semiotic modes had been classified at the same level of the dimension. Alternatively, when one of the modes had been classified at one of the extreme levels and the other had been classified at the moderate level (e.g., a unit of analysis involving a verbal text of high address and a visual image of moderate address) their relation was characterized as “complementarity”. Last, relations of “divergence” between the two semiotic modes appeared when the verbal and the visual part of a unit of analysis had

been classified at the two opposite extreme levels (e.g., a unit of analysis with the verbal text indicating small social distance and the image evoking large social distance).

For example, the verbal mode of the multimodal text in Figure 4 consists of the following clauses: “Wings help the airplane fly”, “Does the airplane move its wings?” [105]. The verbal text in this unit of analysis indicates “low address” since it combines a declarative (first) and an interrogative (second) clause, while the verbs used are in the third person. “Low address” is also promoted visually, since the depicted participant (child) does not gaze towards the reader. Therefore, verbal text and image are characterized by convergence regarding the dimension of address. Furthermore, the verbal text promotes small social distance, since the verbs used are in the active voice (“fly”, “move”) and there are no subordinate clauses. The image indicates moderate social distance because the child is depicted in a medium shot (the child’s body is “cut” to the chest). Therefore, in this unit of analysis verbal text and image are characterized by complementarity as far as social distance is concerned. Last, the verbal text promotes weak involvement since the third-person possessive pronoun (“its”) is used, while the image denotes strong involvement as the child is depicted in a frontal angle. Therefore, the two semiotic modes in this unit are characterized by the relation of divergence regarding involvement [105].



Figure 4. Example of a multimodal text indicating different verbal text–image relations regarding each dimension of interpersonal meaning. Reproduced with permission from R. Spiro, *Geia sou poulaki!* [105]; published by Psychogios Publications: Athens, Greece, 2018.

7. Results

Table 4 presents the observed frequencies of verbal text–image relations for each level of address, social distance, and involvement in the analyzed units. Therefore, regarding address, the majority of units of analysis are characterized by a relation of convergence, with the two semiotic modes mostly indicating low address towards the reader. Furthermore, concerning social distance, relations of complementarity, and divergence were the most common in the sample, with verbal text tending to promote small social distance between the reader and represented participants and images indicating moderate or large social distance. Moreover, as far as involvement is concerned, the analyzed multimodal texts tend to be characterized by convergence, with verbal text and image mostly indicating weak reader involvement. These dominant tendencies are highlighted in bold in Table 4.

Table 4. Interpersonal meaning dimension levels for verbal text and image per category of relationship.

Verbal Text –Image Relation	Address			Social Distance			Involvement		
	Verbal Text	Image	N	Verbal Text	Image	N	Verbal Text	Image	N
Convergence	Low	Low	120	Small	Small	18	Weak	Weak	98
	Moderate	Moderate	13	Moderate	Moderate	19	Moderate	Moderate	17
	High	High	18	Large	Large	13	Strong	Strong	35
	Total		151	Total		50	Total		150
Complementarity	Low	Moderate	23	Small	Moderate	79	Weak	Moderate	25
	Moderate	Low	21	Moderate	Small	16	Moderate	Weak	27
	High	Moderate	4	Large	Moderate	15	Strong	Moderate	8
	Moderate	High	11	Moderate	Large	27	Moderate	Strong	5
	Total		59	Total		137	Total		65
Divergence	Low	High	70	Small	Large	100	Weak	Strong	64
	High	Low	20	Large	Small	13	Strong	Weak	21
	Total		90	Total		113	Total		85

In order to answer the research question, statistical analysis using the chi-square goodness of fit test for each dimension of interpersonal meaning was performed. Results indicated that the three categories of verbal text–image relation were not equally distributed concerning the dimension of address ($\chi^2(2, N = 300) = 43.82, p < 0.001$), social distance ($\chi^2(2, N = 300) = 40.38, p < 0.001$) and involvement ($\chi^2(2, N = 300) = 39.50, p < 0.001$).

Further chi-square tests (with an alpha level adjusted to 0.017) were conducted for each dimension of interpersonal meaning to determine statistically significant differences between pairs of categories (i.e., between convergence and complementarity, between complementarity and divergence, and between convergence and divergence). Regarding address, results indicated that relations of convergence were observed more frequently than expected compared to complementarity ($\chi^2(1, N = 210) = 40.31, p < 0.001$) and divergence ($\chi^2(1, N = 241) = 15.44, p < 0.001$). Moreover, divergence was observed more frequently than expected compared to complementarity ($\chi^2(1, N = 149) = 6.45, p = 0.011$). These results support our first hypothesis (H1) that science-related multimodal texts would mostly involve convergent verbal and visual meanings regarding address. However, this convergence relation is not in the expected direction since verbal text and image tend to indicate low address (Table 4) and not high address as expected (H2). For example, as already indicated in Figure 4 above, verbal text and image are characterized by convergence [105], with the two semiotic modes indicating low address toward the reader.

Concerning social distance, results demonstrated that convergence appeared less frequently than expected compared to complementarity ($\chi^2(1, N = 187) = 40.48, p < 0.001$) and divergence ($\chi^2(1, N = 163) = 24.35, p < 0.001$). No statistically significant differences were observed between frequencies of complementarity and divergence relations. The results do not support the hypothesis that science multimodal texts would be mostly characterized by relations of convergence regarding social distance (H1). Specifically, the hypothesis that verbal text and image would tend to promote small social distance (H3) was not confirmed. Moreover, in both prevalent relation categories, the verbal text tended to induce small social distance, while images tended to indicate moderate social distance in complementarity relations and large social distance in divergence relations (see Table 4). Figure 5 presents an example of a multimodal text characterized by complementarity, with the verbal text “Touch the little elephant’s ear and hear him!” [106] inducing small social distance (use of active voice and parataxis) and the image promoting moderate social distance (the elephant’s full body is depicted occupying more than 50% of the image space).



Figure 5. Example of a multimodal text indicating complementarity regarding social distance and convergence regarding involvement. Reproduced with permission from Hegarty, *Ta pio glika zoakia* [106]; published by Susaeta Hellas Publications: Athens, Greece, 2016.

With regard to involvement, the results indicated that relations of convergence were observed more frequently than expected compared to complementarity ($\chi^2(1, N = 215) = 33.61, p < 0.001$) and divergence ($\chi^2(1, N = 235) = 17.98, p < 0.001$). Therefore, these results support the hypothesis (H1) that science multimodal texts would mostly involve relations of convergence regarding involvement. However, this convergence is not in the expected direction, since verbal text and image tend to indicate weak involvement (Table 4) instead of strong reader involvement as expected (H4). For example, in Figure 5, verbal text and image are characterized by convergence, with the two semiotic modes indicating weak involvement of the reader (absence of possessive pronouns in the verbal text “Touch little elephant’s ear and hear him!” [106] and oblique angle in the image).

8. Discussion

This study explored the relations between the verbal and the visual mode at the level of interpersonal meaning (i.e., in terms of address, social distance, and involvement) in multimodal science texts for preschool children. The results presented in the previous section indicate that the two semiotic modes tend to promote convergent meanings regarding address, primarily serving the purpose of simply presenting knowledge to readers and assigning them a passive role of information receivers (Figure 6). Similar findings have also been reported by previous studies that analyzed the verbal [33,107] and both the verbal and the visual [29] mode in science texts for primary and secondary students.

Verbal text and image also converge in respect to the dimension of involvement, by mostly discouraging the child to become engaged with what is represented (Figure 6). These findings do not align with the socio-cognitive assumptions according to which children are agents in meaning construction, which presupposes their action, participation, and engagement in the learning process [36,37,41,48]. Therefore, as far as address and involvement are concerned, despite producing convergent meanings (which confirms H1), the two semiotic modes tend to refute the socio-cognitive perspective on which H2 and H4 have been based.

Moreover, as the results presented in the previous section indicate, verbal text and image promote either complementary or divergent interpersonal meanings regarding social distance. More particularly, while the verbal mode seems to consistently promote a relationship of familiarity and intimacy between the child and the represented knowledge, the visual mode does not follow this trend, by conveying either a neutral relationship (in the case of complementary meanings), a finding also reported previously regarding images in books for primary students [52], or a relationship of alienation (in the case of divergent meanings) between the child and represented knowledge (Figure 6).

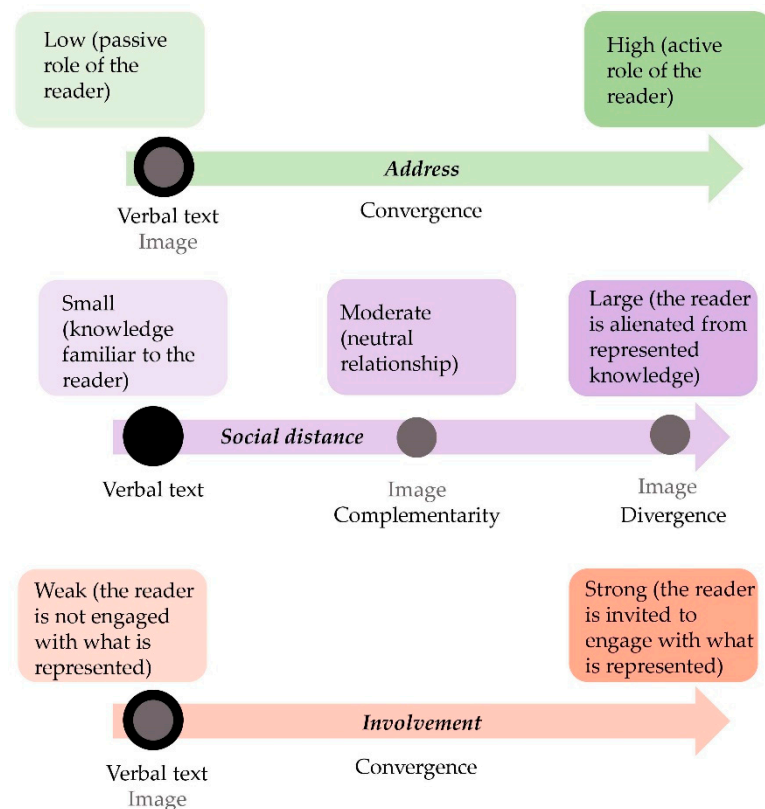


Figure 6. Verbal and visual selections in analyzed texts regarding address, social distance, and involvement.

The socio-cognitive perspective on science teaching and learning advocates that knowledge is not objective or isolated from the learner, but closely related to children's lives [39,48,50]. The analyzed texts do not seem to pay the necessary attention to the interpersonal meanings they disclose through the verbal text and the images they include, but also through the relations between them. The analyzed multimodal science texts for young children tend to promote pedagogically appropriate meanings only in terms of social distance and through the verbal mode, and sporadically through images. This seems to be a paradox, taking into account that preschool children, who are not yet 'literate' in the traditional sense, primarily rely on images in order to 'read' and derive meaning from multimodal texts [108,109]. Moreover, the tendencies of the analyzed material to promote verbal–visual relations of complementarity and divergence do not support young children's understanding. On the one hand, meaning complementarity between different semiotic modes increases the level of complexity in a text, posing challenges in its comprehension [69]. On the other, meaning divergence between different semiotic modes evokes confusion, making interconnection between them problematic [71] and obstructing—instead of supporting—the construction of scientifically adequate meanings [2]. Therefore, the significant number of divergent meaning instances in terms of social distance found in the present study could provoke confusion regarding the position of young readers in respect to the represented knowledge and their relationship with it.

This underestimation of the interpersonal meaning in the analyzed children's science texts might be due to the fact that in such informational texts the focus may lie on the representational meaning which carries the scientific information, i.e., content knowledge. Furthermore, these texts do not seem to adequately take into consideration the consistency between different semiotic modes and promote verbally and visually convergent interpersonal meanings, but rather focus on each of these modes independently—after all, the author and the illustrator of children's books are typically different people. The apparent inattention to interpersonal meanings and the relation between modes might also relate to the informational character of the analyzed texts and their audiences' characteristics,

such as restricted relevant knowledge. Thus, the selection of long or medium shot images—implying moderate or large social distance at the interpersonal level—might be considered as most appropriate at the representational level, since it allows the depiction of entities or phenomena in their entirety. Similarly, an oblique angle of image (indicating weak reader involvement) might be preferred so as to provide a general and ‘objectified’ view of what is represented. Likewise, the represented participants’ gazes could be directed to other elements in an image (e.g., systems, and processes) instead of the reader (low address) in order to direct young readers’ attention to specific and essential parts of the depicted knowledge.

Furthermore, the linguistic choices identified in the analyzed texts, could be associated with the young age of their intended readership. These selections involve the preference to active voice and parataxis (implying small social distance) and the emphasis on the representational meaning level by reference to the presented information through declarative clauses (denoting low address) and through the use of the third person in verbs and possessive pronouns (promoting weak involvement).

Therefore, semiotic selections apparently driven by an overemphasis on representational meaning have an impact on the interpersonal meanings promoted by verbal text and image, as well as on the relations between them. These in turn allow for the emergence of complementary or even divergent interpersonal meanings from the two modes, considered as inappropriate for young children, or for convergence relations that are incompatible with the socio-cognitive perspective for science learning in early childhood.

These findings are remarkable because informational science books are prevalent mediating tools in preschool education and constitute a significant part of young children’s first learning experiences with science concepts and phenomena. The interpersonal meanings promoted by texts such as the analyzed ones largely determine the quality and effectiveness of these learning experiences [30,31,33]. Furthermore, these experiences play an important role in the development of interests [20,110–112]; thus, interpersonal meanings promoted by such texts are expected to influence children’s science-related attitudes and aspirations.

One of the limitations of the current study is that it has explored verbal text–image relations regarding dimensions of one of the three different levels of meaning realized in a multimodal text, namely the interpersonal meaning. Further studies could also investigate the interaction between the three different levels of meaning, i.e., representational, interpersonal, and compositional. Specifically, it is of importance to investigate how interpersonal meanings promoted by verbal text or/and images are further strengthened or weakened through messages related to compositional meaning dimensions. For instance, in terms of compositional meaning, the image in Figure 4 through its large size and special position on the right, receives high salience and high information value, respectively. Therefore, it is presented as the most prominent element on the page and as the New, namely the information at issue, to which the reader’s attention is directed [25]. In this way, the pedagogically desirable interpersonal meaning (strong involvement) promoted by the image is further reinforced by the image’s compositional meanings, contrary to the weak involvement promoted by the verbal text, which is compositionally undervalued through the low salience (small size) and low information value (placement on the left) attributed to it. Similarly, in Figure 5, the distinctive features (bold font, colored letters, and framing) of verbal elements attribute high salience to the verbal text, attracting the reader’s attention [109]. Therefore, the pedagogically desirable interpersonal meaning of high address (use of imperative clause and second person) promoted by the verbal text, is further strengthened by the high salience it receives in terms of compositional meaning, counterbalancing the low address promoted by the image and its high salience (large size and distinctive texture of the animal’s ear) [53].

A further limitation is that the analysis framework developed and applied in this study as well as the discussion and interpretation of the findings regarding the semiotic selections, were based on specific theories of linguistic and visual communication, the Systemic Functional Grammar [24] and the Grammar of Visual Design [25]. This does not

mean that the authors of a text intend to convey the specific meanings or that the readers understand and interpret the various semiotic (verbal and visual) choices in the same way as suggested by the aforementioned theories. Specifically, how preschool children interpret interpersonal meanings promoted by verbal text and images or the relationship between the two semiotic modes in terms of interpersonal meaning when they interact with science multimodal texts, may form critical questions for future research. Last, due to the sampling procedure, the conclusions drawn from the findings concern the population from which the sample was drawn and no attempt is made to generalize the findings to the wider population of multimodal science texts for young children.

9. Conclusions

Our findings indicate that a greater emphasis on the interpersonal meaning and a more systematic analysis of the interaction and interrelation between verbal text and image in science texts for young children is required, aiming at facilitating understanding and supporting them in learning science. Furthermore, the analysis outlined in the present study may have considerable implications for the design, selection and use of multimodal science teaching material.

Concerning the selection and design of such material, attention to verbally and visually convergent meanings is recommended for every interpersonal meaning dimension, especially when a material is intended for very young, emerging readers. In fact, it has been shown that verbal text–image convergence facilitates successful association and coordination of the two modes, thus supporting children’s comprehension of multimodal science texts [66]. Additionally, it is suggested that convergence of verbal text and image harmonizes with the socio-cognitive perspective on teaching and learning, which implies that educational material should encourage—both verbally and visually—children’s activation, involvement, and familiarity with what is represented in teaching material.

Moreover, the analysis framework used in this study could support teachers in using multimodal science texts in their classroom in order to gradually initiate children in the critical analysis and interpretation of the interpersonal meanings conveyed by means of verbal text, image, and their interrelation in these texts. Children, from an early age, can be supported to effectively read and understand such material and use their relevant competencies to produce multimodal texts with scientific content [12,13,53]. Supporting children in developing multimodal communication competencies is a critical aspect of their scientific (visual) literacy and a broad objective of science education, for which preschool education can lay the foundations [9,41,111,112].

Author Contributions: Conceptualization, M.K. and V.C.; Data curation, M.K.; Formal analysis, M.K.; Methodology, M.K., V.C., M.P. and F.B.; Resources, M.K. and V.C.; Supervision, V.C.; Validation, M.K. and V.C.; Visualization, M.K.; Writing—original draft, M.K. and V.C.; Writing—review & editing, M.P. and F.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data sharing is not applicable to this article.

Acknowledgments: The authors wish to thank Psychogios Publications, rightsholders of *Geia sou poulaki*, for their permission to use the copyrighted material appearing in Figure 4 of the paper. Furthermore, the authors wish to thank Susaeta Hellas Publications, rightsholders of *Ta pio glika zoakia*, for their permission to use the copyrighted material appearing in Figure 5 of the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Jewitt, C.; Bezemer, J.; O’Halloran, K. *Introducing Multimodality*; Routledge: London, UK, 2016.
2. Givry, D.; Pantidos, P. Toward a multimodal approach of science teaching. *Skhole* **2012**, *17*, 123–129.

3. Jaipal, K. Meaning making through multiple modalities in a biology classroom: A multimodal semiotics discourse analysis. *Sci. Educ.* **2010**, *94*, 48–72. [[CrossRef](#)]
4. Kress, G.; Jewitt, C.; Ogborn, J.; Tsatsarelis, C. *Multimodal Teaching and Learning: The Rhetorics of the Science Classroom*; Continuum: London, UK; New York, NY, USA, 2001.
5. Pantidos, P. Narrating science in the classroom: The role of semiotic resources in evoking imaginative thinking. *J. Sci. Teach. Educ.* **2017**, *28*, 388–401. [[CrossRef](#)]
6. Pantidos, P. Epistemic, cognitive and semiotic significations in science teaching: The case of sound. *Eur. J. Educ. Stud.* **2019**, *6*, 210–231.
7. Pozzer-Ardenghi, L.; Roth, W. On performing concepts during science lectures. *Sci. Educ.* **2007**, *91*, 96–114. [[CrossRef](#)]
8. Ardenghi, L.P.; Roth, W.-M. *Staging & Performing Scientific Concepts: Lecturing Is Thinking with Hands, Eyes, Body, & Signs*; Sense Publishers: Rotterdam, The Netherlands, 2010.
9. Tytler, R.; Prain, V.; Peterson, S. Representational issues in students learning about evaporation. *Res. Sci. Educ.* **2007**, *37*, 313–331. [[CrossRef](#)]
10. Tang, K.; Delgado, C.; Moje, E.B. An integrative framework for the analysis of multiple and multimodal representations for meaning-making in science education. *Sci. Educ.* **2014**, *98*, 305–326. [[CrossRef](#)]
11. Cope, B.; Kalantzis, M. A grammar of multimodality. *Int. J. Learn.* **2009**, *16*, 361–425.
12. Coleman, J.M.; Bradley, L.G.; Donovan, C.A. Visual representations in second graders' information book compositions. *Read. Teach.* **2012**, *66*, 31–45. [[CrossRef](#)]
13. Danish, J.A.; Phelps, D. Representational practices by the numbers: How kindergarten and first-grade students create, evaluate, and modify their science representations. *Int. J. Sci. Educ.* **2011**, *33*, 2069–2094. [[CrossRef](#)]
14. Liu, Y. Teaching Multiliteracies in scientific discourse: Implications from symbolic construction of chemistry. *K@Ta* **2009**, *11*, 128–141. [[CrossRef](#)]
15. O'Halloran, K.L.; Tan, S.; Marissa, K.L.E. Multimodal analysis for critical thinking. *Learn. Media Technol.* **2017**, *42*, 147–170. [[CrossRef](#)]
16. Papadopoulou, M.; Gorla, S.; Manoli, P.; Pagkourelia, E. Developing multimodal literacy in tertiary education. *J. Vis. Lit.* **2018**, *37*, 317–329. [[CrossRef](#)]
17. Liu, Y.; Khine, M.S. Content analysis of the diagrammatic representations of primary science textbooks. *Eurasia J. Math. Sci. Technol. Educ.* **2016**, *12*, 1937–1951.
18. Pappas, C.C. The information book genre: Its role in integrated science literacy research and practice. *Read. Res. Q.* **2006**, *41*, 226–250. [[CrossRef](#)]
19. Mantzicopoulos, P.; Patrick, H. Reading picture books and learning science: Engaging young children with informational text. *Theory Pract.* **2011**, *50*, 269–276. [[CrossRef](#)]
20. Sackes, M.; Trundle, K.C.; Flevares, L.M. Using children's literature to teach standard-based science concepts in early years. *Early Child. Educ. J.* **2009**, *36*, 415–422. [[CrossRef](#)]
21. Bezemer, J.; Kress, G. Writing in multimodal texts: A social semiotic account of designs for learning. *Writ. Commun.* **2008**, *25*, 166–195. [[CrossRef](#)]
22. Lemke, J. Multiplying meaning: Visual and verbal semiotics in scientific text. In *Reading Science: Critical and Functional Perspectives on Discourses of Science*; Martin, J.R., Veal, R., Eds.; Routledge: London, UK, 1998; pp. 87–113.
23. Coleman, J.M.; Dantzler, J.A. The frequency and type of graphical representations in science trade books for children. *J. Vis. Lit.* **2016**, *35*, 24–41. [[CrossRef](#)]
24. Halliday, M.; Matthiessen, C. *An Introduction to Functional Grammar*, 3rd ed.; Arnold: London, UK, 2004.
25. Kress, G.R.; Van Leeuwen, T. *Reading Images: The Grammar of Visual Design*; Routledge: London, UK, 2006.
26. Xiao, J. Learner–content interaction in distance education: The weakest link in interaction research. *Distance Educ.* **2017**, *38*, 123–135. [[CrossRef](#)]
27. Anagnostopoulou, K.; Hatzinikita, V.; Christidou, V.; Dimopoulos, K. PISA test items and school-based examinations in Greece: Exploring the relationship between global and local assessment discourses. *Int. J. Sci. Educ.* **2013**, *35*, 636–662. [[CrossRef](#)]
28. Herbel-Eisenmann, B.; Wagner, D. A Framework for uncovering the way a textbook may position the mathematics learner. *Learn. Math.* **2007**, *27*, 8–14.
29. Hermawan, B.; Rahyono, F. Ideational meanings of science and interpersonal position of readers in science textbooks for basic level in Indonesia. *Indones. J. Appl. Linguist.* **2019**, *9*, 38–47. [[CrossRef](#)]
30. Dimopoulos, K.; Koulaidis, V.; Sklaveniti, S. Towards a framework of socio-linguistic analysis of science textbooks: The Greek case. *Res. Sci. Educ.* **2005**, *35*, 173–195. [[CrossRef](#)]
31. Elmiana, D.S. Pedagogical representation of visual images in EFL textbooks: A multimodal perspective. *Pedagog. Cult. Soc.* **2019**, *27*, 613–628. [[CrossRef](#)]
32. Ge, Y.P.; Unsworth, L.; Wang, K.H.; Chang, H.P. What images reveal: A comparative study of science images between Australian and Taiwanese junior high school textbooks. *Res. Sci. Educ.* **2018**, *48*, 1409–1431. [[CrossRef](#)]
33. Hoang, V.V. An interpersonal analysis of a Vietnamese middle school science textbook. In *Discourses of Southeast Asia: A Social Semiotic Perspective*; Rajandran, K., Manan, S.A., Eds.; Springer: Singapore, 2019; pp. 129–144.

34. Koulaidis, V.; Tsatsaroni, A. A pedagogical analysis of science textbooks: How can we proceed? *Res. Sci. Educ.* **1996**, *26*, 55–71. [[CrossRef](#)]
35. Martin, J.R.; Veel, R. *Reading Science: Critical and Functional Perspectives on Discourses of Science*; Routledge: London, UK; New York, NY, USA, 1998.
36. Ravanis, K.; Bagakis, G. Science education in kindergarten: Sociocognitive perspective. *Int. J. Early Years Educ.* **1998**, *6*, 315–327. [[CrossRef](#)]
37. Ravanis, K.; Koliopoulos, D.; Hadzigeorgiou, Y. What factors does friction depend on? A socio-cognitive teaching intervention with young children. *Int. J. Sci. Educ.* **2004**, *26*, 997–1007. [[CrossRef](#)]
38. Delsérieys, A.; Jégou, C.; Boilevin, J.M.; Ravanis, K. Precursor model and preschool science learning about shadows formation. *Res. Sci. Technol. Educ.* **2018**, *36*, 147–164. [[CrossRef](#)]
39. Ravanis, K. Early childhood science education: State of the art and perspectives. *J. Balt. Sci. Educ.* **2017**, *16*, 284–288.
40. Robbins, J. ‘Brown paper packages’? A sociocultural perspective on young children’s ideas in science. *Res. Sci. Educ.* **2005**, *35*, 151–172. [[CrossRef](#)]
41. Christidou, V.; Kazela, K.; Kakana, D.; Valakosta, M. Teaching magnetic attraction to preschool children: A comparison of different approaches. *Int. J. Learn.* **2009**, *16*, 115–128.
42. Impedovo, M.A.; Delsérieys-Pedregosa, A.; Jégou, C.; Ravanis, K. Shadow formation at preschool from a socio-materiality perspective. *Res. Sci. Educ.* **2017**, *47*, 579–601. [[CrossRef](#)]
43. Rau, M.A. Conditions for the effectiveness of multiple visual representations in enhancing STEM learning. *Educ. Psychol. Rev.* **2017**, *29*, 717–761. [[CrossRef](#)]
44. Shepardson, D.P. Learning science in a first grade science activity: A Vygotskian perspective. *Sci. Educ.* **1999**, *83*, 621–638. [[CrossRef](#)]
45. Poimenidou, M.; Christidou, V. Communication practices and the construction of meaning: Science activities in the kindergarten. *Creat. Educ.* **2010**, *2*, 81–92. [[CrossRef](#)]
46. John-Steiner, V.; Mahn, H. Sociocultural approaches to learning and development: A Vygotskian framework. *Educ. Psychol.* **1996**, *31*, 191–206.
47. Canedo-Ibara, S.P.; Castelo-Escandell, J.; Garcia-Wehrle, P.; Morales-Blake, A.R. Precursor models construction at preschool education: An approach to improve scientific education in the classroom. *Rev. Sci. Math. ICT Educ.* **2010**, *4*, 41–76.
48. Fler, M. Understanding the dialectical relations between everyday concepts and scientific concepts within play-based programs. *Res. Sci. Educ.* **2009**, *39*, 281–306. [[CrossRef](#)]
49. Fragkiadaki, G.; Ravanis, K. Genetic research methodology meets early childhood science education research: A cultural-historical study of child’s scientific thinking development. *Cult. Hist. Psychol.* **2016**, *12*, 310–330. [[CrossRef](#)]
50. Kambouri-Danos, M.; Ravanis, K.; Jameau, A.; Boilevin, J.M. Precursor models and early years science learning: A case study related to the water state changes. *Early Child. Educ. J.* **2019**, *47*, 475–488. [[CrossRef](#)]
51. Bungum, B. Textbook images: How do they invite students into physics? *Phys. Educ.* **2013**, *48*, 648. [[CrossRef](#)]
52. Lemoni, R.; Lefkaditou, A.; Stamou, A.G.; Schizas, D.; Stamou, G.P. Views of nature and the human-nature relations: An analysis of the visual syntax of pictures about the environment in Greek primary school textbooks—Diachronic considerations. *Res. Sci. Educ.* **2013**, *43*, 117–140. [[CrossRef](#)]
53. Koutsikou, M.; Christidou, V. The interplay between interpersonal and compositional meanings in multimodal texts about animals for young children. *Punctum* **2019**, *5*, 114–137.
54. Macken-Horarik, M. Interacting with the multimodal text: Reflections on image and verbiage in art express. *Vis. Commun.* **2004**, *3*, 5–26. [[CrossRef](#)]
55. Guijarro, A.J.M. The role of semiotic metaphor in the verbal–visual interplay of three children’s picture books: A multisemiotic systemic-functional approach. *Atlantis* **2016**, *38*, 33–52.
56. Haiyan, L. Image–text relations in junior high school EFL textbooks in China: A mixed-methods study. *J. Lang. Teach. Res.* **2018**, *9*, 1177–1190.
57. Hua-bing, L. Exploring image text relation in college English textbooks. *Sino-US Engl. Teach.* **2015**, *12*, 176–181. [[CrossRef](#)]
58. Painter, C.; Martin, J.R.; Unsworth, L. *Reading Visual Narratives: Image Analysis of Children’s Picture Books*; Equinox: Sheffield, UK, 2013.
59. Ainsworth, S. DeFT: A conceptual framework for considering learning with multiple representations. *Learn. Instr.* **2006**, *16*, 183–198. [[CrossRef](#)]
60. Hannus, M.; Hyönä, J. Utilization of illustrations during learning of science textbook passages among low-and high-ability children. *Contemp. Educ. Psychol.* **1999**, *24*, 95–123. [[CrossRef](#)]
61. Henderson, G. Learning with diagrams. *Aust. Sci. Teach. J.* **1999**, *45*, 17–25.
62. Jian, Y.C.; Ko, H.W. Influences of text difficulty and reading ability on learning illustrated science texts for children: An eye movement study. *Comput. Educ.* **2017**, *113*, 263–279. [[CrossRef](#)]
63. Mason, L.; Tornatora, M.C.; Pluchino, P. Integrative processing of verbal and graphical information during re-reading predicts learning from illustrated text: An eye-movement study. *Read. Writ.* **2015**, *28*, 851–872. [[CrossRef](#)]
64. Mason, L.; Tornatora, M.C.; Pluchino, P. Do fourth graders integrate text and picture in processing and learning from an illustrated science text? Evidence from eye-movement patterns. *Comput. Educ.* **2013**, *60*, 95–109. [[CrossRef](#)]

65. McTigue, E.M. Does multimedia learning theory extend to middle-school students? *Contemp. Educ. Psychol.* **2009**, *34*, 143–153. [[CrossRef](#)]
66. Koutsikou, M.; Bonoti, F.; Christidou, V. The effect of explanatory captions on understanding a scientific explanation. *Int. J. Res. Educ. Methodol.* **2015**, *7*, 1127–1138.
67. Daly, A.; Unsworth, L. Analysis and comprehension of multimodal texts. *Aust. J. Lang. Lit.* **2011**, *34*, 61–80.
68. Meneses, A.; Escobar, J.P.; Véliz, S. The effects of multimodal texts on science reading comprehension in Chilean fifth-graders: Text scaffolding and comprehension skills. *Int. J. Sci. Educ.* **2018**, *40*, 2226–2244. [[CrossRef](#)]
69. Unsworth, L.; Chan, E. Bridging multimodal literacies and national assessment programs in literacy. *Aust. J. Lang. Lit.* **2009**, *32*, 245–257.
70. Mayer, R.E. *The Cambridge Handbook of Multimedia Learning*, 2nd ed.; Cambridge University Press: New York, NY, USA, 2014; p. 930. [[CrossRef](#)]
71. Chan, E. Integrating visual and verbal meaning in multimodal text comprehension: Towards a model of intermodal relations. In *Semiotic Margins: Meanings in Multimodalities*; Hood, S., Dreyfus, S., Stenglin, M., Eds.; Continuum: London, UK; New York, NY, USA, 2011; pp. 144–167.
72. Rearden, K.T.; Broemmel, A.D. Beyond the talking groundhogs: Trends in science trade books. *J. Elem. Sci. Educ.* **2008**, *20*, 39–49. [[CrossRef](#)]
73. Baldry, A.; Thibault, P.J. *Multimodal Transcription and Text Analysis: A Multimedia Toolkit and Coursebook*; Equinox: London, UK, 2006.
74. Martínez, M.; Harmon, J.M. Picture/text relationships: An investigation of literary elements in picturebooks. *Lit. Res. Instr.* **2012**, *51*, 323–343. [[CrossRef](#)]
75. Shin, D.; Cimasko, T.; Yi, Y. Development of metalanguage for multimodal composing: A case study of an L2 writer’s design of multimedia texts. *J. Second Lang. Writ.* **2020**, *47*, 100714. [[CrossRef](#)]
76. Manoli, P.; Papadopoulou, M. Greek students’ familiarity with multimodal texts in EFL. *Int. J. Literacies* **2013**, *19*, 37–46. [[CrossRef](#)]
77. Mills, K.A.; Unsworth, L. Multimodal Literacy. In *Oxford Research Encyclopedia of Education*; Oxford Research Encyclopedias: Oxford, UK, 2017.
78. Sipe, L.R. Revisiting the relationships between text and pictures. *Child. Lit. Educ.* **2012**, *43*, 4–21. [[CrossRef](#)]
79. Arifianto, R. Analysis of Verbal and Visual Interpersonal Meanings in Children Picture Books Entitled “Waldo and the Desert Island Adventure” (1986), “Let’s Be Friend Again” (1988) and “Never Lonely Again” (1988) by Hans Wilhelm. Ph.D. Thesis, Indonesia University of Education, Bandung, Indonesia, 2015.
80. Guijarro, A.J.M. A multimodal analysis of “The tale of peter rabbit” within the interpersonal metafunction. *Atlantis* **2010**, *32*, 123–140.
81. Guijarro, A.J.M. Engaging readers through language and pictures: A case study. *J. Pragmat.* **2011**, *43*, 2982–2991. [[CrossRef](#)]
82. Hermawan, B.; Sukyadi, D. Ideational and interpersonal meanings of children narratives in Indonesian picturebooks. *Indones. J. Appl. Linguist.* **2017**, *7*, 404–412. [[CrossRef](#)]
83. Setyowati, L.A. Interpersonal Meaning Analysis in Short Story of Hans Christian Andersen The Real Princess. Ph.D. Thesis, Universitas Muhammadiyah Surakarta, Surakarta, Indonesia, 2016.
84. Yefymenko, V. Text–image relationships in contemporary fairy tales. *Linguist. LingBaW* **2017**, *3*, 216–228. [[CrossRef](#)]
85. Zohrabi, M.; Dobakhti, L.; Mohammadpour, E. Interpersonal meanings in children’s storybooks. *Iran. J. Lang. Teach. Res.* **2019**, *7*, 39–64.
86. Hatnikita, V.; Dimopoulos, K.; Christidou, V. PISA test items and school textbooks related to science: A textual comparison. *Sci. Educ.* **2008**, *92*, 664–687. [[CrossRef](#)]
87. Jian, Y. Fourth graders’ cognitive processes and learning strategies for reading illustrated biology texts: Eye movement measurements. *Read. Res. Q.* **2016**, *51*, 93–109.
88. Pentucci, M. Evolutions in the multimodal design of Italian textbooks. *Soc. Sci. Learn. Educ. J.* **2020**, *5*, 6–14. [[CrossRef](#)]
89. Peterson, M.O. Schemes for integrating text and image in the science textbook: Effects on comprehension and situational interest. *Int. J. Environ. Sci. Educ.* **2016**, *11*, 1365–1385.
90. Postigo, Y.; López-Manjón, A. Images in biology: Are instructional criteria used in textbook image design? *Int. J. Sci. Educ.* **2019**, *41*, 210–229. [[CrossRef](#)]
91. Royce, T. Multimodality in the TESOL classroom: Exploring visual-verbal synergy. *TESOL Q.* **2002**, *36*, 191–205. [[CrossRef](#)]
92. Unsworth, L. Image/text relations and intersemiosis: Towards multimodal text description for multiliteracies education. In *Proceedings of the 33rd IFSC International Systemic Functional Congress, Sao Paulo, Brazil, 10–15 July 2006*; Pontificia Universidade Católica de Sao Paulo: Sao Paulo, Brasil, 2007; pp. 1165–1205.
93. Vinisha, K.; Ramadas, J. Visual representations of the water cycle in science textbooks. *Contemp. Educ. Dialogue* **2013**, *10*, 7–36. [[CrossRef](#)]
94. Yopp, R.H.; Yopp, H.K. Young children’s limited and narrow exposure to informational text. *Read. Teach.* **2012**, *65*, 480–490. [[CrossRef](#)]
95. Cole, M.; Cole, S.R. *The Development of Children*; W. H. Freeman: New York, NY, USA, 1996.
96. Feldman, R.S. *Development across the Life Span*, 4th ed.; Pearson Education: Auckland, New Zealand, 2006.

97. Lombard, M.; Snyder-Duch, J.; Bracken, C.C. Content analysis in mass communication: Assessment and reporting of intercoder reliability. *Hum. Commun. Res.* **2002**, *28*, 587–604. [[CrossRef](#)]
98. Neuendorf, K.A. *The Content Analysis Guidebook*; Sage Publications: London, UK; New Delhi, India, 2002.
99. Bonoti, F.; Misalidi, P. Social emotions in children's human figure drawings: Drawing shame, pride and jealousy. *Infant Child Dev.* **2015**, *24*, 661–672. [[CrossRef](#)]
100. Kleinelumern-Depping, A.; Langner, C. *To Proto Mou Lexiko me Eikones: Ta Zoa tou Zoologikou Kipou [My First Dictionary with Pictures: The Animals of the Zoo]*; Patakis Publications: Athens, Greece, 2016.
101. Godwin, S. *Fotizi Fotizi o Ilios. Mia Proti Matia Sto Fos [It Shines, the Sun Shines. A First Look at the Light]*; Patakis Publications: Athens, Greece, 2015.
102. Godwin, S. *Eksafanistike, Eksafanistike i Kampia. Mia Proti Matia Ston Kiklo Zois Tis Petaloudas [The Caterpillar Has Disappeared. A First Look at the Life Cycle of Butterfly]*; Patakis Publications: Athens, Greece, 2017.
103. Roca, N.; Isern, C. *O Ilios [The Sun]*; Tziampiris-Piramida Publications: Thessaloniki, Greece, 2017.
104. Collective Work. *Atlantas gia Paidia: Fantastiko Anthropino Soma [Atlas for Children: Fantastic Human Body]*; Dioptra Publications: Athens, Greece, 2017.
105. Spiro, R. *Geia sou Poulaki! [Hello Birdy!]*; Psychogios Publications: Athens, Greece, 2018.
106. Hegarty, P. *Ta Pio Glika Zoakia [The Cutest Animals]*; Susaeta Hellas Publications: Athens, Greece, 2016.
107. Schleppegrell, M.J. *The Language of Schooling: A Functional Linguistics Perspective*; Routledge: Mahwa, NJ, USA, 2004.
108. Papadopoulou, M.; Christidou, V. Multimodal text comprehension and production by preschool children. *Int. J. Learn.* **2004**, *11*, 917–927.
109. Papadopoulou, M.; Manoli, P.; Zifkou, E. Typography, how noticeable is it? Preschoolers detecting typographic elements in illustrated books. *Int. J. Early Child. Learn.* **2014**, *20*, 23–36. [[CrossRef](#)]
110. Osborne, J.; Simon, S.; Collins, S. Attitudes towards science: A review of the literature and its implications. *Int. J. Sci. Educ.* **2003**, *25*, 1049–1079. [[CrossRef](#)]
111. Samarapungavan, A.; Mantzicopoulos, P.; Patrick, H. Learning science through inquiry in kindergarten. *Sci. Educ.* **2008**, *92*, 868–908. [[CrossRef](#)]
112. Spektor-Levy, O.; Baruch, Y.K.; Mevarech, Z. Science and scientific curiosity in pre-school: The teacher's point of view. *Int. J. Sci. Educ.* **2013**, *35*, 2226–2253. [[CrossRef](#)]