Production of multimodal texts in Secondary Education: A case study

Marina Rodosthenous-Balafa, Agni Stylianou-Georgiou, and Elisavet Pitri

This paper describes the theoretical framework and the methodology followed by three researchers from different disciplines (Literature, Art Education and Educational Psychology) in order to present the challenges 12-15-year-old students faced in the production of a multimodal text in a school context. It examines the way students address specific dimensions of meaning (representational, social and organizational) within and across linguistic and visual-spatial modes, when producing a multimodal text. Specifically, it demonstrates students' challenges in addressing the different dimensions of meaning and achieving synergy among different semiotic systems in creating multimodal artefacts. Qualitative methods were utilized to analyze how groups addressed linguistically and visually the representational, social and organizational dimensions of meaning for the production of their multimodal texts and how they reflected on the production of their multimodal meaning making. The results of this study shed light on the difficulties the students have faced in the production of multimodal meaning making. Moreover, the study stresses the need to support students through systematic approaches on how to address the different dimensions of meaning in multimodal text production. Finally, several suggestions are provided regarding further research on studying in depth the process of producing multimodal texts in school context.

KEYWORDS   multimodal texts, secondary school, framework, multimodal meaning making

1. Introduction

Although multimodality is not a contemporary phenomenon, during the last decades it has been approached anew as a growing interdisciplinary field of research. Adami considers

multimodality as 'a phenomenon of communication', which 'defines the combination of different semiotic resources, or modes, in texts and communicative events, such as still and moving image, speech, writing, layout, gesture, and/or proxemics' (2017: 451). Modes have 'an organizing principle' (Adami 2017: 451) and require various researchers from different disciplines and backgrounds to study and add new perspectives to the way these modes work together, interact and often contradict one another in order to communicate information and render narratives (Serafini 2014: 15). One of the main assumptions that recipients of multimodal texts should bear in mind is that, when modes concur, each acquires a specialized role in meaning (Jewitt 2008). Each mode 'offers distinct possibilities and constraints' (Jewitt, Bezemer & O'Halloran 2016: 3), or, in Adami's words, modes have 'a specific functional load to the meaning made by the overall text – and as such they deserve attention' (2017: 452).

Apart from the linguistic modes of writing and speech, other modes, through which contemporary societies experience the world, are the visual ones (Albers 2007; Kress 2003). The spectrum of visual production and communication is becoming broader. All that is humanly formed and sensed through vision or visualization shape the way we live our lives, and they are defined as visual culture. Visual culture points to connections between popular and fine art forms, and is considered inherently interdisciplinary and increasingly multimodal (Freedman 2003). Understanding the visual structures that exist within images is as important as comprehending written text (Harste 2003). Berger (1972) suggested that reading images involves a process of seeing and taking notice of what we choose to look at. Viewers, through immersion and recognition, learn to notice and identify relationships across visual text. Research has shown that viewers' interests and individual differences may affect the reading path they choose when they encounter an image, which will ultimately reflect their effort in visual reading (Cope and Kalantzis 2009).

As mentioned above, each mode in an ensemble carries only part of the message, so it is indispensable to study the interplay between modes in a certain multimodal artefact and to examine the way each semiotic system interacts with and conduces to the others (Bezemer and Jewitt 2009: 9). Bezemer and Jewitt observe that, during the interplay of the modes, the reader may realize meaningful tensions between the various textual aspects, which lead to reflection and finally expand the meaning (2009: 9).

In 2015, Theo van Leeuwen, in his paper 'Multimodality in Education: Some Directions and Some Questions', stresses the significant impact multimodality has had on educational theory and practice and formulates a series of questions on multimodality in education, which need to be addressed. Some of the questions he raises are the following: 'Should we see multimodality as something in which students are already literate, or is there a role for educators in developing multimodal literacies [...]. If so, in which directions, and why?' (2015: 582). Acknowledging van Leeuwen's directions for the study of multimodality in educational contexts,

we designed an empirical case study to explore the difficulties that students may have when producing a multimodal text.

This paper will discuss the results of a case study on multimodal text production in a school context. The study was initiated when the researchers were invited by a secondary school to plan a workshop on multimodal production, in order to assist students to participate in a competition organised by the Ministry of Education. The paper will demonstrate students' challenges in addressing the different dimensions of meaning and achieving *synergy* (Sipe 1998: 98-99) among semiotic systems in creating multimodal artefacts. The main research question is: How do students address specific dimensions of meaning (representational, social and organizational) within and across linguistic and visual-spatial modes, when producing a multimodal text? The results of this study shed light on the difficulties the students have faced at the production of multimodal meaning making and indicated the need to support students grappling with the dimensions of multimodal meaning. Several suggestions are provided regarding further research on in-depth studies of the complex process of producing multimodal texts in school contexts.

2. Theoretical framework

The main theoretical framework guiding this study comprises the 'Grammar of Multimodality' by Cope and Kalantzis (2009). The approach of 'Slow Looking', introduced by Tishman (2018), is utilized to discuss how such a framework could illuminate the multiple dimensions of meaning and uncover the complexity of interconnections among different modes.

Multiliteracies, in an era where communications are multimodal and an image culture has risen (Kress 2009), pose challenges for learners, since 'new structures of agency' emerge in the meaning-making process (Kalantzis and Cope 2015: 17). Cope and Kalantzis (2009) suggested a multimodality grammar to emphasize the parallels and the affordances shared among different modes of meaning. They raise important questions about meaning-making, which relate to five dimensions:

1. Representational dimension: 'What do the meanings refer to?' (2015: 19).[1] When we write or speak, we represent things with nouns and processes with verbs. In images, representation is achieved through line, form and color. In tactile representation, edges and texture are utilized for representation, whereas in gesture representation, different means are used, such as beat, as well as acts of pointing. Similarly, each mode can uniquely add to the representational aspect of meaning (i.e., a word can refer to a concept, an image can serve as a symbol and a sound can be used to represent the main idea). The act of referring can also establish relations (through prepositions in language, contrast in image) and qualities (adjectives in language, visual attributes in images).

2. Social dimension: 'How do meanings connect the participants in meaning-making?' (2015: 20). In meaning-making, we establish roles (speaker/listener, designer/user, sound maker/hearer), direct orientations (first/second/third person in language, placement of objects in image, pointing to self/others/world in gesture), encounter agency (voice/mood in language, focal planes of engagement in image, openings/barriers in space) and interpret potentials (open/closed texts, realistic/abstract images, directive/turn-taking gestures).

3. Organizational dimension: 'How does the overall meaning hold together?' (2015: 20). Internal cohesion, coherence, and boundedness in meaning are achieved through ordering the atomic meaning units in each mode. Morphemes, picture elements, physical components, and strokes in gesture can be ordered in conventional or inventive ways. For example, ideas can be arranged through sequence in the text and by positioning picture elements in images.

4. Contextual dimension: 'Where is the meaning situated?' (2015: 20). Context is important in meaning-making in the sense that the meaning of an object ('what') is determined by the place it appears ('where'). Therefore, context is part of the meaning. It defines how the meanings fit into the larger world of meaning. For example, an object such as a ring can mean different things depending on its context.

5. Ideological dimension: 'Whose interest is a meaning designed to serve?' (2015: 21). It is important to consider motivation in interrogating meanings. Considering subjectivity or objectivity in verbal or visual modes in a multimodal text requires critical thinking so that text, image, gesture, sound, or space are explored to uncover concealed interests.

Cope and Kalantzis (2009: 363) have used the word *'synaesthesia'* to refer to the process of switching representational modes to convey the same or similar meanings. In a multimodal text, each mode, which constitutes a distinct semiotic resource, has unique meaning potential to recognize or adequately use the meaning and learning potentials inherent in different modes. Representational parallels make synaesthesia possible, because each mode provides an added value in the process of meaning-making when producing a multimodal text.

Producing the multiple dimensions and interconnections of meanings requires skills for slow and attentive viewing. In Tishman's terms, '[s]low looking means taking the time to carefully observe more than meets the eye at first glance. [...] Learning through prolonged observation can occur through all of the senses' (2018: 2). Seeing constitutes a meaning or interpretation through viewing (Berger 1972), which implies that it is an active process. Lancaster (2001) suggests that gaze is an action, influencing what is interpreted in visual and spatial modes.

Tishman (2018) suggests several strategies for slow looking. One of them is the use of certain categories to guide the eye in identifying elements and principles of design in an

image, such as color, shape, line, balance, and pattern. A complementary strategy is switching up one's point of view (physical perspective) to see things in a new way and to describe them from a different angle. This approach, as Tishman (2018) points out, emphasizes looking at things from the vantage point of different physical perspectives – far away or close up, above or below.

Cope and Kalantzis' framework was used to design an empirical case study in a school context. The aim was to guide students to carefully view several dimensions of meaning-making through the production of a multimodal text.

3. Methodology

A. Participants

The participants in this case study were 208 students (46 groups of maximum 5 students), twelve to fifteen years old, at a public junior high school (gymnasium) in Cyprus, and three researchers-academics specializing in Literature, Art Education, and Educational Psychology.

B. Procedure

This study was conducted in two phases. In phase one, students were introduced to the core idea that when producing a multimodal text, meaning-making operates at several levels. In order to emphasize the parallels shared among the different modes of meaning, researchers presented examples of the ways multimodal texts (comic books, picture books and advertisements) are read, explaining how each dimension was addressed and emphasizing the importance of synergy among them. Researchers elaborated on how meaning is created through the relation among modes and the social values these meanings transmit to the text. This introductory presentation concluded with the theoretical framework of Cope and Kalantzis (2009) regarding a 'Grammar of Multimodality'. Specifically, students were introduced to important questions about how meaning is created in a multimodal text through five dimensions (representational, social, organizational, contextual, ideological).

Thereafter, students were presented with a task requiring them to collaborate in groups to create a multimodal text during a 60-minute school session, utilizing a provided image outline and a matrix. More specifically, students were asked to combine at least three semiotic modes by enriching the visual mode (i.e., adding colors, lines, additional patterns), developing a hand-written text of approximately 50 words (either prose or verse), and adding movement or sound or both in relation to their given transformed outline.[2] A certain guideline required the students to identify the narrator of their linguistic text with any inanimate figure.

The choice to have an image as a starting point was based on Kress (2003), considering

the reading path of an image being more open than that of writing. As explained by Cope and Kalantzis, the reading path in a written text is clear even if it demands some semantic filling, because its syntax provides 'stable blueprints for semantic interpretation' (2009: 364). In the case of images, the syntax is in the hands of the viewer, since they are not characterized by linearity. Researchers aimed to challenge students to experience the production of a multimodal text through a non-linear process for meaning-making.

During phase two, students were presented with the multimodal texts created collaboratively during phase one. The purpose of this phase was to have students reflect on their previous work and examine the synergy between linguistic and visual modes of their multimodal artefacts. Specific prompts were used to direct students' attention to the representational, social and organizational dimensions of meanings in their texts, in order to proceed to written reflections on their multimodal productions.

C. Instruments

Image outline for producing the group multimodal text

The outline used as a starting point for students' multimodal text production was an illustration from the picture book *Erotokritos* (Diaplasi 2015).[3] The text was an adaptation of the homonymous Cretan Renaissance romance (sixteenth-seventeenth centuries) by Vitsentzos Kornaros, whose length is more than 10,000 fifteen-syllable coupled verses. The poetic adaptation of the picturebook was made by Katerina Mouriki and Ioanna Kyritsi-Tzioti, the illustrations by Iris Samartzi,[4] and the music was specially composed for this multimodal text by Nikos G. Papadogiorgos.
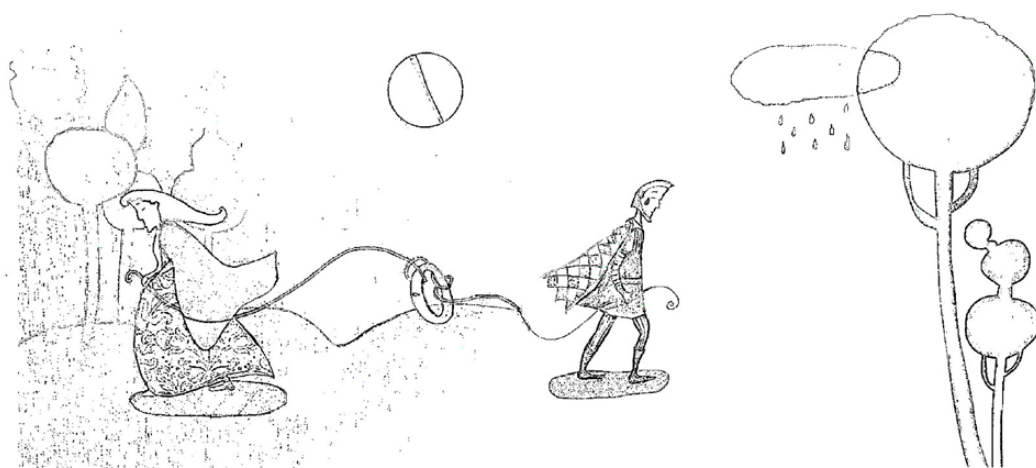
Figure 1. Outline of a picture book image as a starting point for students' multimodal meaning making

Although the students were familiar with the prototype text of *Erotokritos*, since particular passages of the particular work are taught at gymnasium, researchers did not unveil the source of the illustration. That is because, they aimed at encouraging students to recreate a new context of the image based on their own reading of the outline. In any case, the original illustration was modified in order to present to the students a black-and-white outline (see Figure 1). Freedman (2003) describes how approaching visual culture in education requires consideration of the relationship between form, feeling, and knowing to learning. Researchers chose an image that was neither simplistic nor predictable, neither chaotic nor unrelated to anything that the students have seen before.

Matrix introducing the grammar of multimodality

A matrix with guiding questions about the representational, social, and organizational dimensions of meaning (Cope and Kalantzis 2009) in different modalities (linguistic, visual-spatial, gestural, audio) was provided to students (see Table 1). The matrix aimed at introducing the complexity of creating meaning within and across modes.

Specifically, with regards to the representational dimension, students were prompted to think about who speaks, to whom he/she/it addresses, and what he/she/it says (linguistic mode). They were also asked to observe the details of each element included in the outline-image and think about the placement of objects in the image space, the significance of space among elements, the colors they would use to paint the objects, as well as why the two figures were facing opposite directions (visual-spatial mode). As far as the gestural mode was concerned, students were asked to think about the elements' implied movement and how they would represent movement in the image. In case students would choose to add an audio mode to their multimodal text, they would have to think about how to produce sound and rhythm for selected elements of the image.

To address the social dimension of meaning in their multimodal text, students were guided to think how each element of the image speaks based on its social status or physical position and the relationship among the elements (linguistic mode). Furthermore, they were prompted to think about the perspective of each element based on its position (visual-spatial mode), how elements would move according to their feelings, attributes, social status, or physical position (gestural), or sound based on the elements' attributes (audio mode).

The organizational dimension was also introduced in the matrix, prompting students to think how the meanings are interconnected across the different modes. Students were asked to decide whether they would compose a monologue or a dialogue among elements (linguistic mode), the meaning of each element's location (still or moving), possible interactions among elements, and whether there was harmony, balance, and contrast in their relationships (visual-spatial mode). There were also guiding questions to think about how to present ges-

tures, facial expressions, posture (gestural mode), and the kind of music or sound that elements would produce (audio mode).

Dimensions of meaning	Linguistic	Visual-spatial	Gestural	Audio
Representational What do the meanings refer to;	Who speaks, whom does he/she/it address to and what does he/she/it say?	What details do you see in each element? Why each element is placed in a certain space at the image? Why do the figures face opposite direction? What is the significance of space among elements? In what colours would you paint each element?	Is movement implied by elements? How would you represent movement in the image?	Produce sound and rhythm for selected elements of the image.
Social How do the meanings connect the elements they involve?	How does each element of the image speak based on its social status or physical position? What is the relationship among the elements?	What is the perspective of each element? Why are elements placed in the specific space?	How do they move according to their feelings, attributes, social status or physical position?	How does each element sound based on its attributes? Mood, volume
Organizational How do the meanings interconnect?	Is there a monologue or a dialogue among the elements?	Where are the elements placed in the image and what does their particular location mean? Are the elements still or moving? Do elements interact with each other? And if so, is there harmony, balance, contrast in their relationships? (If you wish, make modifications in the image.)	Gestures, facial expressions, posture?	What kind of music or sound would you produce?

Table 1. Matrix introducing the grammar of multimodality

Self-reflective questions about multimodal meaning-making

Students were asked to reflect on their choices for creating meaning in their multimodal artefacts, describing the following dimensions:

a. Representational: What are the meanings of your multimodal text? Which visual

elements (such as color: warm, cold; type of lines: curve, straight, dotted; direction: vertical, diagonal; composition: adding new elements) are used to represent these meanings? Which verbal elements of your multimodal text are associated with the meanings you assign to the text?

b. Social: How did you consider perspective or focal planes or emphasis in order to draw the reader-viewer of your text into your meaning?

c. Organizational: How do visual elements relate to speech? How are verbal elements linked to the image?

Students' reflections were followed by the researchers' insights concerning technical aspects of how visual and linguistic characteristics of the multimodal texts could represent examples of meanings related to the initial black-and-white outline of the illustration.

D. Data analysis

Qualitative methods of analysis were utilized in the study. A matrix based on an adapted version of the grammar of multimodality framework by Cope and Kalantzis (2009) was used to analyze how groups addressed linguistically and visually the representational, social, and organizational dimensions of meaning for the production of their multimodal texts.

Analysis of the representational dimension of meaning in the linguistic mode focused on identifying the content, through certain themes and motives in the written text, whereas, in the visual-spatial mode, the analysis focused on identifying specific elements and principles of design in students' drawings. More specifically, students' choices of color properties, lines, and form were described.

Analysis of the social dimension of the linguistic text was based on roles and directed orientations (person of narration, vantage point, roles, attributes of the figures, status), as revealed by the students' views. As far as the visual-spatial mode was concerned, the social dimension was studied based on the visual emphasis of each element added to the given image outline in relation to its role in the written text.

Analysis of the organizational dimension of group multimodal texts focused on the interconnection of meanings, on how that was achieved through the linguistic mode (monologue or dialogue among the elements prose, rhyming or free verse, dialogic form), and through the visual-spatial mode (composition, harmony, balance, contrast).

The two researchers specialized in Literature and Art examined the linguistic and visual modes respectively and analyzed interconnections between the two modes. For reliability control, the two researchers discussed with the third researcher their findings until one hundred per cent agreement was reached.

Students' responses to the self-reflective questions were analyzed to check whether the

visual and linguistic texts denoted synergy among the modes and the dimensions of meanings. More specifically, researchers examined group responses' clarity and depth in stating the dominant meanings of their multimodal texts and in defining both the linguistic and visual choices.

4. Results

Analysis of students' multimodal artefacts is presented based on the three dimensions of meaning (representational, social, and organizational). Each dimension is discussed in relation to the linguistic, the visual mode, and the synergies attained within and across modes.

a.   Representational dimension of meaning in students' multimodal texts

Content is the aspect that relates to the representational dimension of the linguistic mode. The largest number of texts referred to an intentional split, because one of the characters of the couple could not live in harmony with the other person. Although the separation was deliberate, it was mentioned, in some cases, that the characters were still in love. Moreover, a few noted that the characters remained united or that they would reunite. Almost a quarter of the texts referred to an unintentional separation that was due to war, death, rival kingdoms, different religion, and different social status between the couple. The sociocultural differences caused parents' disapproval of the couple's unity.

In relation to the overall analysis of students' visual-spatial elaboration of the text, focusing on the representational dimension showed that, even if some students had developed technical knowledge of elements and principles (such as achieving color tone gradient, color mixing, intensity, variety of lines, and textures), the above characteristics in the majority of the drawings seemed random, with no obvious symbolism or connection to the linguistic text. Symbolism and relevance to the text was, to a small degree, achieved when students chose to add elements and shapes to the image that were not included in the initial outline, rarely leading to original additions and abstraction. For example, a group of students chose to add the outline shape of a lightning below the top-centered split circle, directed towards the ring (see Figure 2). This group also transformed the outlines of the trees and the cloud into abstract human figures, and the shadow of the figures into red, broken hearts, which were also viewed as symbols of separation.

Each element was signified by its shape and position in the composition of the given outline, and it often affected the choice of color. For example, in all drawings, the circle in the top center of the image was assigned the meaning of a sun or moon, or a combination of both, due to the dividing diagonal line in the original outline, and was colored accordingly.

Sometimes, the students added rays. The oval shape in the lower center of the image, between the two figures, was assigned the meaning of an engagement/wedding ring. Therefore, it was colored yellow or orange in the majority of drawings to signify gold, or an extra element was added, such as a diamond. Additionally, the drops under the cloud, on the top-right side of the image, represented raindrops, but in some images, they were connected to the tear on the male figure's face. In the majority of drawings, the specific cloud was painted gray or black, symbolizing the male figure's negative emotional state.

Figure 2. Synergy of linguistic and visual modes achieved in students' multimodal text

Color choices, mainly in relation to inanimate elements, were not connected to the story of the linguistic text. It was observed that there was no consistency between the content and verbal and visual modes. For example, more surrealistic elements were incorporated in the visual text, such as pink or multicolor trees, a blue moon, orange clouds, or personified trees, whereas no relevant diction and style were found in the linguistic mode.

b. Social dimension of meaning in students' multimodal texts

The person of the narration refers to the social dimension of multimodal meaning-making. The majority of the texts were written in third-person singular. However, it was difficult to

understand which object narrates the story, since it was rarely directly mentioned or visually denoted. The narrator's identification could have also been revealed through the style of the supposed 'idiolect' of the object's language (e.g., How does a tree or a ring talk? Are there particular ways to make a moon speak?). However, clarifying an inanimate narrator seemed to be challenging for students, possibly because students might not frequently represent the speech of such types of narrators. Special practice is needed in order to illustrate or depict in stylistic verbal terms personified perspectives of objects. The element chosen from the image to use first-person singular was frequently denoted either in words (for example 'Cloud: I do not like ...') or by visually transferring the particular words of the narrator in added speech bubbles. This addition demonstrated an organizational link across modes to identify the narrator. The first-person plural was mainly used when the trees spoke.

Apart from the first-person narration, the students used the second person singular in order to give voice to the objects, especially when they needed the objects to directly counsel the couple. Frequently, the phenomenon of polyphony appeared, in the sense that more than one object talked. However, in some cases the polyphony was counterpart in the text, meaning that the two objects speaking to the couple gave different advice to each of them. For example, the cloud advised the man to walk away from the woman and never come back, whereas the sun advised him to remain in the relationship. In some cases, apart from the direct speech of the objects, the inanimate elements shared a dialogue between themselves. Frequently, the students thought that the first or even the second-person narration was not enough to complete the text, so they also incorporated a third person narration in conjunction with the direct speech of the objects.

Two extraordinary narrators that were observed when analyzing the multimodal texts were God and a cloud-poet. In the case of the narrator, who was presented to be God, an extra visual element (a dome) was added, which also showed a good organizational relationship between the representational and social dimensions of the two modes (verbal and visual). God's advice to the couple stylistically imitates a phrase taken from the *New Testament*, which priests often use in church. The other case is when the cloud invoked the sun to recount the story of the resourceful male character of the image:

Σύννεφο: 'Τον άντρα, ήλιε, τον πολύτροπο να μου ανιστορήσεις
Που έφτασε μέχρι την άκρη του φωτός για να την αντικρίσει!'
[Cloud: 'Recount, sun, for me the story of the resourceful man,
                Who reached the edge of light to meet his beloved.']

This passage has clear intertextual elements with Homer's *Odyssey*. More specifically, it recalls the way the poet invokes the muse at the beginning of this exemplary epic. Students' use of the most representative adjective of Odysseus, as appeared in the first line of *Odyssey*, 'πολύτροπον', shows clearly this intertextuality along with the invocation of a celestial power. It

is noteworthy to mention that the literal meaning in Greek of the adjective the students used (πολύτροπο) is 'multimodal'. Although one cannot be certain whether the choice of the particular word was intentional in relation to the multimodal practice the students experienced, it makes us realize how many centuries the etymology of this growing field of multimodality (πολυτροπικότητα) dates back to, despite that, in the *Odyssey*, the term is used in a different context.

Studying the social dimension of the visual-spatial mode, the researchers focused on identifying elements and principles that enhance aspects of the characters' physical appearance that symbolize feelings, attributes and social status as described in the texts. More specifically, design elements that affect the perspective in an image, such as color value, lines, axes and shapes, were examined in relation to the written text in order to identify students' intentional use of elements to emphasize the role of each character of their story. It was found that, in the majority of drawings, groups employed either color or line to symbolize the social status of the figures. For example, the figures of happy princes and princesses were in bright colors, whereas, sad or belligerent figures were either black or other dark colors. Additionally, straight lines in the same direction depicting smooth textures were utilized to symbolize positive feelings, whereas, swirling lines and rough textures symbolized negative attitudes. Students chose earth tones to represent poverty.

In the drawings that had "separation" as the central idea, students emphasized it by visually separating the two figures in the image with two colors in contrasting hues (cold-warm) or values (light-dark) for the background, or for the circle in the top center of the image, or on the ring between the figures. Visual separation of the picture plane was achieved in a few cases by drawing a straight or zigzag line vertically to divide the surface in half.

In general, excluding the drawings where speech bubbles were added, the narrator could not be visually or linguistically identified. Partial synergy between modes regarding the social dimension of students' multimodal texts was sometimes achieved when students gave voice to objects, such as the cloud or a tree, in their written text, while also visually personifying these objects by adding facial features to them with relevant expressions.

c.   Organizational dimension of meaning in students' multimodal texts

As far as the organizational dimension in the linguistic mode is concerned, most of the texts were written in prose. A few of them had a clear dialogic form, whereas some others combined dialogic form in direct speech and prose, writing in continuous speech. In only five cases was there an attempt to write poetry, in either rhyming couplets, such as the passage cited above that adopts Homeric elements, or free verse in stanzas, or rhyme without versification.

Apart from a very few exceptions, the style in third person narration was relevant to the style of story-telling, and often the students chose to start their texts with the formulaic fairy-

tale opening phrase 'Once upon a time' to briefly narrate the whole story. The vocabulary was usually simple without particular adjectives or sophisticated phrases. In dialogic parts, a colloquial style was adopted.

The vast majority of drawings depicted either the initial scene or the final scene of the story. Only one drawing depicted both the beginning and the end. Two different drawings depicted mainly the beginning and part of the evolution of events. However, there was one drawing (see Figure 3) where the complete story evolves visually throughout the surface. Also, a similar visual flow and unity of elements was achieved by placing the written words in between the existing visual elements of the outline. In this way, a vertically positioned semi-circular axis is created, enriching the organizational dimension.

Figure 3. Visual flow and unity of elements in students' multimodal text

In most multimodal texts, student groups either limited their written text to a verbal description of the image or did not include important visual elements in the linguistic mode. Moreover, some groups tried to explain directly what the visual mode described and did not try to compose a parallel linguistic story in order to create meaningful interplay between modes. Analysis of students' multimodal texts and their reflections on how the linguistic and visual elements were interconnected to support the meanings revealed that there was a lack of synergy across modes. Even though the researchers in phase one analyzed and distributed a matrix in an attempt to show the importance of the modal interdependence layering of meaning, students faced difficulties in applying it in the production of their multimodal texts.

In several cases, students' choices were random and fragmented, indicating challenges in attributing synergy of the three aspects of meaning across the two modes.

Analysis of students' reflections regarding the choices made in the production of their multimodal texts (phase two) demonstrated that students could not adequately justify the choices they made to produce meaning in each semiotic system. Moreover, their main difficulty was in explaining how they interconnected the modes in the process of multimodal meaning-making.

Conclusions and suggestions for further research

This study explored multimodal text production by groups in a local school context. The process was initiated by an invitation to the researchers from a specific secondary school to provide students with guidance for participating in a competition organized by the Ministry of Education. External time limitations affected the design of this study and raised important issues regarding the process of scaffolding students to develop necessary skills for multimodal meaning-making. The study allowed an initial investigation of the challenges that exist in producing a multimodal text, which reveal useful directions for further research that will be discussed below.

Production of multimodality is a multilayered approach, since it requires understanding of: a) how the visual mode leads to incorporating multiple paths for orchestrating meanings bounded by image properties; b) how each mode emerges in the moment as the group members collaborate; c) how cross-checking meanings across modes is achieved; and, d) how students engage in self-correcting, when meanings are not working together. Our study indicated that students' approach to meaning-making was fragmented and often one-dimensional. The fact that we asked students, in a limited time frame, to start their production from an open image outline brought to the fore difficulties caused by the ambiguity of wordless images. Viewers could be given more time to make sense of multiple indefinite meanings and choose a narrative path among many. Assisting students to avoid limiting the text to a verbal description of the image and to write a creative story and helping them to introduce all the essential visual elements in their linguistic text, by making the appropriate synergies between the modes, students should engage in long-term processes of reflection, dialogue and justification of choices in multimodal meaning-making.

Admittedly, the limited time frame of the study did not allow for slow and engaged looking, as suggested by Tishman, which would enable students to pause and 'uncover complexities that can't be grasped in a quick glance' (2018: 141). This is a practice of vital importance for constructive multimodal text production.

Working in groups for a limited amount of time seemed to cause additional difficulties for students' collaborative production of multimodal texts. Transforming an image outline into a

multimodal text is in itself, as mentioned above, an open process, which especially welcomes multiple perspectives. This process becomes even more open when the producer engages in dialogue with others. If we would redesign this action, we would provide more time for this 'interpersonal function' (Kalantzis and Cope 2015: 19) of meaning-making, which involves students engaging in dialogue with themselves and others in relation to the visual and verbal texts in progress.

Reflecting on the production of multimodal texts from a researcher's perspective could be further studied by recording and analyzing the process of creating multimodal texts. For example, comprehending a visual mode involves decoding image properties (salience, positioning, use of color, vectors), that is, how one reads and reacts to an image. When images are used, as in this case study, video data of viewers' interactions with the image, while they dialogically transact with it, might be important to provide insights into how individual or group meanings are created.

As Lysaker (2019) suggests, in order to experience how readers voice meanings with the special symbolic resource of image, it is necessary to sit by readers and observe all that happens with their eyes, hands, bodies, and words. The fact that such data were not captured and analyzed is one of the limitations of this study. Group members' gaze as they interacted with the image, their facial expression, and their deictic gestures could provide deep insights regarding the process of meaning-making and its communication among group members. Assessing the relative importance of modes in an interaction requires rich context data to provide information for the intensity, complexity and density of modal use. If a mode is intense in a particular interaction, meaning-making is attached to it. Therefore, an interaction has modal complexity when modes are 'intricately intertwined' (Norris 2004: 79).

Multimodal pedagogy requires long-term, scaffolded, and purposeful interdisciplinary synergies and projects among central and active agents in educational contexts (educators, students, academics, and policymakers). We share the opinion of Gardiner, Cumming-Potvin, and Hesterman that the 'simplistic way multimodality is collaged into the curriculum' (2013: 357) underpins teachers' and students' understanding of the complexity and depth of the multimodality phenomenon. We argue that engaged educators and students can reflect and elaborate on the new possibilities of meaning-making utilizing different semiotic modes, especially, in an interdisciplinary level, which would demonstrate clearly the synergies of multiple modules (e.g., Modern Greek Language, Art, Technology) in students' exploration of modal affordances. Students' literacy and educators' professional development regarding multimodality should be part of an enculturation process through participatory experiences, required by educational policies, curriculum, educators' and academics'[5] initiatives, as it was with this case study.

NOTES

1 The five questions about meaning posed by the two researchers regarding multimodal grammar are quoted from Kalantzis' and Cope's 2015 paper. They originally addressed them in detail in their paper published in the *International Journal of Learning* (Cope and Kalantzis 2009: 365-422).

2 We do not discuss students' movement or sound productions in this paper, since no license was provided for audio and video recordings.

3 We thank the publishing house Diaplasi for allowing us to make the outline of an illustration by Iris Samartzi in *Erotokritos* 2015 and to use it for educational and research purposes.

4 The illustrations by Iris Samartzi in the specific picturebook follow the Renaissance motifs and colors. The clothes of the protagonsts give information about the time and the social status to which they belong. In general, her illustration shows the concern of an artist to study the historical period of the original book and to transmit certain cultural elements of the time in her work. The students could analyze some of these aspects even in the black and white outline. For more information about Samartzi's illustrations in *Erotokritos*, see Rodosthenous-Balafa 2017: 106-119.

5 Academics play an important role in the constructive practice of New Literacies in schools. Especially through research programs with which they are engaged. An example is *Dialls*, Dialogue and Argumentation for Cultural Literacy Learning in School (2018-2021), where through multimodal texts students are engaged in discussions, which enable a growing awareness of their own culture but also of the other's culture. This research program is applied to Pre-Primary, Primary and Secondary Education. Ten academic partners are involved (including the University of Nicosia). The leading University of the project is Cambridge University, see https://dialls2020.eu/.

REFERENCES

Adami, Elisabetta 2017. Multimodality. In: Ofelia García, Nelson Flores and Massimiliano Spotti (eds.) *The Oxford Handbook of Language and Society*. Oxford: Oxford University Press, 451-472.

Albers, Peggy 2007. Visual discourse analysis: An introduction to the analysis of school- generated visual text. In: Deborah Wells Rowe, Robert T. Jimenez, Donald L. Compton, David K. Dickinson, Youb Kim, Kevin M. Leander and Victoria J. Risko (eds.) *Fifty Sixth Yearbook of the National Reading Conference,* vol. 56, Oak Creek, WI: National Reading Conference.

Berger, John 1972. *Ways of Seeing*. New York: Penguin.

Bezemer, Jeff and Jewitt, Carey 2009. Social Semiotics. In: Jan-Ola Östman and Jef Verschuer-

en (eds.) *Handbook of Pragmatics*. Amsterdam: John Benjamins. Also available online https://doi.org/10.1075/hop.13.soc5

Cope, Bill and Mary Kalantzis. 2009. A Grammar of Multimodality. *The International Journal of Learning,* 16 (2): 361-425.

Harste, Jerome C. 2003. What do we mean by literacy now. *Voices in the Middle*, 10 (3): 8-12.

Freedman, Kerry 2003. *Teaching Visual Culture*. New York, NY: Teachers College Press.

Jewitt, Carey 2008. Multimodality and Literacy in School Classrooms. *Review of Research in Education* 32: 241-267.

Jewitt, Carey, Josephus Johannes Bezemer and Kay L. O'Halloran. 2016. *Introducing Multimodality*. London: Routledge. Available from: http://search.ebscohost.com/login.aspx?direct=true&AuthType=ip,sso&db=nlebk&AN=1204264&site=ehost-live [accessed 25 April 2019].

Gardiner, Veronica, Wendy Cumming-Potvin and Sandra Hesterman 2013. Professional learning in a scaffolded 'multiliteracies book club': Transforming primary teacher participation. *Issues in Educational Research* 23 (3): 357-374.

Kalantzis, Mary and Bill Cope 2015. Regimes of Literacy. In: Mary Hamilton, Rachel Hayden, Kathryn Hibbert and Roz Stooke (eds.) *Negotiating Spaces for Literacy Learning: Multimodality and Governmentality*. London: Bloomsbury, 15-24.

Kress, Gunther 2003. *Literacy in the New Media Age*. London: Routledge.

Lancaster, Lesley 2001. Staring at the page: The functions of gaze in a young child's interpretation of symbolic forms. *Journal of Early Childhood Literacy,* 1 (2): 131-152.

Leeuwen, van Theo 2015. Multimodality in Education: Some directions and some questions. *Tesol Quarterly* 49 (3): 582-589.

Lysaker, T. Judith 2019. *Before words. Wordless picture books and the development of reading young children*. New York: Teachers College Press Columbia University.

Mouriki, Katerina and Ioanna Kyritsi-Tzioti 2015. Erotokritos [IN GREEK]. Illustr. Iris Samartzi, Mus. Nikos G. Papadogiorgos, Athens: Diaplasi.

Rodosthenous-Balafa, Marina 2017. 'Her face was the moon': Poetic imagery and illustration in *Erotokritos*'s latest picture book version. In: Rodosthenous-Balafa (ed) *Creative Teaching of Literature. Suggestions and Practices in Primary Education*. Athens: Grigoris, 106-119. [IN GREEK]

Norris, Sigrid 2004. *Analyzing Multimodal Interaction: A Methodological Framework.* New York, NY: Routledge.

Serafini, Frank 2014. *Reading the visual. An introduction to teaching multimodal literacy*. New York: Teachers College Press.

Sipe, R. Lawrence 1998. How Picture Books Work: A Semiotically Framed Theory of Text Picture Relationships. *Children's Literature in Education* 29 (2): 97-108.

Tishman, Shari 2018. *Slow Looking. The Art and Practice of Learning Through Observation.* New York, NY: Routledge.

Marina Rodosthenous-Balafa, Associate Professor in Modern Greek Literature and Teaching Literature, Education Department, University of Nicosia, Cyprus

    E-Mail: rodosthenous.m@unic.ac.cy

Agni Stylianou-Georgiou, Associate Professor in Educational Psychology, Education Department, University of Nicosia, Cyprus

    E-Mail: stylianou.a@unic.ac.cy

Elisavet Pitri, Assistant Professor in Teaching Art, Education Department, University of Nicosia, Cyprus

    E-Mail: pitri.e@unic.ac.cy