Training, Language and Culture
Volume 9 Issue 3, 2025, pp. 113–125
rudn.tlcjournal.org
doi: 10.22363/2521-442X-2025-9-3-113-125

Original Research

Cognitive processing of educational polycode text: An experimental eye-tracking study

by Inna V. Tubalova[1] and Alena V. Garina[1]

[1]Tomsk State University, Tomsk, Russia

This study investigates the cognitive processing of educational PCTs by Russian-speaking readers, addressing the research gap concerning how this process is executed, regulated, and influenced by various factors. The primary aim is to explore the principles of text-image integration during the reading of an educational PCT by native Russian-speaking readers. Employing the concept that eye movements correspond to the logic of cognitive processing, the study uses eye-tracking methods to analyse this process. It relies on eye-movement data from 67 Russian students in the socio-humanities who read an educational polycode biology text in Russian. Findings reveal that the verbal component plays a dominant role in the cognitive processing of this text type, suggesting that this dominance is not dependent on the national educational culture. At the same time, the reading pattern is influenced by the learners' experience with educational texts of a specific content type — in this case, biology. The study also confirms and identifies the mechanisms of the regulatory function of the verbal component during the reading of an educational PCT. The implications of this study are both theoretical and practical: it contributes to the understanding of how the cognitive processing of educational texts is carried out and what factors influence it and underscores the importance of the mindful use of images in the design of educational materials. Moreover, its findings can be applied in developing principles for teaching visual literacy.

KEYWORDS: *educational text, PCT, polycode text, eye-tracking, native readers, Russian, cognitive processing*

CRediT AUTHOR STATEMENT: Inna V. Tubalova: Conceptualisation, Methodology, Supervision, Project administration, Funding acquisition, Writing — Original Draft. Alena V. Garina: Methodology, Formal analysis, Investigation, Validation, Data Curation, Writing — Review & Editing.

CONFLICT OF INTEREST: The authors declared no conflict of interest.

DATA AVAILABILITY STATEMENT: The data supporting this study's findings are available from the corresponding author, Alena V. Garina, at alyonushka050700@gmail.com, upon reasonable request.

FUNDING: This research was supported by Ministry of Science and Higher Education of the Russian Federation, project No FSWM-2025-0016.

ARTICLE HISTORY: Submitted June 30, 2025 | Revised August 18, 2025 | Accepted September 12, 2025

FOR CITATION: Tubalova, I. V., & Garina, A. V. (2025). Cognitive processing of educational polycode text: An experimental eye-tracking study. *Training, Language and Culture, 9*(3), 113–125. https://doi.org/10.22363/2521-442X-2025-9-3-113-125

1. INTRODUCTION

The study of the polycode text (PCT), including the educational PCT, has been an active area of research since the end of the 20th century. Recently, this research area has gained new momentum due to the high demand for information presented in the form of units from different semiotic systems and transmitted through various channels, which characterises both the modern reader and the modern learner. The purpose of this study is to identify the principles of semantic integration of the verbal component and the image during the reading of an educational PCT in the native language by Russian-speaking readers.

An analysis of the eye-movement activity of readers of an educational anatomy PCT, with the verbal component in Russian, made it possible to reveal its reading patterns and the reader orientation mechanisms, which are determined by the verbal component and aimed at establishing a connection with the pictorial component. The novelty of the results obtained is primarily determined by the focus on educational text material in the Russian language. In addition, this study for the first time reveals the principles underlying the regulatory function of the verbal component of an educational PCT in Russian during its perception by Russian-speaking readers.

Volume 9 Issue 3, 2025, pp. 113–125 doi: 10.22363/2521-442X-2025-9-3-113-125

'A special role in obtaining significant information about the processes of PCT perception is played by eye-tracking data, the interpretation of which makes it possible to analyse the sequence of processing different elements of textual information and the proportion of attention allocated to them. Studies aimed at identifying patterns of cognitive processing of PCTs with a verbal component in Russian are currently scarce, and eye-tracking data on the reading processes of Russian speakers engaging with educational PCTs in their native language have been scarcely addressed by researchers'

2. THEORETICAL BACKGROUND

The concept of the PCT aligns with the idea of non-linearity and openness in text organisation, the origins of which trace back to the works of Bakhtin (1986), Barthes (2004), the French school of discourse analysis (Lacan, 1981; Pêcheux et al., 1995; Sériot, 1997; Foucault, 1971), post-structuralist text studies (Derrida, 2016; Kristeva, 2009), critical discourse analysis (Wodak, 2004; van Dijk, 2015; Fairclough, 2023), and others. Within this logic, the formation of the text's content structure is not limited to its verbal component alone but is ensured by the conditions of its functioning.

The opposition of the PCT to the monocode text emerged in linguistics in the 1970s. As a result, researchers argue that *'PCTs in a broad semiotic sense should include cases of combining the natural language code with the code of any other semiotic system (image, music, etc.)'* (Eiger & Yukht, 1974, p. 107). Consequently, within the concept of text non-linearity, a specific aspect arises: the non-linearity of a text's content organisation is determined by the interaction of its verbal component with units of other code systems, which aligns with Paivio's (1986) Dual Coding Theory.

It should be noted that the heterogeneity of code units involved in organising the text's content can be facilitated by a single mode of its transmission or by multiple modes – as a *'socially and culturally conditioned semiotic resource for making meaning'* (Kress, 2009). Forceville (2006) connects them with the five senses, which provide channels for information transmission: (1) the pictorial or visual mode; (2) the aural or sonic mode; (3) the olfactory mode; (4) the gustatory mode; and (5) the tactile mode. Many scholars differentiate the modes of text transmission in a similar way (Kibrik, 2010; van Leeuwen, 2005; Schnotz & Horz, 2009; Lemke, 1998; Jewitt, 2009; O'Halloran, 2004; Machin, 2007; Scollon & Scollon, 2003).

Thus, researchers fundamentally distinguish between the codes of text organisation and the modalities of its transmission: *'If a PCT is one that combines different semiotic codes, then a multimodal text can be called a text that is perceived using different modalities – channels of information perception'* (Nekrasova, 2014, p. 45).

The present study focuses on the specific nature of perceiving a polycode monomodal text, organised based on the interaction of verbal and pictorial codes transmitted through the visual modality. This type of text has been studied in greatest detail both in terms of its internal organisation, which determines the formal-semantic coherence of its heterogeneous components and the communicative unity of their functioning (Anisimova, 2003; Bernatskaya, 2000; Valgina, 2003; de Beaugrande & Dressler, 1981; Halliday & Hasan, 1976; Levin, 1981; Levin et al., 1987; Chernyavskaya, 2009), and in terms of its perception (Kunitsyna, 2010; Sonin, 2005; Sorokin & Tarasov, 1990; Mason et al., 2013; Rayner et al., 2001).

A special role in obtaining significant information about the processes of PCT perception is played by eye-tracking data, the interpretation of which makes it possible to analyse the sequence of processing different elements of textual information and the proportion of attention allocated to them. Studies aimed at identifying patterns of cognitive processing of PCTs with a verbal component in Russian are currently scarce (Blinova & Shcherbakova, 2020), and eye-tracking data on the reading processes of Russian speakers engaging with educational PCTs in their native language have been scarcely addressed by researchers.

Studies on the processes and results of PCT perception, in turn, are based on developed theories of reading, for instance, the concept by Hoover and Gough (1990), who consider decoding and comprehension as two main stages in reading comprehension. Decoding is viewed as a basic process of word recognition that provides access to the mental lexicon and the retrieval of a word's semantic information. Comprehension is a higher-level process that involves integrating words and constructing a mental model of the text. Jian (2013) and Lee and Wu (2017) suggest that this reading model can be applied to the analysis of a PCT that includes verbal and pictorial components.

The present study is aimed at identifying the specific features of cognitive processing of educational PCTs. The results of examining the specifics of their reading may prove useful in the development of educational practices, primarily in addressing the problem of effective acquisition of educational information presented in the PCT format (Levin, 1981; Levin et al., 1987; Mayer & Gallini, 1990; Mayer, 2005, 2009; Schnotz, 2005).

Solving this problem relies on theoretical conclusions drawn from psycholinguistic studies of cognitive processing of PCTs by readers.

The most actively developed issue concerns the distribution of significance between the verbal component and the image in PCT perception. Several previous studies investigating the reading of text–picture combinations have found that reading behaviour appears to be largely directed toward the verbal component (Hegarty & Just, 1993; Rayner et al., 2001; Kosenko, 2019).

This conclusion has been reached both in the study of educational texts (Hegarty & Just, 1993; Rayner et al., 2001) and texts of other types – for example, advertising (Rayner et al.,

'Studies on the processes and results of PCT perception, in turn, are based on developed theories of reading, for instance, the concept by Hoover and Gough (1990), who consider decoding and comprehension as two main stages in reading comprehension. Decoding is viewed as a basic process of word recognition that provides access to the mental lexicon and the retrieval of a word's semantic information. Comprehension is a higher-level process that involves integrating words and constructing a mental model of the text. Jian (2013) and Lee and Wu (2017) suggest that this reading model can be applied to the analysis of a PCT that includes verbal and pictorial components'

2001) or artwork descriptions (Kosenko, 2019). These conclusions were based on various data, the main ones being the following.

1. Data on the sequence of processing the verbal and pictorial components of an educational PCT, reflecting its reading strategies corresponding to eye-movement patterns, allowed for the identification of three main strategies: (a) text-first with references to the image (readers process the verbal part of the PCT, periodically referring to the image); (b) text-first (readers turn to the illustration only after processing the verbal component); and (c) image-first (readers turn to the verbal component only after processing the illustration) (Hannus & Hyönä, 1999; Hegarty, 1992; Hegarty & Just, 1993; Jian, 2015, 2016; Jian & Wu, 2015). The dominance of the verbal component in this case is evidenced by the priority use of strategies (a) and (b).

2. Data on the distribution of visual attention indicated that the proportion of time spent by most readers processing the verbal component of the educational PCT out of the total processing time was greater than that spent processing the image (Hannus & Hyönä, 1999; Jian, 2015).

The proposition concerning the dominance of the verbal component of PCTs, including in the reading of educational texts, is not disputed by researchers, and in the present study we anticipated its confirmation.

At the same time, the question of the mechanisms underlying the integration of the content of heterogeneous components of educational PCTs, as well as the factors determining these mechanisms, remains unresolved.

A number of studies analysing the nature of reading educational PCTs on the basis of eye-tracking data take into account the specifics of their subject domain, although the question of whether reading strategies for educational PCTs depend on the subject domain has been explicitly raised in only a few studies. For example, Jian (2015) draws attention to the importance of visual literacy in extracting meaning from a PCT, which, according to the author, does not depend on its subject area but is determined by general reading and writing practices. In contrast, Lee and Wu (2017), who study the features of reading geometry texts, believe that in terms of the distribution of visual attention between the components of a PCT, geometric reading is different and is largely focused on the figure. Although they, like their predecessors, conclude that the reading pattern for illustrated geometric descriptions is verbally-oriented (most readers first looked at the text and then studied the corresponding elements in the figure), they emphasise that fixations on geometric figures took up a significant proportion of the participants' time – almost 40%, since geometry is the study of shape and space.

In substantiating the mechanisms of integrating heterogeneous components of PCTs, researchers identify readers' prior experience with PCTs of a certain type as one of the significant factors. Such experience determines the anticipated outcome of perception (Jian, 2015; Lee & Wu, 2017; Zlokazov & Lipnitsky, 2018; Kosenko, 2019; Sonin, 2005; Vashunina et al., 2020). An educational PCT may be considered 'typical' or 'familiar to readers' and therefore processed in a standard way, provided that it corresponds to the principle of organising heterogeneous components characteristic of a given subject domain. For example, Jian and Wu (2015) characterise the images included in their stimulus materials, taking into account their specific features. Drawing on Kress and van Leeuwen (2020), they point out that such illustrations are used to represent part-whole relations in the structure of an object, to indicate the names of individual components of the object, and to display their spatial arrangement. As an example, they cite an illustration of the internal structure of a cell, including the nucleus, the cell membrane, and so forth. Importantly, such illustrations, first, correspond to the semantic and spatial representation of the object and, second, according to Slough et al. (2010), are the most frequently used type in science textbooks. Since educational texts with a verbal component in Russian and their processing by Russian-speaking readers have not previously attracted scholarly attention, it is particularly important in their analysis to consider the typical ways in which information is presented in Russian textbooks.

Differences in the cognitive processing of PCTs across subject domains, revealed through eye-tracking data, are observed in two respects.

First, while not disputing the dominant role of the verbal component in processing educational PCTs, some researchers differentiate the degree of its significance when reading texts from various subject areas. Thus, Jian and Wu (2012), Lin and Lin (2014), and Lee and Wu (2017) noted that in reading geometry texts, fixation durations on figures were relatively high.

Second, according to researchers, data on the cognitive processing of educational PCTs from different subject domains differ with respect to the presence or absence of preferred reading strategies. For example, Hegarty (1992) and Hegarty and Just (1993) documented the absolute priority of the 'text-with-reference-to-picture' strategy when reading PCTs in mechanics. By contrast, Jian and Wu (2012, 2015) found that, when working with materials in biology, a substantial number of readers employed the 'text first' or 'picture first' strategies. They compared

Volume 9 Issue 3, 2025, pp. 113–125 doi: 10.22363/2521-442X-2025-9-3-113-125

their findings with earlier studies and suggested that one possible explanation lies in differences in the content: materials on mechanical systems can only be understood by relying on images that capture the spatial relations between elements, whereas materials in biology allow readers to extract information about spatial relations from the verbal component of the PCT. Therefore, the 'text-with-reference-to-picture' strategy is not necessarily dominant for readers of biology texts (Jian & Wu, 2015). Later, however, Lee and Wu (2017) confirmed the absolute priority of the 'text-with-reference-to-picture' strategy when reading geometry PCTs.

Thus, previous studies have provided important observations on how the mechanism of integrating heterogeneous components of educational PCTs is implemented and what it depends on.

However, no definitive conclusions have been reached concerning the specific factors that influence these mechanisms (e.g., particular subject domains, prior experience with them, or the thematic focus of the material).

Since visual perception processes in reading are largely unconscious and not subject to self-control, it seems that valuable insights into this issue can be gained from examining the moments in which readers shift their visual attention from the verbal component of the PCT to the image – that is, when integration occurs. This question can be posed within the framework of analysing the 'text-with-reference-to-picture' reading strategy. In studies focused on educational PCTs in biology, the verbal component has not been considered in this respect. In analysing a geometry text, however, Lee and Wu (2017, p. 711-712) noted that *'the timing of transferring to the figure was the reference to new geometric elements in the text'*, and readers turned to the figure in order *'to make sure that the spatial relation of geometric elements was in accord with the spatial image formed by reading the text'.*

In our study, we intended to devote special attention to those units of the verbal component of PCTs that direct the reader to the image.

3. MATERIAL AND METHODS

3.1. Research questions and hypotheses

The first question addressed in this study is whether the reading process of Russian-speaking learners of a PCT, typical of Russian biology textbooks, is verbally- or illustration-oriented. Most previous research has found that reading of educational illustrated texts is verbally directed. However, we assume that Russian learners may have a different experience with such texts. Previous studies did not focus on the specifics of educational experience.

In differentiating reading patterns, we used the methodology of Jian and Wu (2015). If a reader initially scans the illustration, fixating at least three times, this movement indicates an image-oriented reading pattern. In this case, we can conclude that the process of examining the figure stimulates the subsequent process of reading the text. Otherwise, we can conclude that the reading of illustrated PCTs remains text oriented. To answer this question, we used eye-tracking technology to record the readers' eye-movement patterns.

Furthermore, to identify the dominant role of the verbal component or the illustration, we considered an eye-movement index such as the mean processing time of the verbal component and the image (the ratio of the total processing time of one of the heterogeneous components of the PCT for all respondents to the number of respondents).

The second research question relates to the specifics of processing an educational text in which the verbal and pictorial components, placed in the same frame, may or may not have a semantic link. The formulation of this question is determined by the fact that in textbooks, an image is often placed next to a verbal description that does not correspond to it, which can cause additional difficulties for learners. In this regard, we posed the question of how the processing of the frame changes if the image does not semantically correspond to the verbal description. We determined whether the mean fixation duration on the verbal component and the image differs when processing frames of the two designated types and calculated their level of processing difficulty using the method of Anisimov and Luzhin (2020).

Answering this question also allowed us to make assumptions about the verbal mechanisms for regulating the reading process of an educational PCT. First, we hypothesised that the verbal component contains certain typical elements that trigger a reference to the image. We designated them as verbal predictors of reference to the image. If the hypothesis that there are verbal signals determining the reader's attention shift to the image is confirmed, it can be argued that the verbal component of an educational PCT regulates the reader's activity of integrating the meaning of its heterogeneous components, and that the regulatory mechanism is set by the organisation of the verbal component.

Second, we hypothesised that the significance of the regulatory function of the verbal component of the PCT is so high that even in the absence of a semantic link between the verbal component and the image, the verbal predictors will continue to perform their function. An alternative hypothesis was that, upon detecting a semantic mismatch between the verbal component and the image in the early stages of reading, respondents might refuse to refer to the image.

To identify the verbal predictors of reference to the image, we analysed the composition of the fixation zones from which there are progressive saccades to the image and identified the average number of saccades for each such fixation (the ratio of the total number of fixations for each zone to the number of respondents).

Third, we analysed the nature of the verbal units that most regularly referred to the image and attempted to determine whether their content is typical and how it relates to the subject matter being studied. If the content proves to be subject-specific, this will confirm the idea put forward by earlier researchers that

'Since visual perception processes in reading are largely unconscious and not subject to self-control, it seems that valuable insights into this issue can be gained from examining the moments in which readers shift their visual attention from the verbal component of the PCT to the image – that is, when integration occurs. This question can be posed within the framework of analysing the 'text-with-reference-to-picture' reading strategy. In studies focused on educational PCTs in biology, the verbal component has not been considered in this respect'

the nature of educational PCT processing depends on the subject content. Additional insights into these questions may be provided by comparing the data obtained from processing frames with images that semantically correspond and do not correspond to the verbal description: do the verbal predictors identified in the first and second cases coincide?

3.2. Approach and research tools

Due to the lack of reference to contrastive material – educational texts from other subject areas – we did not plan to draw significant conclusions about cognitive processing of PCT in biology. However, in interpreting the results, we used the conclusions obtained in the aforementioned studies, i.e. in mechanics (Hegarty, 1992; Hegarty & Just, 1993), natural sciences (Jian, 2013; Jian & Wu, 2015), and geometry (Lee & Wu, 2017).

At the same time, we paid special attention to the features of the educational text, as it seems to us that the specifics of its functioning and the reader's experience of its perception, acquired in the practices of participating in educational discourse, influenced the resolution of the questions posed.

Within this logic, an educational text is defined as a speech product that *'serves as a means for the primary and secondary textual activities of learners and is understood as methodological and didactic educational material'* (Kuryanovich, 2022, p. 55). Researchers note that the purpose of an educational text is to organise the activities of both the teacher and the student, ensuring the systematisation of the content of the studied discipline (Babakova, 2022, p. 42). First and foremost, an educational text is intended to convey the fundamentals of scientific knowledge corresponding to the curriculum. It is an important source of knowledge for students and a guide for teachers (Tyutkova, 2016, p. 122), reflecting the system of knowledge at a given level of education. Significant features of the educational text for this study are *'(1) the ability to be a carrier of general and professional knowledge'; (2) an orientation towards unambiguous perception and construction according to the laws of logical thinking; (3) a special principle of organisation, suggesting that 'such texts have the same set of informative units (semantic blocks), the same set of lexical-grammatical units that realise these meanings'* (Sabinina, 2009, p. 223).

The educational text used by us belongs to texts that have a particular complexity of informational content and a significant multiplicity of information details that the recipient must assimilate.

Given that in most of the studies our work is based on, the type of PCT under consideration is termed an *illustrated scientific text*, we will pay special attention to the relationship between the concepts of *scientific text* and *educational text*.

The use of the term *scientific text* indicates the type of content it represents, reflecting the results of scientific cognition. This type of content implies its presentation in accordance with the laws of logical thinking and an orientation towards unambiguous perception – that is, it has a number of features discussed by researchers when characterising an educational text. At the same time, an educational text has distinctions significant for our research. In our study, we designate the object of the investigated cognitive processing as an *illustrated educational text*, thereby emphasising the purpose and sphere of its use – in educational discourse, and specifically, in a textbook. Within educational discourse, a scientific text possessing these features can be used as an educational text (van Dijk, 1997; Sweller, 1994).

In this regard, the purpose of presenting an educational text – as a product of educational discourse – is the dissemination of knowledge that is new to the learner. Text-based learning is the addition of new information to the reader's background knowledge. The discursive conditions of using an educational text presuppose specifics in its presentation and perception. For this study, particular significance is attached to such conditions of its presentation as a special substantive selection of material corresponding to the reader's background knowledge, and the manner of its presentation, which reflects the typical features of educational experience. The placement of a text in a textbook, for example, presupposes the assimilation of its content by the target group of users of that textbook, and its author selects the material accordingly, arranging it in accordance with the typical attitudes of the educational culture, in particular, with the attitudes of interest to us regarding the use of a PCT that includes a verbal component and an image. Thus, in Russian literature textbooks, it is common to illustrate information presented in verbal form with portraits of writers, reproductions of paintings based on literary works, and so on; in textbooks on technical disciplines – with diagrams of mechanisms' operation, and so on.

Special principles for using PCT also exist in biology textbooks, the material of which is used as a stimulus in the present study. Typical for the educational culture is illustrating the characteristics of a biological object with an image of its structure (we will focus on this in more detail when characterising the stimulus material).

The reader's mindset, according to researchers (Luria, 2009; Shelestyuk, 2010), also plays an important role in the process of text perception. The perception of an educational text also has certain features, and for this study, the readers' mindset for subsequent monitoring of the assimilation of the information

Volume 9 Issue 3, 2025, pp. 113–125 doi: 10.22363/2521-442X-2025-9-3-113-125

received is important. This mindset is conditioned by a discursive goal – the focus of educational discourse on the dissemination of information, which for the learner as its participant is expressed in the acquisition of new knowledge. The novelty of the information for the recipient and the detail of its presentation is a discursive condition for the functional significance of an educational text: *'the meaning of the main text is the quintessence of teaching a subject, the basis of the professional properties that we wish to instil in the student to realise the competency-based paradigm of education'* (Kalugin & Prokhorov, 2023).

We took these features of the educational text into account when analysing its cognitive processing during reading.

We used these features of the educational text when interpreting the results of the experiment that recorded the eye movements of a reader of an educational PCT. In its design and implementation, we relied on the extensive tradition of eye-tracking studies of reading, described, for example, in Rayner (1998). Most of these studies were conducted in the so-called 'default' reading mode, that is, '*when comprehension occurs without difficulty and the eyes continue to move forward along the line of text'* (Reichle et al., 2009, p. 9), i.e., with a relatively small number of regressions. This reading mode is certainly used by some respondents at certain stages of reading an educational PCT. However, eye behaviour when perceiving the verbal component and the image during learning activities can differ significantly from that in 'default' reading modes, and it is precisely for this reason that many researchers of PCT reading raise the question of the sequence of cognitive processing of the heterogeneous components of an educational PCT, which form its various reading patterns. Moreover, most eye-tracking studies have examined English-speaking readers, not the Russian-speaking readers discussed in this article, whose behaviour may differ significantly.

An NTrend-ET500 was used as the eye-tracker, which has no chin rests, helmets, or other distracting factors, ensuring the most natural user behaviour possible. The scanning frequency can be varied in the range of 60 to 520 Hz. Additionally, the device was equipped with binocular tracking (rather than single-eye tracking) and a face camera.

3.3. Participants

The subjects were students of socio-humanitarian faculties of Tomsk State University (Tomsk, Russia), native Russian speakers (n = 67). The subjects were divided into 2 groups, each of which was presented with a separate stimulus. One group was the main group (stimulus 1), and the second was the control group (stimulus 2). The main group included 45 participants, and the control group included 25. Data on the eye movements of three participants from the main group and two from the control group were excluded due to technical flaws in their recording. The subjects read on-screen text consisting of verbal and pictorial components.

All participants signed the relevant ethical and personal information consent forms.

3.4. Stimulus material

The experimental material was an illustrated educational text, adapted from a biology textbook for medical students used in Russia (Sapin & Bilich, 2010, p. 98). The topic of the text was the structure and functions of human skin. The text consisted of 254 Russian words. For stimulus 1, it was accompanied by an illustration from this textbook that semantically corresponded to it (Table 1), and for stimulus 2, by an illustration from the same textbook (Sapin & Bilich, 2010, p. 99) that did not semantically correspond to it (Table 2). The illustration for stimulus 2 was chosen so that its external form corresponded as much as possible to the authentic illustration: it was created in the same visual style (e.g., colour scheme), also reflected the structure of a human body organ, and did not allow the semantic discrepancy to be detected at a first cursory glance. The illustrations for both stimuli were placed above the text, which could stimulate their processing at the beginning of reading, in accordance with the standard pattern of eye movement on a page.

The text conveyed detailed special information but did not contain fragments whose full understanding would be impossible without reliance on specialised background knowledge; that is, the information contained within them was presented in such a way that it could be understood by a non-specialist in the field. The influence of such prior knowledge had to be excluded to achieve maximum attention from readers to all fragments of the text.

The text consisted of three parts: the first, introductory part provided a general characterisation of the layered structure of human skin; the second and third parts described the structure and functions of each layer. The statements that made up the text, in their content, included (1) a description of the spatial relationships between the internal elements of the skin, as in *Под сосочковым слоем находится сетчатый слой* [Under the papillary layer is the reticular layer]; (2) a description of the external features of its individual elements, as in *Благодаря наличию сосочков на поверхности кожи видны гребешки, разделенные бороздками* [Because of the presence of papillae, scallops separated by grooves are visible on the surface of the skin]; (3) an indication of the functions of these elements, as in *Этот слой играет важную роль в терморегуляции и является жировым депо организма* [This layer plays an important role in thermoregulation and is a fat depot for the body]; (4) information supplementary to the structure and functions of the organ, as in *Изучение деталей рельефа кожи (папиллярных линий и узоров) получило название дерматоглифики* [The study of the details of the skin's relief (papillary lines and patterns) is called dermatoglyphics].

The participants were asked to review the material displayed on the screen and answer some questions about its content. The respondents received the following instruction: *'Before you on the screen you will see a text followed by several questions. Imagine that you are taking a test, and your answers to the questions will be graded. Please read the text carefully and write your answers on the answer sheet'.* The instruction for subsequent

monitoring of the material assimilation allowed the situation model to be brought as close as possible to the educational conditions for mastering an educational PCT. The content of the questions, as well as the results of the answers, were not considered in this study, which was aimed at studying the process (not the result) of mastering an educational PCT.

It is important to note that, despite the absence of professional knowledge among the subjects – students of socio-humanitarian fields – regarding the content of this text, they had significant experience in mastering similarly organised educational PCTs, gained during their school education. An analysis of three Russian school biology textbooks (Dragomilov & Mash, 2008; Kolesov et al., 2016; Rokhlov & Trofimov, 2007) showed that in terms of illustration content (schematic representation of the structure of a human body organ) and the structure of the corresponding verbal component, PCTs of this type account for 58% to 92% of the content in these textbooks.

Based on this, we can assert that our respondents relied not only on perceptual conditions but also on their discursive experience in establishing a connection between the components of the PCT and had a certain image of the predicted result of its perception, the significance of which is actively discussed by researchers (Vashunina et al., 2020).

4. STUDY AND RESULTS

The results of our study showed the presence of all three previously identified strategies for reading an educational PCT, but for most respondents, the priority of the verbal component of the PCT and the dominance of the 'text-first with references to the image' strategy were confirmed. The absolute majority of respondents used this strategy both when processing stimulus 1 (36 out of 42; 86%), where the illustration content corresponded to the verbal component content, and when processing stimulus 2 (15 out of 20; 75%), where the illustration content did not correspond. Among the respondents of both groups, 82% of readers used this strategy.

In addition, 4 respondents from the group reading stimulus 1 and 3 respondents reading stimulus 2 used the 'text-first' strategy, thus also focusing on the verbal component of the PCT. Only 2 respondents from each group used the 'image-first' strategy, meaning they showed at least three fixations on the illustration at the beginning of reading.

Furthermore, the dominance of the verbal component is indicated by data on the average processing time of the verbal component and the image. For both stimulus 1 and stimulus 2, the dwell time in the verbal component zone was significantly predominant (Table 1).

Table 1
Mean dwell time in the zones of the verbal component of the PCT and the image

	TOTAL MEAN READING TIME OF PCT, MS	TOTAL MEAN READING TIME OF VERBAL COMPONENT, MS	TOTAL MEAN READING TIME OF VERBAL COMPONENT, %	TOTAL MEAN READING TIME OF IMAGE, MS	TOTAL MEAN READING TIME OF IMAGE, %
Stimulus 1	83091	78306	94.2	4785	5.8
Stimulus 2	101858	92110	90.4	9748	9.6

To address the question of how educational PCTs are read when the accompanying image either corresponds or does not correspond to the content of the verbal component, we analysed the indicators of the overall mean processing time for stimuli 1 and 2, as well as the indicators of mean subjective complexity during their reading.

The indicator of the overall mean processing time for stimuli 1 and 2 was considered significant because readers were allowed to choose a comfortable pace of reading and were not restricted in the time allotted for working with the PCT. As shown in Table 1, this indicator was considerably higher for stimulus 2.

The indicator of mean subjective text complexity was calculated by dividing the number of fixations followed by regressive saccades by the total number of fixations, where a fixation is defined as the concentration of the reader's gaze on specific areas of the text material (Anisimov & Luzhin, 2020). This indicator was also computed for the processing of stimuli 1 and 2. The results demonstrated that processing stimulus 1 required, on average, a smaller number of regressive fixations (mean = 85), and the coefficient of subjective complexity was lower (14.7094%), whereas processing stimulus 2 involved a larger number of regressive fixations (mean = 147) and showed a higher coefficient of subjective text complexity (23.1864%).

Separately, we analysed the mechanisms regulating the process of reading educational PCTs when employing the strategy 'verbal component with references to the image.' An analysis of fixation zones from which progressive saccades to the image were made revealed that the maximum mean number of saccades (five or more per participant) originated from fixations on text components that corresponded in content to statements of the following types: (1) a description of the spatial relationships between the internal elements of the skin, and (2) a description of the external features of its individual elements. Moreover, the analysis of the verbal component of the PCT demonstrated that, for participants in the control group (stimulus 2 – the image does not correspond to the verbal component),

Volume 9 Issue 3, 2025, pp. 113–125 doi: 10.22363/2521-442X-2025-9-3-113-125

all fragments of this type presented in the text were included as references to the image, while for participants in the experimental group (stimulus 1 – the image corresponds to the verbal component), only three such components were not included (Table 2). The total number of saccades to the image recorded during the processing of stimulus 1 was, on average, 85 saccades per reader, and for stimulus 2, it was 147.

Consequently, the saccades to the image initiated from the identified fixation zones constituted 79% for stimulus 1 and 77% for stimulus 2. Saccades from other fixation zones were isolated (no more than two per zone on average), and therefore can be attributed to random eye movements, and the verbal units forming the corresponding fixation zones cannot be considered typical.

Table 2
Mean (per participant) number of saccades to the image from text components

No.	VERBAL COMPONENT ZONES OF THE PCT	MEAN NUMBER OF SACCADES TO THE IMAGE (STIMULUS 1)	MEAN NUMBER OF SACCADES TO THE IMAGE (STIMULUS 2)
1.	В ней различают сосочковый и сетчатый слои [In it, the papillary and reticular layers are distinguished]	15	20
2.	Сосочковый слой находится под базальной мембраной эпидермиса [The papillary layer is located under the basement membrane of the epidermis]	12	18
3.	Он сформирован рыхлой волокнистой неоформленной соединительной тканью, которая расположена в виде сосочков [It is formed by loose fibrous unformed connective tissue, which is located in the form of papillae]	10	15
4.	Благодаря наличию сосочков на поверхности кожи видны гребешки [Because of the presence of papillae, scallops separated by grooves are visible on the surface of the skin]	8	12
5.	Гребешки, соответствующие возвышениям сосочков дермы... [The ridges, corresponding to the elevations of the dermis papillae...]	7	10
6.	Строение кожного рельефа широко используется для идентификации личности в криминалистике [The structure of the skin relief is widely used for personal identification in forensics]	–/–	8
7.	В сосочковом слое имеются миоциты... [In the papillary layer, there are myocytes...]	–/–	5
8.	Под сосочковым слоем находится сетчатый слой [Under the papillary layer is the reticular layer]	9	10
9.	Наряду с коллагеновыми в сетчатом слое имеется сеть эластических и небольшое количество ретикулиновых волокон [Along with collagen fibers, the reticular layer contains a network of elastic and a small number of reticulin fibers]	6	7
10.	Пучки коллагеновых волокон сетчатого слоя переходят в подкожную основу... [Bundles of collagen fibers of the reticular layer pass into the subcutaneous base...]	–/–	8
	TOTAL	67	113

'As expected, our material confirmed the dominance of the verbal component in reading educational PCTs. The data obtained from readers whose educational experience was shaped within the Russian educational environment do not differ from previously reported findings, which demonstrated the dominance of the verbal component in PCTs: in the process of processing a PCT aimed at mastering its content for subsequent control, readers concentrated on its verbal component. At the same time, our data on reading strategies differ from those obtained by Jian and Wu (2015), who reported that a considerable proportion of readers adopted either 'text-first' or 'image-first' strategies when processing educational PCTs in biology'

5. DISCUSSION

As expected, our material confirmed the dominance of the verbal component in reading educational PCTs. The data obtained from readers whose educational experience was shaped within the Russian educational environment do not differ from previously reported findings, which demonstrated the dominance of the verbal component in PCTs: in the process of processing a PCT aimed at mastering its content for subsequent control, readers concentrated on its verbal component.

At the same time, our data on reading strategies differ from those obtained by Jian and Wu (2015), who reported that a considerable proportion of readers adopted either 'text-first' or 'image-first' strategies when processing educational PCTs in biology.

Since in our stimulus material we employed a typical representation of PCTs from Russian biology textbooks, which corresponded to their typical representation in the textbooks analysed by Lee and Wu (2017), it was not possible to assess the impact of educational background.

In addressing the second research question, based on the average reading time of stimuli 1 and 2 as well as their mean subjective complexity, we confirmed the earlier and rather obvious conclusion that presenting verbal text and a non-matching image in a single frame creates additional difficulties for readers in text processing.

This conclusion gains further importance when interpreted in light of the mechanisms of integration of verbal text and image that we identified. These mechanisms are expressed in the unconscious search for correspondence between components during the detection of verbal predictors.

Analysis of fixation zones from which progressive saccades were directed to the image confirmed the presence of predictors within the verbal component that prompt references to the image. These predictors regulate reading when the 'verbal component with references to image' strategy is applied, determining the readers' intention to establish a link between the heterogeneous components of the PCT. Their composition proved to be stable not only for stimulus 1 but also for stimulus 2. This demonstrates the regulatory function of the verbal component – namely, its role in guiding readers' integration of heterogeneous elements of an educational PCT and the mechanisms by which this regulation is implemented.

We cannot claim that the specific textual fragments we identified will serve as predictors for all PCTs with different subject content. However, for biology PCTs aimed at assimilating knowledge of biological objects, such predictors are clearly manifested. The experience of educational practice directs readers to search in the pictorial component for information about structure and visual features, and this search is triggered by the content of the verbal component. The stable reader needs to turn to the image to confirm information received verbally is realised not spontaneously, but precisely in accordance with the organisation of the verbal component. This indicates not only its dominance but also its regulatory function.

In addressing the issue of whether subject specificity plays a role in the implementation of integration mechanisms of heterogeneous components of educational PCTs, the analysis of the composition of the identified predictors becomes particularly significant.

This content in its representation in Russian is organised by certain verbal units that semantically encode:

(1) the spatial position of a part of the described object relative to other parts: сосочковый и сетчатый слои (papillary and reticular layers), сосочковый слой находится под... (the papillary layer is located under...), на поверхности кожи видны гребешки (ridges are visible on the surface of the skin), в сосочковом слое имеются... (in the papillary layer there are...), под сосочковым слоем находится... (under the papillary layer there is...), в сетчатом слое имеется... (in the reticular layer there are...), пучки коллагеновых волокон сетчатого слоя переходят в подкожную основу (bundles of collagen fibers of the reticular layer pass into the subcutaneous base);

(2) the presence of specific external features of these parts: расположена в виде сосочков (located in the form of papillae), на поверхности кожи видны гребешки (ridges are visible on the surface of the skin), as well as these external features themselves: соответствующие возвышениям сосочков дермы (corresponding to the elevations of the papillae of the dermis). We draw attention to the character of verbal predictors referring to the image. Hegarty and Just (1993) noted that the primary purpose of inspecting a diagram when reading an educational PCT in mechanics is to confirm the spatial relations presented in the verbal description. Similar observations were made by Lee and Wu (2017) when analysing the cognitive processing of geometry texts.

In this study's material, predictors prompting reference to the image include descriptions not only of spatial relations but also of external characteristics of the object, which are likely determined by readers' experience with presenting such information in biology textbooks. This also explains the absence (or rarity) of saccades to the image from verbal fragments that describe

Volume 9 Issue 3, 2025, pp. 113–125 doi: 10.22363/2521-442X-2025-9-3-113-125

'The conducted study contributes to the development of the concept of perception of an educational PCT, and specifically, it reveals the mechanisms by which a reader establishes a connection between its heterogeneous elements. For the first time, such mechanisms of integration were identified. Moreover, this type of study, conducted using the eye-tracking method, was for the first time carried out on data concerning the cognitive processing of an educational PCT in Russian by Russian-speaking readers.
This study confirmed the dominance of the verbal component of the PCT in conditions that orient towards subsequent monitoring of the knowledge obtained. This conclusion corresponds to the results obtained earlier from data on the processes of reading educational PCTs'

spatial relations but do not refer to the biological structure of the object (e.g., *На стопе, локтях, концевых фалангах пальцев, подвергающихся постоянному давлению, ячейки сети широкопетлистые*). Thus, the need to establish a link between verbal and pictorial components of educational PCTs is regulated not only by the content and organisation of the verbal component but also by readers' experience.

Both when a semantic link between heterogeneous PCT elements is present and when it is absent, the overall proportion of saccades to the image from the identified predictors was approximately the same (79% and 77%), and their intensity decreased during reading.

Furthermore, particular interest is raised by the results of comparing the composition of verbal predictors prompting references to the image with the frequency of such references when readers processed stimuli 1 and 2 (Table 2). The number of saccades to the image in the absence of semantic correspondence between components of the PCT was considerably higher than in its presence, which indicates the greater difficulty of processing such material for readers (subjective complexity measures are presented above).

Of greatest interest, however, for addressing the research questions is the distribution of saccades to the image during the processing of stimuli 1 and 2. The data show that when a semantic link between verbal and pictorial components is present, the number of saccades to the image decreases as reading progresses (from an average of 15 saccades at the first predictor to 6 saccades at the last; the exception is the eighth predictor (9 saccades), which marks the beginning of a new part of the text and introduces new content). Some fragments that are formally and semantically capable of referring to the image cease to serve as predictors in the latter part of the text (zones 6, 7, 10). The possibility of inconsistent reference to the image may be explained by the observations of Jian and Wu (2015), who, comparing their data with those of Hegarty and Just (1993), noted that in processing biology texts, unlike mechanics texts, readers were relatively less dependent on the information presented by images. In this regard, having established a link between components of a PCT at early stages of processing, readers subsequently lose the need for regular confirmation of this link, although when they do confirm it, they rely on the identified verbal predictors.

When no semantic link is present between the components of a PCT, a relative decrease in the mean number of saccades to the image also occurs from the beginning to the end of reading. However, all fragments that are formally and semantically capable of referring to the image consistently function as predictors, despite the potential independence from the image's content. Thus, the failure of early attempts to establish a link between verbal and pictorial components of a PCT does not diminish the functionality of verbal predictors. Encountering them during reading, respondents subconsciously continue to actively refer to the image.

6. CONCLUSION

The conducted study contributes to the development of the concept of perception of an educational PCT, and specifically, it reveals the mechanisms by which a reader establishes a connection between its heterogeneous elements. For the first time, such mechanisms of integration were identified. Moreover, this type of study, conducted using the eye-tracking method, was for the first time carried out on data concerning the cognitive processing of an educational PCT in Russian by Russian-speaking readers.

This study confirmed the dominance of the verbal component of the PCT in conditions that orient towards subsequent monitoring of the knowledge obtained. This conclusion corresponds to the results obtained earlier from data on the processes of reading educational PCTs (Jian & Wu, 2015; Lee & Wu, 2017). It is based on the absolute prevalence of the chosen reading strategy 'text-first with references to the image', as well as the significantly larger proportion of time spent processing the verbal component compared to the proportion of time spent processing the image. Differences between the strategies used by readers were not considered in this study.

At the same time, a comparison with the results of the aforementioned studies, conducted on the basis of an analysis of the processes of reading texts of various subject content, as well as taking into account the typical principles of placing a PCT in biology textbooks, gives reason to believe that the nature of reading is influenced by the learners' experience gained when working with educational texts of a certain content type. The coincidence of the results of using this experience is determined by the unity of information presentation in biology textbooks.

This assumption, of course, needs further clarification with the involvement of data on the reading of educational texts of other content types.

Furthermore, the study confirmed the regulatory function of the verbal component of an educational PCT during its reading. This function is manifested in the presence of a stable set of

verbal predictors of reference to the image, which determine the reader's need to establish a connection between the heterogeneous elements of the PCT in situations where information about the structure of the described object is found in its verbal component, which he is accustomed to receiving in the form of a PCT within his educational experience. The composition of the verbal predictors that realise the regulatory function corresponds in content to the potentially visualisable information that is usually presented in the image in textbooks. This conclusion is confirmed, firstly, by the consistent focusing of saccades to the image on verbal units of a certain type, and secondly, by the preservation of their regulatory function in the absence of a semantic link between the heterogeneous elements of the PCT, and even by the strengthening of its activity. The activity of the regulatory function of the verbal component in the absence of its semantic link with the image confirms the idea that readers implement the mindset to establish a connection with the image already at the first stage of text understanding – at the decoding stage, that is, at the stage of extracting semantic information from individual components of the text.

Additionally, the study showed that the discrepancy between the information presented in the image and the content of the verbal component of the PCT greatly complicates the reading process. This is especially important because, as our data show, the reader cannot ignore the image and unconsciously continues an active search for a connection throughout the entire time of processing the verbal component.

The results of the study are important for the organisation of educational practices, in particular for the processes of teaching visual literacy in a specific subject domain, as well as for the effective placement of educational text presented in polycode form.

References

Anisimov, V. N., & Luzhin, A. O. (2020). *Method of determining efficiency of visual presentation of text materials.* Patent No. 2722440 C1. Russian Federation. EDN: JAVCJT

Anisimova, E. E. (2003). *Text linguistics and intercultural communication (on the material of creolized texts)*. Academia.

Babakova, T. A. (2022). Scientific-methodical project on higher education pedagogy in the context of academic writing. *Lifelong Education: The 21st Century*, *3(39)*, 30-43. https://doi.org/10.15393/j5.art.2022.7846

Bakhtin, M. (1986). The problem of the text. In C. Emerson & M. Holquist (Eds.), *Speech genres and other late essays* (pp. 101-131). University of Texas Press.

Barthes, R. (2004). Introduction to the structural analysis of narratives. In M. Bal (Ed.), *Narrative theory: Critical concepts in literary and cultural studies* (pp. 65-116). Routledge.

Bernatskaya, A. A. (2000). On the problem of text 'creolization': History and current state. *Speech Communication: Specialized Messenger, 3,* 104-110.

Blinova, E., & Shcherbakova, O. (2020). Patterns of intellectual behaviour manifested when working with verbal and iconic texts. In A. V. Shaboltas (Ed.) *Psychology of the XXI century: Current challenges and achievements* (pp. 78-88). Skifiya-Print.

Chernyavskaya, V. E. (2009). *Text linguistics: Polysemy, intertextuality, interdiscursivity.* LIBROKOM.

de Beaugrande, R., & Dressler, W. U. (1981). *Introduction to text linguistics*. Longman.

Derrida, J. (2016). *Of grammatology.* Jhu Press.

Dragomilov, A. G., & Mash, R. D. (2008). *Biology. 8th grade: Textbook for students of general educational institutions.* Ventana-Graf.

Eiger, G. V., & Yukht, V. L. (1974). Towards the construction of a typology of texts. In *Proceedings of Research Conference on Text Linguistics* (pp. 103-109). Moscow State Linguistic University Press.

Fairclough, N. (2023). Critical discourse analysis. In M. Handford & J. P. Gee (Eds.), *The Routledge handbook of discourse analysis* (pp. 11-22). Routledge. https://dx.doi.org/10.4324/9781003035244

Forceville, C. (2006). Non-verbal and multi-modal metaphor in a cognitivist framework: Agendas for research. In G. Kristiansen, M. Achard, R. Dirven, & F. J. R. de Mendoza Ibáñez (Eds.), *Cognitive linguistics: Current applications and future perspectives* (pp. 379-402). Mouton de Gruyter.

Foucault, M. (1971). Orders of discourse. *Social Science Information, 10*(2), 7-30. https://dx.doi.org/10.1177/053901847101000201

Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. Longman.

Hannus, M., & Hyönä, J. (1999). Utilization of illustrations during learning of science textbook passages among low- and high-ability children. *Contemporary Educational Psychology*, *24*(2), 95-123. https://doi.org/10.1006/ceps.1998.0987

Hegarty, M. (1992). Mental animation: Inferring motion from static diagrams. *Spatial Cognition & Computation*, *2*(3-4), 211-235.

Hegarty, M., & Just, M. A. (1993). Constructing mental models of machines from text and diagrams. *Journal of Memory and Language*, *32*(6), 717-742. https://dx.doi.org/10.1006/jmla.1993.1036

Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing*, *2*(2), 127-160. https://doi.org/10.1007/BF00401799

Jewitt, C. (Ed.). (2009). *The Routledge handbook of multimodal analysis*. Routledge.

Jian, Y.-C. (2013). How do students read diagrams with text in high school biology textbooks: An eye-tracking study. *International Journal of Science Education*, *35*(3), 467-496.

Jian, Y.-C. (2015). Visual literacy does not simply mean reading diagrams: The role of text-diagram integration and cognitive processing in comprehending illustrated science text. *Journal of Science Education and Technology*, *24*, 43-55.

Jian, Y.-C. (2016). Fourth graders' cognitive processes and learning strategies for reading illustrated biology texts: Eye movement measurements. *Reading Research Quarterly, 51*(1), 93-109. https://doi.org/10.1002/rrq.125

Jian, Y.-C., & Wu, C.-J. (2012). The effect of arrows in an illustration when reading scientific text: Evidence from eye movements. *Chinese Journal of Psychology*, *54*, 385-402.

Volume 9 Issue 3, 2025, pp. 113–125 doi: 10.22363/2521-442X-2025-9-3-113-125

Jian, Y.-C., & Wu, C.-J. (2015). Using eye tracking to investigate semantic and spatial representations of scientific diagrams during text- diagram integration. *Journal of Science Education and Technology*, *24*(1), 43-55. https://dx.doi.org/10.1007/s10956-014-9519-3

Kalugin, Y. E., & Prokhorov, A. V. (2023). Understanding of educational texts. *Universum: Psychology and Education*, *4*(106), 46-49.

Kibrik, A. A. (2010). Multimodal analysis of discourse: A new area of research and a training course. *Bulletin of Moscow State University. Series 9: Philology, 5,* 114-125.

Kolesov, D. V., Mash, R. D., & Belyaev, I. N. (2016). *Biology. Human: Textbook for the 8th grade of general educational institutions.* Drofa.

Kosenko, V. S. (2019). The impact of a creolized text non-verbal and verbal parts on its perception (an experimental study). *Polylinguality and Transcultural Practices*, *16*(3), 310-321. https://doi.org/10.22363/2618-897X-2019-16-3-310-321

Kress, G. (2009). *Multimodality: A social semiotic approach to contemporary communication*. Routledge. https://dx.doi.org/10.4324/9780203970034

Kress, G., & van Leeuwen, T. (2020). *Reading images: The grammar of visual design*. Routledge. https://dx.doi.org/10.4324/9781003099857

Kristeva, J. (2009). *Narration et transformation.* De Gruyter Mouton.

Kunitsyna, E. Y. (2010). Features of the perception of creolized texts (an experimental study on the material of agitational-political texts). *Political Linguistics, 2*, 115-119.

Kuryanovich, A. V. (2022). Educational texts as didactic and methodological material in the professional activity of a Russian language teacher. In *Proceedings of the 20th International Research Conference 'Personal Development in the Educational Space'* (pp. 55-60). Shukshin Altai State Humanitarian Pedagogical University.

Lacan, J. (1981). *The language of the self: The function of language in psychoanalysis.* JHU Press.

Lee, W.-K., & Wu, C.-J. (2017). Eye movements in integrating geometric text and figure: Scanpaths and given-new effects. *International Journal of Science and Mathematics Education*, *16*(4), 699-714. https://dx.doi.org/10.1007/s10763-016-9790-2

Lemke, J. L. (1998). Multiplying meaning: Visual and verbal semiotics in scientific text. In J. R. Martin & R. Veel (Eds.), *Reading science: Critical and functional perspectives on literacies of science* (pp. 87-113). Routledge.

Levin, J. R. (1981). On functions of pictures in prose. In F. J. Pirozzolo & M. C. Wittrock (Eds.), *Neuropsychological and cognitive processes in reading* (pp. 203-228). Academic Press. https://doi.org/10.1016/B978-0-12-185030-2.50013-5

Levin, J. R., Anglin, G. J., & Carney, R. N. (1987). On empirically validating functions of pictures in prose. In D. M. Willows & H. A. Houghton (Eds.), *The psychology of illustration: Basic research* (pp. 51-85). Springer-Verlag. https://dx.doi.org/10.1007/978-1-4612-4674-9_2

Lin, J. J. H., & Lin, S. S. (2014). Cognitive load for configuration comprehension in computer-supported geometry problem solving: An eye movement perspective. *International Journal of Science and Mathematics Education, 12*(3), 605-627. https://doi.org/10.1007/s10763-013-9479-8

Luria, A. R. (2009). *The main problems of neurolinguistics.* LIBROKOM.

Machin, D. (2007). *Introduction to multimodal analysis*. Hodder Arnold.

Mason, L., Tornatora, M. C., & Pluchino, P. (2013). Do fourth graders integrate text and picture in processing and learning from an illustrated science text? Evidence from eye-movement patterns. *Computers & Education*, *60*(1), 95-109. https://doi.org/10.1016/j.compedu.2012.07.011

Mayer, R. E. (2005). *The Cambridge handbook of multimedia learning*. Cambridge University Press.

Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). Cambridge University Press.

Mayer, R. E., & Gallini, J. K. (1990). When is an illustration worth ten thousand words? *Journal of Educational Psychology*, *82*(4), 715-726. https://doi.org/10.1037/0022-0663.82.4.715

Nekrasova, E. D. (2014). On multimodal perception of the text (a psycholinguistic experiment). *Tomsk State University Journal*, *378*, 45-48. EDN: RZCHTJ

O'Halloran, K. L. (Ed.). (2004). *Multimodal discourse analysis: Systemic-functional perspectives*. Continuum.

Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford University Press.

Pêcheux, M., Hak, T., Helsloot, N., & Macey, D. (1995). *Automatic discourse analysis.* Rodopi.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*(3), 372-422. https://doi.org/10.1037/0033-2909.124.3.372

Rayner, K., Rotello, C. M., Stewart, A. J., Keir, J., & Duffy, S. A. (2001). Integrating text and pictorial information: Eye movements when looking at print advertisements. *Journal of Experimental Psychology: Applied*, *7*(3), 219-226. https://doi.org/10.1037/1076-898X.7.3.219

Reichle, E. D., Rayner, K., & Pollatsek, A. (2009). The E-Z reader model of eye-movement control in reading. *Behavioral and Brain Sciences, 26*(4), 445-476. https://dx.doi.org/10.1017/s0140525x03000104

Rokhlov, V. S., & Trofimov, S. B. (2007). *Biology. 8th grade: Man and his health. Textbook for general educational institutions.* Mnemozina.

Sabinina, A. A. (2009). Training text: Structure and pragmatics. *Herzen University Journal of Humanities & Sciences, 97*, 222-225. EDN: KJBLDZ

Sapin, M. R., & Bilich, G. L. (2010). *Normal human biology: Book 1.* Medical Information Agency.

Schnotz, W. (2005). An integrated model of text and picture comprehension. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 49-69). Cambridge University Press.

Schnotz, W., & Horz, H. (2009). On the role of embodiment in the comprehension of static and animated displays. In R. Lowe & W. Schnotz (Eds.), *Learning with animation: Research and implications for design* (pp. 89-111). Cambridge University Press.

Scollon, R., & Scollon, S. W. (2003). *Discourses in place: Language in the material world*. Routledge. https://dx.doi.org/10.4324/9780203422724

Sériot, P. (1997). Ethnos et demos: la construction discursive de l'identité collective. *Langage & Société, 79*(1), 39-51.

Shelestyuk, E. V. (2010). Laws and stages of text comprehension. *Issues of Cognitive Linguistics, 2(23),* 85-90. EDN: MVDFNZ

Slough, S. W., McTigue, E. M., Kim, S., & Jennings, S. (2010). Science textbooks' use of graphical representation: A descriptive analysis of four sixth-grade science texts. *Reading Psychology*, *31*(3), 301-325

Sonin, A. G. (2005). *Modelling the mechanisms of understanding polycode texts* [Doctoral dissertation, Institute of Linguistics, Russian Academy of Sciences]. Russian State Library. https://search.rsl.ru/en/record/01003258360

Sorokin, Y. A., & Tarasov, E. F. (1990). Creolized texts and their communicative function. In R. G. Kotov (Ed.), *Optimisation of speech impact* (pp. 180-186). Nauka.

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction*, *4*(4), 295-312. https://doi.org/10.1016/0959-4752(94)90003-5

Tyutkova, I. A. (2016). *Pedagogical thesaurus: A study guide.* Institute for Lifelong Learning.

Valgina, N. S. (2003). *Text theory.* Logos.

van Dijk, T. A. (1997). *Discourse as structure and process*. Sage.

van Dijk, T. A. (2015). Critical discourse analysis. In D. Tannen, H. E. Hamilton, & D. Schiffrin (Eds.), *The handbook of discourse analysis* (pp. 466-485). Wiley. https://dx.doi.org/10.1002/9781118584194.ch22

van Leeuwen, T. (2005). *Introducing social semiotics: An introductory tectbook*. Routledge.

Vashunina, I. V., Tarasov, E. F., Nistratov, A. A., & Matveev, M. O. (Eds.). (2020). *Creolized text: Semantic perception.* Institute of Linguistics, Russian Academy of Sciences.

Wodak, R. (2004). Critical discourse analysis. *Qualitative Research Practice, 185,* 185-204.

Zlokazov, K. V., & Lipnitsky, A. V. (2018). Violence in a creolized text: A study of the patterns of understanding. *Political Linguistics*, *71*, 135-143.

ABOUT THE AUTHORS

Inna V. Tubalova
DSc in Linguistics, Dean of the Faculty of Philology, Professor
Russian Language Department, Faculty of Philology, Tomsk State University, Tomsk, Russia
Postal address: 634050, Tomsk, Lenin Ave., 36.
Email: tina09@inbox.ru
ORCID ID: https://orcid.org/000-0001-6519-0739

Alena V. Garina
Assistant Lecturer
Russian Language Department, Faculty of Philology, Tomsk State University, Tomsk, Russia
Postal address: 634050, Tomsk, Lenin Ave., 36.
Email: alyonushka050700@gmail.com
ORCID ID: https://orcid.org/0009-0006-9821-1675