



Article

Metaphor as a Device for Seeing: Exploring Meaning-Making in Audio Description and Phraseological Pedagogy

Elena Isakova

The Department of Foreign Languages and Crosscultural Communication, Tver State University, 33 Zhelyabova Str., 170100 Tver, Russia; isakova.em@tversu.ru; Tel.: +7-(4822)-34-24-52

How To Cite: Isakova, E. (2026). Metaphor as a Device for Seeing: Exploring Meaning-Making in Audio Description and Phraseological Pedagogy. *Studies of Applied Linguistics in Asia*, 1(1), 209–220. <https://doi.org/10.53941/sala.2026.100017>

Received: 2 June 2026

Revised: 17 August 2026

Accepted: 17 September 2026

Published: 9 October 2026

Abstract: This conceptual article explores metaphor as a cognitive and semiotic mechanism underlying meaning construction across two seemingly distinct domains: audio description of picturebooks and metaphor-based visualization in foreign language teaching. Both fields are united by a common epistemological concern—how meaning becomes accessible when it is not directly available through either visual perception or compositional linguistic analysis. The article addresses two questions: (1) What cognitive mechanism is shared by the verbalization of visual in audio description and the visualization of verbal in idiom pedagogy? and (2) How can these practices be understood within a common framework of multimodal meaning-making? In this article metaphor is approached not as a decorative linguistic device but as a fundamental “device for seeing”, enabling the transition from perceptual input to conceptual understanding. In the case of picturebooks, where visual and verbal modes interact in a co-dependent structure, audio description raises the problem of transforming visual meaning into verbal form while preserving interpretative openness. In foreign language pedagogy, idiomatic expressions often resist literal interpretation, requiring alternative strategies for meaning activation through conceptual metaphor and visual support. Drawing on examples from the bilateral research project on metaphor-based idiom visualization and illustrative cases from picturebook audio description, the article argues that both practices rely on the same underlying mechanism: guided cross-modal inference, whereby a mediator selects and organizes cues in one modality to facilitate the construction of meaning that is implicit or inaccessible in another. From this perspective, visualizing the verbal and verbalizing the visual are complementary forms of multimodal mediation. The article proposes a conceptual framework that positions metaphor as a central mechanism of multimodal meaning-making in both accessibility practices and foreign language pedagogy.

Keywords: metaphor; conceptual metaphor theory; multimodality; audio description; picturebooks; idioms; phraseology; foreign language teaching; visualization; cognitive linguistics

1. Introduction

A metaphor can easily be seen as a bridge etymologically “carrying over” from one thing to another (Cortazzi & Jin, 1999). One of the things that all scholars agree on in metaphor research is that metaphor can be defined as “a device for seeing something in terms of something else” (Burke, 1945, as cited in Cameron, 1999).

In this article we suggest that metaphor is a device for seeing. This definition immediately raises a question: “What if we can see words, but we do not see the meaning?” Or even more radically: “What if we cannot see at all?” These questions will be developed in this article in two directions. The first direction—the picturebooks as a unique literary form and genre, and the audio description of them. Picturebooks are a unique art form where visual



and verbal texts are co-equal and meaning emerges in the gap between them, in their synergy. What happens if this visual layer is inaccessible? Audio description (AD) attempts to solve this problem by transforming images into words. Its goal is not only to describe, but to make images experienceable—to allow the audience to construct meaning independently. However, here we face tension. Theory is often disconnected from user-experience. In Russian tradition audio description is expected to be strictly objective and a picturebook is not treated as an object of audiovisual translation itself.

The second direction arises from a classroom challenge. How can we help learners *see* the meaning where it is not immediately visible? Idioms are a perfect example where learners see the words but not the meaning. As early as Katz, idioms were described as expressions whose meaning cannot be derived from individual words (Katz, 1973). Recent research tells us that idioms are non-arbitrary features of a language (Gibbs, 1990; Thyab, 2016; O'Reilly et al., 2025). Most of them are grounded in conceptual metaphors. For example, when we understand ANGER as a HEATED FLUID IN A CONTAINER, expressions like *blow your stack*, *flip your lid* suddenly become meaningful. Learners begin to see the logic. However, very often learners and teachers do not see this metaphor (see Glucksberg et al., 1993). This is where metaphorically motivated visualization becomes helpful. This claim will be developed further.

It is noteworthy that in this article the term *metaphor* is used in three related but distinct senses. First, we adopt the cognitive-linguistic view of conceptual metaphor as a mechanism of understanding one conceptual domain in terms of another (see Lakoff & Johnson, 1981). Second, we discuss visual or multimodal metaphors, in which conceptual mappings are expressed through the interaction of verbal and visual modes (see Peng, 2026 for an overview). Finally, we examine metaphorical language as one possible strategy in audio description, where figurative expressions may guide recipients' interpretation of visual content (see Luque Colmenero & Soler Gallego, 2020). Although these phenomena differ in their semiotic realization, Bolognesi (2019) argues that they may converge at deeper levels of conceptual representation, particularly at the level of image schemas. From this perspective, both can be viewed as supporting multimodal meaning-making through shared cross-domain mappings.

Finally, it could be argued that to visualize the verbal and to verbalize the visual are the processes that are closely linked together and are at the core of the overall perception and interpretation of the world around us.

2. Visual and Verbal in Meaning-Making

The process of integrating verbal and visual signs in meaning-making has become a common practice in the present multimodal world where people usually merge

words and images in the production of messages (Khadka, 2025). In most theories of multimodality, meaning-making is approached—implicitly or explicitly—from the perspective of a speaker, a producer, or an author (Bucher, 2017). Hans-Jürgen Bucher puts the recipients at its center, asking the question: How do recipients integrate the different modalities like text, picture, sound, and design into a coherent meaning? Bucher compares perceiving these multimodal arrangements to a conversation or a dialogue between the recipient and the medium, where the recipients act as if the medium is a partner who continuously provides the necessary information for deeper comprehension (Bucher, 2017).

By using visuals to explain an idiom, a teacher also engages students in the process of multimodal meaning-making. Students actively construct meaning by relating the visual representation to the linguistic expression. Rather than decoding the idiom in isolation, students engage in an inferential process in which meaning is constructed through the interaction of multiple modes, namely the linguistic mode, represented by the verbal constituents of the idiom, and the visual mode, which may depict its literal meaning, figurative meaning, or a combination of both. A teacher acquires a role of a mediator in what Bucher calls a *dialogue* where the meaning-making happens (see Figure 1). At the core of a multimodal arrangement there is a process of *visualizing the verbal*.



Figure 1. The multimodal arrangement of understanding an idiom's meaning.

Another example where the role of a mediator is crucial for the multimodal meaning-making is *reading a picturebook* by the blind and visually impaired people. Picturebooks are complex multimodal texts (Sherif, 2022). They entail both visual and verbal modes, which affect meaning and shape the reader (Painter et al., 2013). In this case, the multimodal arrangement is represented by the visual code of picturebook illustrations and the verbal code of an audio description (AD). The mediator can be a translator, who creates an audio description for the picturebook (see Figure 2). Central to a multimodal arrangement in this case is the process of *verbalizing the visual*.

Since both words and images make their own relatively autonomous contribution to the overall semantic, aesthetic and emotional effect of the picturebook, it has often been observed that the picturebook is closer to other mixed narrative forms such as drama or film than to verbal fiction (Wu, 2014). In this

article we support the claim that the picturebook is the righteous object of audiovisual translation (AVT) (Isakova, 2026). As a multimodal text, a picturebook becomes audiovisual through read-aloud performance or audio description. In the next section this will be discussed in greater detail.

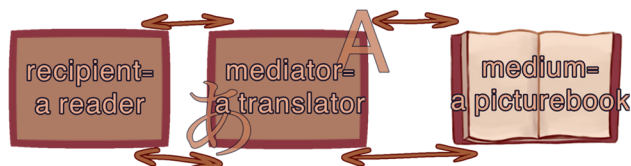


Figure 2. A multimodal arrangement of reading a picturebook.

The reason for comparing two meaning-making activities also lies in the fact that the so-called secondary audience of AD comprises sighted people, for example students, cognitively challenged people, and any people who need audio support of their reading practices. Recent research showed that adding audio description to educational materials helped guiding viewer's attention to important elements in audiovisual content (Perego, 2016). Some of the studies on audio description tasks demonstrate that revoicing can stimulate metalinguistic reflection and contribute to the development of lexical and phraseological competence (Formentelli, as cited in Lertola, 2019). AD has started to be considered as an effective task in the foreign language classroom (Lertola, 2019). Like an audio describer, a teacher mediates learners' understanding by selecting and organizing multimodal cues that facilitate meaning-making. This functional similarity makes audio description a useful point of comparison for examining the teacher's mediating role in supporting the comprehension of multimodal picturebooks.

Taken together, these observations suggest that, teaching idiomatic expressions and creating audio descriptions can both be understood as forms of guided multimodal meaning construction. In both cases, a mediator, such as a teacher or translator, selects and organizes multimodal cues that guide the recipient's inferential meaning-making. The following section examines the process by which the mediator selects how to present the medium to recipients.

3. Selection Process

The teaching strategy that teachers, acting as mediators in the learning process, may employ in their classroom depends on many factors. Teacher beliefs are central to their instructional decisions and significantly impact educational quality and student learning outcomes (Li & Ma, 2025). Beliefs are understood as the principles teachers hold regarding what is important in language teaching and how best to facilitate learning (Paquet et al., 2026). These beliefs play a critical role in determining

instructional decisions, particularly in the field of language education, where attitudes toward language use, learning processes, and learner diversity directly influence pedagogical choices (Wesely et al., 2024, as cited in Paquet et al., 2026).

Accordingly, teachers' beliefs about idiom instruction are also likely to influence the methods they adopt in the classroom. Second language teachers differ considerably in their views on the usefulness of idioms for L2 learners (Ramonda, 2016), and idioms have long been recognized as a challenge for language instruction (Heidari & Aliyar, 2025). For this reason, it is important that teachers' instructional decisions be informed by current research in idiom pedagogy. For example, studies have shown that visualizations encouraging *metaphorical elaboration* are among the most effective ways of supporting idiom learning (Ramonda, 2016). Kris Ramonda defines *metaphorical elaboration* as the process through which metaphorical or metonymical pictorial input is used as a means to activate deep processing (Cermak & Craik, 1979, as cited in Ramonda, 2016) and dually code (Paivio, 1986) the form-meaning link of metaphorical idioms for the purposes of elucidating their figurative meaning or enhancing their recall (Ramonda, 2016). From the perspective of the teacher as a mediator, this raises an important question: how should such multimodal cues be selected and organized? The following section, "Visualizing idioms—verbalizing the meaning", addresses this question by outlining three ways in which teachers can organize multimodal cues to facilitate idiom learning.

The strategies that an audio describer employs to mediate readers' understanding of a multimodal text, such as a picturebook, also depend on many factors. AD is a cognitively complex phenomenon, since it involves three different perspectives (Holsanova, 2016 as cited in Vermeulen & Ibáñez Moreno, 2017): (1) A production perspective, focused on the sighted audio describer; (2) a reception perspective, focused on all kinds of receivers; and (3) the 'meeting of minds' perception, focused on the exchange between both the sighted audio describer and the users, blind or sighted users. In their role as social intermediaries, audio describers need to have a deep understanding of the world and reality of their target audience (sensorially, socially or culturally challenged) (Vermeulen & Ibáñez Moreno, 2017). Their descriptive choices are also influenced by how they conceptualize the source medium. For example, if a picturebook is regarded as a multifaceted art form (Arizpe, 2021), the audio describer may adopt conventions associated with the audio description of visual art, such as museum or exhibition audio description. In other words, an audio describer will have to select an audio description (AD) style. Colmenero (2024) in his recent overview of existing AD styles, points out that subjectivity has been a point of contention within the domain of AD, particularly due to the inconsistency

between various guidelines. Some advocate for a neutral or objective approach, while others argue in favor of more subjective descriptions, which have shown benefits for certain blind and partially sighted (BPS) visitors. The objective style aligns with the recommendations provided in existing guidelines for creating AD for visual art. At the same time new styles of AD (metaphoric, gist, poetic, synesthetic, and interpretative voices) can be found in museums across Europe (Colmenero, 2024). In the description of the metaphorical style Colmenero mentions that metaphor is used in the AD guides of art museums with the communicative function of transferring through the linguistic code that cannot be perceived through the visual channel (Colmenero, 2024). The section “Verbalizing the visual—making the unseen accessible” illustrates how a picturebook page can be audio described using both objective and metaphorical AD styles.

4. Visualizing Idioms—Verbalizing the Meaning

Human interaction is multimodal in its nature; however, vision often plays a central role in coordinating attention, supporting joint action, and constructing meaning in interaction. As Goodwin argues, participants engaged in social activity attend not only to one another but also to relevant phenomena in their shared environment, and vision can be central to this process (Goodwin, 2000). Consequently, visual representations have become an important resource for communication and meaning-making across educational and cultural contexts. In both Russian and English, one can find idioms that reflect the importance of visual perception for humans:

“The eyes are the mirror of the soul” is an expression emphasizing that one can see a person’s feelings and emotions through their gaze.

“A sight for sore eyes” is someone or something pleasant to look at.

“The apple of one’s eye” is a dear person.

When teaching foreign language idioms, a legitimate question arises: how can knowledge about the characteristics of human visual perception be most effectively utilized for educational purposes? It is worth recalling that one theory of relevance to second language acquisition is Dual Coding Theory (Paivio, 1986). According to this approach, verbal and nonverbal information are processed through different channels and represented differently in the human mind. Consequently, dual coding of input provides an alternative pathway for subsequent retrieval. Information presented through both visual and verbal codes will be stored and retrieved more easily than information presented through only one modality. Thus, despite the depth and influence of visual perception, the most effective way to present information will be a combination of *verbal + visual*. Therefore,

employing both verbal and visual pathways could facilitate the interpretation and retention of idiomatic language, which was already demonstrated in some studies (Boers et al., 2009; Cuevas, 2016; Sadoski, 2005; Sadoski & Paivio, 2013, as cited in Samburskiy, 2020).

However, Dual Coding Theory does not imply that any visual accompaniment will automatically enhance learning. Rather, its pedagogical value depends on how effectively visual and verbal information are integrated to support meaning construction. Promising methods for teaching English idioms which demonstrate that not all visualizations are equally effective are Chris Ramonda’s *metaphorical elaboration method* (Ramonda, 2022) and *the visual metaphorization method* (Isakova, 2023). The latter is being developed within the framework of the international project of Tver State University and the University of Klagenfurt in Austria, “Phraseological Aspect in Foreign Language Teaching Based on Visual Metaphorization” (Isakova, 2023). Both methods aim to create the most effective image (illustration of the idiom) that will accompany the verbal component of the idiomatic expression and stimulate its memorization during the process of learning a foreign language. One of the significant experiments that is part of the bilateral project “Phraseological Aspect in Foreign Language Teaching Based on Visual Metaphorization” is described in the article *Putting images of English idioms to the test: L1, L2, and learner ratings on visualization types of idioms* (Papitsch et al., 2025). It is explained there that as part of a larger research project (see <https://www.aau.at/en/english/research-projects/metaphors-in-english-idioms/>), we have opted for a picture design that is as unified as possible across the following three types:

- Type A: Literal

The image visualizes the literal meanings of the content words making up the idiomatic expressions, i.e., its literal scene.

- Type B: Meaning

The image captures the overall meaning of the idiomatic expression.

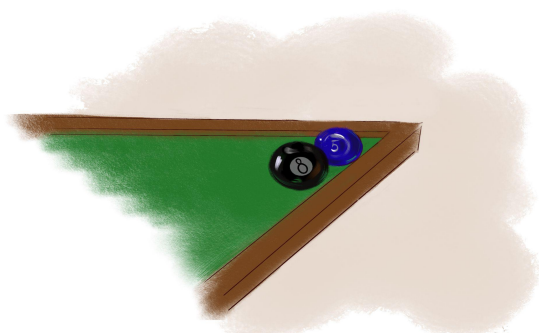
- Type C: Combination

The image combines visual elements relating to the literal meanings of the content words making up the idiom with the depiction of the overall meaning of the idiom.

An example of the three types for the idiom *behind the eight ball* (“in a highly disadvantageous position”) is given below (see Figure 3). These examples are illustrative and are not intended to constitute a full report of the empirical study; rather, they serve to demonstrate how different types of idiom visualization can be conceptualized within the framework proposed in this article.

Type A reflects the literal meaning of the components of an idiom. The image depicts the game of pool (billiards), where one of the balls is trapped directly

behind the black eight ball. Type B reflects the meaning of the idiom where one of the people in a picture is in a disadvantageous position. Type C unites both the plane of expression and the plane of content of the idiom (meaning + metaphorical component): a person is located in a position where a giant ball is approaching. The position is disadvantageous. One possible conceptual metaphor that can motivate the idiom is *BAD IS BEHIND*.



Type A



Type B



Type C

Figure 3. An example of the three types of visuals for the idiom *behind the eight ball*.

Thus, in the idiom-visualization examples, conceptual metaphor refers to the underlying mapping that motivates the figurative meaning of the idiom, whereas the image constitutes a pictorial or multimodal realization of this mapping. Depending on the design, the image may depict the literal source domain, represent the figurative target domain, or combine elements of both.

A teacher may select any of the visuals to guide students' multimodal meaning-making. The team of researchers from Austria and Canada have analysed the data they obtained during the above-mentioned bilateral project and concluded that combining imagery that depicts the literal and figurative meanings of idioms is most successful at leveraging learning for both idioms and their meanings. Imagery that depicts just the literal meanings of idioms seems to work second best—performing as successfully as combined imagery for aiding learning of just idioms' forms, and almost as well as the combined imagery with idioms' meanings (and better than with figurative imagery). Imagery depicting only the figurative (i.e., idiomatic) meaning of idioms was the least successful at helping L2 English students learn idiom forms and their meanings (Rasse, in press).

If teachers as mediators in the meaning-making process select to use the combined imagery, they can use one of the project deliverables, namely a booklet for learners and teachers of English containing 59 Advanced English idioms each accompanied by a combined image type that has been published as a freely accessible resource online (https://www.aau.at/wpcontent/uploads/2025/09/59_Advanced_English_idioms_compressed_copy.pdf).

Some teachers may prefer to continue using traditional sources of idioms' visualizations, such as dictionaries and educational course books. Idiom dictionaries also use different visual representations to support the meaning clarification of an idiom. First, we will compare two dictionaries: Cambridge Dictionary of American Idioms and Cambridge Idioms Dictionary. Thus, in the Cambridge Dictionary of American Idioms, some idiomatic units and dictionary entries are accompanied by an illustration that includes both the literal and figurative meaning of the idiom. So, for the idiom *sweep something under the carpet* (Cambridge University Press, 2006b, p. 58), the image of a woman who sweeps some scrolls under the carpet is used (the direct or literal meaning). The scrolls symbolize such abstract concepts as injustice, slavery, etc. (the figurative meaning of the idiom).

The Cambridge Idioms Dictionary also contains pictures of idioms. However, here we can see the classical method of visualizing only the idiomatic meaning of a unit. Thus, for the idiom *lend somebody a hand* we get an illustration of helping in a store (Cambridge University Press, 2006a, p. 236).

Now, let us compare the use of visual aids in educational literature. *How Idioms Work* by Yvonne Clarke is intended for adult English language learners who, despite being proficient in English, find idiomatic language confusing or incomprehensible. This fully photocopiable reference contains images of the literal meaning of high-frequency idioms, chosen specifically for their lexical value in the modern world. Ten eight-page sections include flashcards, definitions, a reading activity section (using authentic newspaper articles), and a review section. The

guide also includes tips and advice for teachers, 20 suggested activity types, an answer key, and templates for developing additional activity cards (Clarke, 2010).

English for Everyone: English Idioms from the publishing house DK combines a visual teaching method and presents hundreds of expressions in context, with clear definitions and attractive illustrations that depict the figurative meaning of the idiom. The literal meaning of the idiom is not depicted. The book also covers English phrases and words that are often confused. Each training module is followed by highly focused practical exercises that help you remember the material in the manual (DK, 2019).

Comic strips and memes can serve as another way to teach idioms with the help of their visualizations (Nadiya et al., 2023). The interplay between literal and figurative meanings in the comic strip below (see Figure 4) may serve both to introduce the new idiom and to facilitate its retention in memory.

The comic strip presented above depicts two characters from the book “Crocodile Gena and his friends”—Crocodile Gena himself and Cheburashka (a teddy bear-like creature with big ears). This book (Uspensky, 1966) is well-known in Russia. The comic effect is achieved when Cheburashka, who naturally has very big ears, replies to his friend Crocodile Gena by saying “I am all ears”. It uses an idiom to say that he is listening to his friend with all the possible attention and at the same time it can be understood literally—Cheburashka’s ears are very noticeable for the viewer to say that he consists of only them.

It is obvious that both educational and reference literature use completely different methods for visualizing idioms, and in a classroom setting to make meaning out of idioms will always be a guided cross-modal inference, whereby a mediator selects and organizes cues in one modality to facilitate the construction of meaning that is implicit or inaccessible in another.

Notably, one of the most important stages in the work on creating an idiom illustration is its verbalization—a detailed description of the future idiom illustration by team members for subsequent drawing by the designer.

As the example documented in Ramonda’s (2016) picture collaboration notes demonstrates, creating a comprehensive image of an idiom (meaning + metaphorical component) requires a clear and detailed description of both its expression plan and content plan. At the same time, both parts must logically combine into a single story to function as a unified illustration. These notes document the collaboration between the researcher and the illustrator in developing the visualizations: the researcher proposed a visual representation of an idiom, the illustrator described how the proposed idea could be realized, and the researcher provided follow-up comments and suggestions for refinement (Ramonda, 2016, p. 241).



Figure 4. Visualizing “be all ears” through the interplay of literal and figurative meanings. Note: Original pedagogical illustration created by Anna Isakova for this article using characters from Uspensky’s *Krokodil Gena i ego друз’ya* (1966).

Thus, in the process of verbalizing a visual image, the following stages can be identified:

- (1) Verbalization of the idiom’s meaning. For example, for the idiom “turn heads”, the meeting participants described its meaning: “*If someone or something turns heads, they are so beautiful, unusual, or impressive that people notice them and admire them.*” (Ramonda, 2016, p. 241)
- (2) Verbalization of the expression plan and the content plan of the idiom for one illustration—a task for the designer to draw: “*There could be two or three heads turning because there is an incredibly attractive woman.*” (Ramonda, 2016, p. 241)
- (3) Correction of the verbalization of the visual image—additions or creating a new version of the illustration if, for some reason, the first version of the illustration did not work: “*If it is possible add in some motion marks so that «turn» is more emphasized.*” (Ramonda, 2016, p. 241)

Verbalization of visual images, i.e., describing an image with verbal expressions, can be used as a standalone or complementary method in teaching the memorization and use of idioms. This method can have a positive effect for the following reasons:

- (1) Verbalization helps to consolidate visual impressions in memory. When we describe an image with words, it creates a connection between verbal

and visual representation, which can enhance and improve memory retention.

- (2) Verbalization is an effective means of exchanging information. When people can accurately describe visual impressions, it simplifies communication and the transmission of thoughts to others. It was precisely the verbalization that helped achieve mutual understanding among the members of the scientific team and create effective illustrations of idiomatic expressions for experiments and educational materials.
- (3) Verbalization of visual images is a key element of the educational process. The teacher can use verbal descriptions to help students better understand and associate with visual materials.
- (4) It is worth noting separately that the verbalization of visual images can be an important means of accessibility for people with disabilities, such as those with low vision or blindness. Describing images with lexical units helps them get an idea of visual images.

Verbalization of visual images is the foundation of audiovisual translation and can be used to create audio descriptions of educational materials with illustrations of idiomatic expressions for blind and visually impaired students. The experience of verbalizing the semantic duality of an idiom, creatively connecting its meaning and metaphorical component before the illustration itself is drawn, can also be used at the stage of creating professional audio descriptions for teaching aids on foreign language idioms.

5. Verbalizing the Visual—Making the Unseen Accessible

Contemporary culture is experiencing what researchers describe as a “visual revolution” (Kovrov & Simonyan, 2024), or a “visual turn” (Ishchenko, 2018; Ershova, 2025)—an era in which the image as an iconic sign and visual language as a form of non-verbal communication are becoming equal, and at times even dominant, means of cognition and communication.

In literature, multimodal genres such as *comics*, *graphic novels*, and *picturebooks* have gained considerable popularity. The picturebook is currently undergoing a “new golden age” (Salisbury & Styles, 2012)—a period in which modern printing technologies make it possible to produce high-quality picturebooks in a wide range of formats, precisely within a genre where the physical form of the book, its orientation, the colour reproduction of its visual and graphic components—in short, every paratextual element—is of crucial importance not only for its overall popularity among readers and its ability to attract their attention, but also for the interpretation of the work as a whole and the transmission of its profound aesthetic potential.

The picturebook is a literary genre that is simultaneously defined both as a form of visual art and as a unique product of intermedial discourse. The metaphorical nature of the genre’s very name (picturebook—a book as a picture, as a form of visual art) appears to predetermine the unique synergy of visual and verbal texts that convey the intentions of the author and illustrator and coexist on equal terms on every spread of the book. This multimodal nature of the picturebook genre has been noted in the works of O. A. Moskalenko (Moskalenko, 2017), A. S. Soina (Soina, 2017), D. Sergeev (Sergeev, 2019), and O. N. Myaeots (Myaeots, 2021). Examples of literary works created in the picturebook genre include *The River* by Elena Safonova, a student of K. S. Petrov-Vodkin, *Neighbours* by Katya Denisovich, and *Hello Lighthouse* by Sophie Blackall. In these works, illustrations and text form a unified and inseparable whole. In some picturebooks, verbal language is entirely absent, and the narrative is constructed through visual language—namely, illustrations. Such picturebooks constitute a distinct subgenre known as silent books.

Many scholars who study a picturebook as a literary genre also emphasize its considerable pedagogical value. Thus, in his article *What Are We Reading? Towards a Definition of the Children’s Picturebook*, Dmitry Sergeev speaks of the “value of the aesthetic experience required for viewing illustrations, as well as the presence of visual literacy, which is formed through continual exposure to aesthetically accomplished images. Visual literacy is becoming one of the most important cultural competencies for the younger generation today, which undoubtedly enhances the pedagogical value of the picturebook as a highly artistic model in the context of the rapid increase in visual material surrounding us” (Sergeev, 2019).

Thus, the picturebook represents a complex object of perception and interpretation that requires not only reading skills but also visual literacy and the ability to verbalise an image. Such verbalisation has much in common with ekphrasis as a device. Yu. V. Yarovikova notes that ekphrasis as a device (rhetorical-narratological, literary, artistic, and visual-artistic) is a “text within a text”, that is, a description of a work of art embedded within another literary genre (Yarovikova, 2019). However, the verbalisation of images in a picturebook is more accurately defined by the formula *text + text*; in other words, the visual element is not supplementary. As noted above, it is inextricably linked to the verbal and graphic text.

During the reading of a picturebook, a multilayered process of intermedial comprehension takes place. Most often, verbalization occurs through inner speech, when readers “translate” visual images into verbal representations, as well as during read-aloud sessions or collaborative discussions. It should be noted, however, that these are not the only moments in which

verbalization arises, nor the only situations in which it becomes necessary.

If we acknowledge the high artistic and pedagogical value of the picturebook, we cannot avoid addressing the highly important issue of the media accessibility of this unique literary genre. By the media accessibility of the picturebook, we mean the ability of all categories of readers to fully comprehend the artistic intent of a work and, on an equal basis with other readers, derive aesthetic pleasure from the reading experience.

In this context, people with visual impairments constitute the most vulnerable category of readers. Ensuring the accessibility of other forms of art, such as those exhibited in museums and art galleries across Russia, is by no means a new phenomenon. This is achieved through the complex process of audiovisual translation known as audio description (or typhlocommentary), which requires experience in ekphrastic interpretation. Audio description makes it possible to transform visual images into verbal text, constituting a form of intersemiotic translation that differs significantly from other, more traditional types of translation (Alexandrova, 2022).

Within this framework, the full appreciation of a picturebook requires the adaptation of its visual narrative in such a way that the information conveyed through images becomes fully accessible to audiences with visual impairments, thereby ensuring the integrity and completeness of their understanding of the work.

If we acknowledge the importance of adapting visual texts for blind readers, the question naturally arises as to why picturebooks accompanied by high-quality audio descriptions are virtually absent from the library collections of our country. In the article *Metaphors we see by: how to make picturebooks accessible through audiodescription* (Isakova, 2026), the researcher considers it important to demonstrate why the picturebook, as a multimodal genre, should be regarded as a fully legitimate object of audiovisual translation and draws attention to the need for developing a methodology for the audio description of this genre. While the picturebook is inherently multimodal, its modes are not equally accessible to all readers. For blind readers, access to the visual component requires its mediation through another sensory channel. Audio description therefore transforms the original multimodal configuration into an audiovisual experience, in which visual information becomes available through verbal-auditory input. This transformation is particularly relevant to the present discussion because audio description does not simply add an auditory channel to the picturebook; it mediates access to visual meaning by verbalizing what cannot be directly perceived. Thus, for blind readers, understanding the multimodal picturebook involves a cross-modal process in which visual meaning becomes accessible through verbal-auditory representation.

In the Russian scholarly tradition, the term *audio description* is often replaced by the term *typhlocommentary*. Both terms are defined as “special forms of conveying visual information through verbal description; they share common goals and objects of description but differ in their origins, target audiences, and modes of implementation. Typhlocommentary represents a broader and more specialised approach to adapting visual information for people with visual impairments, whereas audio description was initially conceived as an assistive technology for all categories of users” (Gorbunova & Gorozhanov, 2025). Another significant conceptual distinction lies in the fact that, according to Russian standards, a typhlocommentary should be as objective as possible, whereas audio description allows greater freedom in expressing the translator’s own perception of the visual material and makes more frequent use of figurative language.

As an illustration, let us compare a typhlocommentary and an audio description of the back cover (see Figure 5) of *Listen: How Evelyn Glennie, a Deaf Girl, Changed Percussion* by Shannon Stocker and Devon Holzwarth, prepared by Ivan Motunov, a professional typhlocommentator, at the request of the author of article *Metaphors we see by: how to make picturebooks accessible through audiodescription* (Isakova, 2026). Thus, the two descriptions presented below are drawn from an empirical study and are used here as illustrative examples to demonstrate how different AD styles may mediate the interpretation of the same visual material; the study itself is not reported here as a systematic empirical comparison.



Figure 5. Back cover of *Listen: How Evelyn Glennie, the Deaf Girl, Changed Percussion* by Shannon Stocker and Devon Holzwarth.

5.1. Typhlocommentary

Back cover. In the lower left corner, a young woman in a scarlet dress stands with her eyelids lowered and her arms outstretched as if dancing. Her cheeks are rosy, her expression is calm, and a simple necklace adorns her neck. She holds a pair of mallets with red tips. From each mallet extends a flame-shaped aura rising upward and to the right. The aura is filled with elongated petal-like shapes and small circles. The two flame-like streams are concentrated, creating a diagonal white space that extends toward the upper right corner. In the upper right corner is another depiction of the young woman, wearing a scarlet skirt and a brownish blouse. Her cheeks are also rosy, and her eyelids are lowered. She holds a pair of mallets above a xylophone suspended in mid-air. Within the white space appears the following text: “This is a story of music. Of obstacles. Of strength and hard work. Of all you accomplish when you dream. If you find your own way to...listen. The true and inspiring story of a girl who lost her hearing and grew up to become a world-famous musician.” The word “listen” in the phrase “If you find your own way to...listen” differs from the rest of the text in its typeface, size, and blue, rather than black, color.

5.2. Audio Description

Back cover. In the lower left corner stands the girl, now grown into a confident young woman. She is wearing a scarlet dress. Her cheeks are still rosy, her eyes are gently closed, and her ears are hidden beneath her dark blond hair. She stands with her arms outstretched, as if dancing. In her hands she holds a pair of mallets with bright red heads. From them flows a calm, harmonious aura, the color of flame, rising upward and to the right. It is filled with multicolored petals and small circles. Between the two streams of flame, a diagonal white corridor is formed. In the upper right corner appears another depiction of the young woman. She is wearing a scarlet skirt and a brownish blouse and is playing a xylophone. A soft scarlet aura also radiates from her, gradually unfolding into flower buds. Inside the white corridor is the following text in black:

“This is a story of music. Of obstacles. Of strength and hard work. Of all you accomplish when you dream. If you find your own way to...listen. The true and inspiring story of a girl who lost her hearing and grew up to become a world-famous musician.” The word “listen” in the phrase “If you find your own way to...listen” is given special emphasis. It is positioned in the center of the cover, printed in a different typeface, and colored blue rather than black. The flame-like auras emanating from the young woman’s music evoke the warmth and vitality of life itself.

It can be assumed that audio description is written in a metaphorical style because it does not simply describe what is visible. It employs the realizations of the visual metaphor AFFECTION IS WARMTH. The possible interpretation can be that when people are content, happy and inspired, they feel that the life is kind to them, they compare this feeling with the bodily experience of warmth. They feel the warmth of life. By selecting metaphorical AD, a translator/audio describer guides meaning-making of the readers. Thus, “metaphorical” refers not only to the visual material itself, but to the audio describer’s verbal interpretation of that material through metaphorical language.

As the examples of audiovisual translation of the picturebook’s back cover demonstrate, both audio description (AD) in its metaphorical style and typhlocommentary (TC), or an objective AD style, possess an inherently ekphrastic nature. The audio description text is more emotionally expressive, while both texts can serve not only the primary target audience of AD, such as readers with visual impairments, but also secondary audiences, including audiobook listeners, students, and readers with developmental differences or without them.

6. Conclusions

As noted above, the picturebook is a multimodal artistic text in which the illustration and the verbal component co-construct meaning in inseparable interaction. The composition and structure of illustrations form the visual text of the picturebook—a complete and meaningful semiotic layer that is not merely supplementary, unlike illustrations in an illustrated edition, for example, *Alice’s Adventures in Wonderland* by L. Carroll. Thus, it can be stated with confidence that the picturebook contains a visual channel of meaning-making (illustrations as a visual text), a verbal channel (written text), and a graphic mode (typography, composition). The semiotic space of the picturebook is constructed through the interaction of these modes in a manner analogous to audiovisual works:

- (1) The interaction between image and text, which on the one hand represents an intermedial interaction of artistic systems within a single work, and on the other hand parallels the audiovisual interaction of frame and dialogue;
- (2) The interaction between the illustrator’s colour choices and the poetics of the text with the reader’s emotional response, which closely resembles musical accompaniment;
- (3) Layout (sequencing and spatial arrangement of illustrations in relation to text), as well as the involvement of paratextual elements of the book (cover, endpapers, dedication pages) in the narrative structure, all of which influence narrative rhythm, similarly to the perception of editing rhythms in cinema.

Consequently, it can be concluded that the picturebook shares the same type of modal organisation as audiovisual texts, albeit realised in a static form. Accordingly, a translator of picturebooks may and should face the same types of challenges as a translator of audiovisual texts. In translating picturebooks, one must take into account the interaction between image and text, visual metaphors, cultural codes, colour symbolism, and graphic structure. This fully corresponds to the objectives of audiovisual translation, where the translator likewise works with the synchronisation of modes (text and image/sound), intermodal relations (irony, metaphor), and the pragmatics of reception (orientation towards a specific audience). The picturebook therefore clearly requires an audiovisual approach to translation, as the translator must reconstruct a multimodal message rather than merely translate verbal text.

The picturebook vividly demonstrates that reading is not only an act of textual interpretation but also an act of visual contemplation. It reflects a tendency towards intermediality—the interaction of different sign systems—and towards an ekphrastic experience, in which the reader verbalises a visual image, transforming perception into creative interpretation. Thus, the picturebook is simultaneously a work of visual art and a literary text that can be regarded as a specific form of audiovisual text, where word, image, and sound (in the case of read-aloud or audio description) create a polyphonic space of perception.

In this article we have clearly seen that to visualize the verbal and to verbalize the visual are the processes that are closely linked together and are at the core of the overall perception and interpretation of the world around us.

In bringing together the study of audio description in picturebooks and the visualization of idioms in foreign language teaching, this article has aimed to demonstrate that both domains are ultimately concerned with a common epistemological and pedagogical problem: how meaning becomes accessible when it is not immediately available through direct perception. Although these fields are traditionally treated as separate—belonging respectively to multimodal accessibility studies and phraseological didactics—they converge at the level of cognitive and semiotic processes involved in meaning construction.

In both cases, metaphor functions not merely as a rhetorical device, but as a fundamental mechanism of cognition and perception, enabling the transition from verbal into visual and from visual into verbal. Audio description of picturebooks requires the transformation of visual structures into verbal representations that preserve interpretative openness, while metaphor-based visualization of idioms seeks to making “visible” the seemingly “invisible” and hidden meaning. In this sense, both practices address the same pedagogical challenge: how to support the emergence of meaning where it is implicit, or inaccessible.

Thus, it can be concluded that in this article, the term metaphor is used at several related but distinct levels. *Conceptual metaphor* refers to the underlying cross-domain mapping between a source and a target domain that can motivate figurative meaning. *Pictorial or multimodal metaphor* refers to the realization of such mappings through visual or multimodal representations, including images that combine literal and figurative elements. *Linguistic metaphor*, including *metaphorical expressions* used in audio description, refers to metaphor realized through verbal language. These forms are related but should not be treated as interchangeable: a conceptual metaphor may motivate a linguistic or pictorial expression, while an audio description becomes metaphorical only when its verbal formulation introduces a metaphorical or interpretive mapping.

By treating metaphor as a “device for seeing”, the article argues for a model of perception and learning in which visual and verbal modalities are not opposed, but dynamically interrelated. This perspective allows us to reconsider both accessibility practices and language pedagogy as complementary forms of meaning-making, in which the act of “seeing” is always already an act of interpretation.

Funding

This research received no external funding.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Acknowledgments

The author would like to acknowledge Gromova Victoria (Saint Petersburg Art Gallery OKNA, Russia) for her contribution as a designer within this project. Her illustrations were used in this article (see Figure 3). The author would also like to acknowledge Anna Isakova, for creating Figures 1, 2 and 4. Their visual design work supported the conceptual development of the study and contributed to the integration of visual and verbal dimensions in the representation of metaphorical structures. The consent to use the illustrations in the article was obtained.

Conflicts of Interest

The author declares no conflict of interest.

Use of AI and AI-Assisted Technologies

During the preparation of this work, the author used ChatGPT to check for grammatical errors. After using this tool, the author reviewed and edited the content as

needed and takes full responsibility for the content of the published article.

References

- Alexandrova, E. V. (2022). Аудиовизуальный текст как источник визуальной метафоризации [Audiovisual text as a source of visual metaphorization]. *Теория и практика лингвистического описания разговорной речи [Theory and practice of linguistic description of colloquial speech]*, 4(33), 15–20. <https://elibrary.ru/item.asp?id=49950786>.
- Arizpe, E. (2021). The state of the art in picturebook research from 2010 to 2020. *Language Arts*, 98(5), 260–272. <https://doi.org/10.58680/la202131213>.
- Bolognesi, M. (2019). How do words vs. images construct and represent metaphor. In *TriCoLore 2018: Creativity, Cognition, Computation* (CEUR Workshop Proceedings, Vol. 2347). CEUR-WS.
- Bucher, H.-J. (2017). Understanding multimodal meaning making: Theories of multimodality in the light of reception studies. In J. Wildfeuer & O. Seizov (Eds.), *New studies in multimodality: Conceptual and methodological elaborations* (pp. 91–123). Bloomsbury. <https://doi.org/10.5040/9781350026544.0011>.
- Cambridge University Press. (2006a). *Cambridge idioms dictionary*.
- Cambridge University Press. (2006b). *Cambridge dictionary of American idioms*.
- Cameron, L. (1999). Operationalising ‘metaphor’ for applied linguistic research. In L. Cameron & G. Low (Eds.), *Researching and applying metaphor* (pp. 3–28). Cambridge University Press. <https://doi.org/10.1017/cbo9781139524704.004>.
- Clarke, Y. (2010). *How idioms work: Resource book*. Garnet Education.
- Colmenero, M. O. L. (2024). Subjectivity and creativity versus audio description guidelines. In A. Marcus-Quinn, K. Krejtz, & C. Duarte (Eds.), *Transforming media accessibility in Europe* (pp 39–52). Springer. https://doi.org/10.1007/978-3-031-60049-4_3.
- Cortazzi, M., & Jin, L. (1999). Bridges to learning: Metaphors of teaching, learning and language. In L. Cameron & G. Low (Eds.), *Researching and applying metaphor* (pp. 149–176). Cambridge University Press. <https://doi.org/10.1017/CBO9781139524704.011>.
- DK. (2019). *English for everyone: English idioms*.
- Ershova, Y. S. (2025). Роль эмодзи и визуального языка в цифровой коммуникации на английском языке [The role of emoji and visual language in digital communication in English]. *Вестник науки [Bulletin of Science]*, 5(86), 1045–1062. <https://cyberleninka.ru/article/n/rol-emodzi-i-vizualnogo-yazyka-v-tsifrovoy-kommunikatsii-na-angliyskom-yazyke>.
- Gibbs, R. W., Jr. (1990). Psycholinguistic studies on the conceptual basis of idiomaticity. *Cognitive Linguistics*, 1(4), 417–452. <https://doi.org/10.1515/cogl.1990.1.4.417>.
- Glucksberg, S., Brown, M., & McGlone, M. S. (1993). Conceptual metaphors are not automatically accessed during idiom comprehension. *Memory & Cognition*, 21(6), 711–719. <https://doi.org/10.3758/BF03197201>.
- Goodwin, C. (2000). Practices of seeing: Visual analysis—An ethnomethodological approach. In T. van Leeuwen & C. Jewitt (Eds.), *Handbook of visual analysis* (pp. 157–182). Sage.
- Gorbunova, K. A., & Gorozhanov, A. I. (2025). Typhlo commentary and audio description: Problems of terminological unification. *Philology. Theory & Practice*, 18(5), 1895–1902. <https://doi.org/10.30853/phil20250268>.
- Heidari, K., & Aliyar, M. (2025). Thirty-five years of research on idioms in second language acquisition: A methodological review. *Research Synthesis in Applied Linguistics*, 1(1), 161–183. <https://doi.org/10.1080/29984475.2025.2486966>.
- Isakova, E. (2026). *Metaphors we see by: how to make picture books accessible through audio description*. SSRN, <https://doi.org/10.2139/ssrn.7254398>.
- Isakova, E. M. (2023). Visualization of metaphorical phraseological units: A cognitive approach. *Bulletin of Tver State University. Series: Philology*, 4(79), 108–113. <https://doi.org/10.26456/vtfilol/2023.4.108>.
- Ishchenko, E. N. (2018). Герменевтика визуального в контексте революции в современном искусствознании [Hermeneutics of the visual in the context of the revolution in contemporary art studies]. In *Актуальные проблемы теории и истории искусства—2018: Тезисы докладов VIII Международной конференции [Current issues in theory and history of art—2018: Abstracts of the VIII International Conference]* (pp. 781–783). НП-Принт. <https://elibrary.ru/item.asp?id=44544323>.
- Katz, J. J. (1973). Compositionality, idiomaticity and lexical substitution. In S. Anderson & P. Kiparsky (Eds.), *A festschrift for Morris Halle* (pp. 357–376). Holt, Rinehart and Winston.
- Khadka, T. B. (2025). Integrating verbal and visual signs in meaning making: A semiotic analysis. *Spectrum of Humanities and Social Sciences*, 1(1), 147–158. <https://doi.org/10.3126/shss.v1i1.79838>.
- Kovrov, E. L., & Simonyan, E. G. (2024). Новые формы работы в эпоху визуальной революции [New forms of work in the era of the visual revolution]. In *Передовые достижения науки в молочной отрасли: Сборник научных трудов по результатам работы VI Международной научно-практической конференции, посвящённой дню рождения Николая Васильевича Верещагина [Advanced achievements of science in the dairy industry: Proceedings of the VI International Scientific and Practical Conference dedicated to the birthday of Nikolai V. Vereshchagin]* (pp. 144–147). Вологодская государственная молочнохозяйственная академия имени Н. В. Верещагина. <https://elibrary.ru/item.asp?id=77059895>.
- Lakoff, G., & Johnson, M. (1981). *Metaphors we live by*. University of Chicago Press. <https://doi.org/10.7208/chicago/9780226470993.001.0001>.
- Lertola, J. (2019). *Audiovisual translation in the foreign language classroom: Applications in the teaching of English and other*

- foreign languages. Research-publishing.net. <https://doi.org/10.14705/rpnet.2019.27.9782490057252>.
- Li, G., & Ma, Y. (2025). Exploring the influencing factors of teacher beliefs and their impact on teacher behaviors. *BMC Psychology*, 13(1), 993. <https://doi.org/10.1186/s40359-025-03095-z>.
- Luque Colmenero, M. O., & Soler Gallego, S. (2020). Metaphor as creativity in audio descriptive tours for art museums: From description to practice. *Journal of Audiovisual Translation* 3(2), 64–78. <https://doi.org/10.47476/jat.v3i2.2020.128>.
- Moskalenko, O. A. (2017). Книжка-картинка в западной и отечественной литературной традиции: перевод или перекодирование? [Picturebook in Western and Russian literary traditions: Translation or recoding?]. In *Перевод в современном мире: Сборник трудов IX Международной научно-практической конференции [Translation in the modern world: Proceedings of the IX International Scientific and Practical Conference]* (pp. 50–58). РИБЭСТ. <https://elibrary.ru/item.asp?id=29181245>.
- Myaeots, O. N. (2021). The tale of “The Three Little Pigs”: Evolution of the fairy-tale plot in twentieth-century picturebooks. *Children’s Readings*, 19(1), 215–234. <https://doi.org/10.31860/2304-5817-2021-1-19-215-234>.
- Nadiya, S., Rahmat, H., & Muhammad, M. (2023). The effectiveness of comic strips on the students idiomatic expression mastery. *Cakrawala Bahasa*, 12(1), 1–10. <https://doi.org/10.33387/j.cakrawala.v12i1.5747>.
- O’Reilly, D., Onysko, A., Rasse, C., Papitsch, L., Colston, H., & van der Horst, I. (2025). High ceilings and ingenuine allies: Tapping into the idiom meaning knowledge of first and second language speakers of English. *Language and Cognition*, 17, e72. <https://doi.org/10.1017/langcog.2025.10029>.
- Painter, C., Martin, J. R., & Unsworth, L. (2013). *Reading visual narratives: image analysis of children’s picture books*. Equinox.
- Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford University Press.
- Papitsch, L., Onysko, A., Rasse, C., Colston, H., Kryukova, N., Isakova, E., Aleksandrova, E., & van der Horst, I. (2025). *Putting images of English idioms to the test: L1, L2, and learner ratings on visualization types of idioms*. SSRN. <https://doi.org/10.2139/ssrn.6779282>.
- Paquet, P.-L., Villabona, N., & Gartziaarena, M. (2026). Diving into language teacher beliefs: Experience and contexts in plurilingual classrooms. *International Journal of Applied Linguistics*. <https://doi.org/10.1111/ijal.70287>.
- Peng, X. (2026). Making meaning across modes: A review of theories and multimodal metaphor in media discourse. *US-China Foreign Language*, 24(1), 23–29. <https://www.davidpublisher.com/index.php/Home/Article/index?id=53254.html>.
- Perego, E. (2016). Gains and losses of watching audio described films for sighted viewers. *Target*, 28(3), 424–444. <https://doi.org/10.1075/target.28.3.04per>.
- Ramonda, K. (2016). *The role of semantic transparency and metaphorical elaboration through pictures for learning idioms in a second language* (Doctoral dissertation, University of Birmingham). <http://etheses.bham.ac.uk/6689/>.
- Ramonda, K. (2022). A double-edged sword: Metaphor and metonymy through pictures for learning idioms. *International Review of Applied Linguistics in Language Teaching*, 60(3), 523–561. <https://doi.org/10.1515/iral-2018-0336>.
- Rasse, C., Colston, H., Onysko, A., & Papitsch, L. (in press). *Using images to aid L2 idiom learning: Metaphorical and structural correspondence derivation versus associative learning and dual coding*. https://www.aau.at/wp-content/uploads/2026/06/Using-Images-to-Aid-L2-Idiom-Learning_submitted_vers.pdf.
- Salisbury, M., & Styles, M. (2012). *Children’s picturebooks: The art of visual storytelling*. Laurence King Publishing.
- Samburskiy, D. (2020). The effect of a dual coding technique on idiom interpretation in ESL/EFL learners. *International Journal of Instruction*, 13(3), 187–206. <https://doi.org/10.29333/iji.2020.13313a>.
- Sergeev, D. (2019). What are we reading? Towards a definition of the children’s picturebook. *Children’s Readings*, 16(2), 400–415. <https://doi.org/10.31860/2304-5817-2019-2-16-400-415>.
- Sherif, A. (2022). Translating picture books: Towards a multimodal approach. *Insights into Language, Culture and Communication*, 2(2), 85–96. <https://doi.org/10.21622/ilcc.2022.02.2.085>.
- Soina, A. S. (2017). Особенности перевода книжки-картинки с английского языка [Features of translating the picturebook from English]. In *Перевод в современном мире: Сборник трудов IX Международной научно-практической конференции*, Севастополь, 11–15 июля 2016 года [Translation in the modern world: Proceedings of the IX International Scientific and Practical Conference, Sevastopol, July 11–15, 2016] (pp. 84–88). РИБЭСТ. <https://elibrary.ru/yooxhh>.
- Thyab, R. A. (2016). The necessity of idiomatic expressions to English language learners. *International Journal of English and Literature*, 7(7), 106–111.
- Uspensky, E. N. (1966). *Крокодил Гена и его друзья [Crocodile Gena and his friends]*. Детская литература. <https://ast.ru/book/krokodil-gena-i-ego-druzya-827642/>.
- Vermeulen, A., & Ibáñez Moreno, A. (2017). Audio description as a tool to enhance intercultural competence. In J. Deconinck, P. Humblé, A. Sepp, & H. Stengers (Eds.), *Transcultural competence in translation pedagogy* (Vol. 12, pp. 133–156). LIT Verlag.
- Wu, S. (2014). A multimodal analysis of image-text relations in picture books. *Theory and Practice in Language Studies*, 4(7), 1415–1420. <https://doi.org/10.4304/tpls.4.7.1415-1420>.
- Yarovikova, Y. V. (2019). On the categorical definition of ekphrasis. *Philology: Scientific Studies*, (1), 144–151. <https://doi.org/10.7256/2454-0749.2019.1.28507>.