RUDN Journal of Language Studies, Semiotics and Semantics
2313-2299 (Print), ISSN 2411-1236 (Online)
Вестник Российского университета дружбы народов.
Серия: ТЕОРИЯ ЯЗЫКА. СЕМИОТИКА. СЕМАНТИКА

2024 Vol. 15 No. 3 664–683
http://journals.rudn.ru/semiotics-semantics

DOI: 10.22363/2313-2299-2024-15-3-664-683
EDN: HLEXET
UDC 81'22:791.43.05

Research article / Научная статья

Event Construal through Spatial Relations in Science Documentaries: Language and Image

Nare A. Ovagimian[1], Maria I. Kiose[1,2] ✉

[1]Moscow State Linguistic University, *Moscow, Russian Federation*
[2]Institute of Linguistics RAS, *Moscow, Russian Federation*
✉ maria_kiose@mail.ru

Abstract. Information construal in cinematic discourse employs different semiotic modes; meanwhile, their variance is yet to be explored. The study is aimed at exploring spatial event construal in the language and image modes of science documentaries. We hypothesise that three major event types — environmental events, human-environment interaction events and interpersonal interaction events — employ different spatial construal patterns in language and image, which results from mode allowances and constraints. To verify the hypothesis, we use spatial image schema topology while identifying the image schemas in the lexical and grammatical structure of language and in the layout of the objects, the manner of their movement and interaction in image. The research data include 353 events in single clauses and in shots extracted from two English-language science documentaries. The results show the prevalence of the source-path-goal schema in both semiotic modes, which consequently prevents it from differentiating between event types both within and across the modes. The schemas scale, straight, and near-far display a tendency to differentiate between events; however, no significant distinctions were observed, presumably due to the ontological nature of events as well as the semiotic characteristics of the modes. Additionally, the study reveals that spatial construal of events can follow parallel alignment, commonly with the schemas source-path-goal, contact, near-far, and complementary alignment with the schemas centre-periphery, scale, up-down, front-back, straight, left-right, which reflects a complex nature of inter-semiotic relations in cinematic discourse. The findings of the article contribute to the understanding of event construal in the multimodal discourse of science documentaries.

Keywords: multimodality, semiotic mode, image schema, spatial relations construal, popular science cinematic discourse, event, documentary film

Conflicts of interest: The authors declare no conflicts of interest.

Authors' contribution: The authors contributed equally to this article.

Article history:
Received: 01.06.2024
Accepted: 15.06.2024

For citation:
Ovagimian, N.A. & Kiose, M.I. (2024). Event Construal through Spatial Relations in Science Documentaries: Language and Image. *RUDN Journal of Language Studies, Semiotics and Semantics,* 15(3), 664–683. https://doi.org/10.22363/2313-2299-2024-15-3-664-683

Конструирование событий через призму пространственных отношений в научно-популярном кино: язык и изображение

Н.А. Овагимян[1], М.И. Киосе[1,2] ✉

[1]Московский государственный лингвистический университет, *Москва, Российская Федерация*
[2]Институт языкознания РАН, *Москва, Российская Федерация*
✉ maria_kiose@mail.ru

Аннотация. Конструирование информации в кинодискурсе происходит с участием разных семиотических модальностей, при этом особенности их варьирования изучены слабо. Исследование направлено на изучение особенностей представления событий в терминах пространства в языковой и визуальной модальностях научно-популярных документальных фильмов. Гипотеза исследования заключается в том, что три основные типа события — окружающей среды, взаимодействия человека и окружающей среды, взаимодействия людей между собой — задействуют разные возможности конструирования пространства в языке и изображении, что обусловлено возможностями и ограничениями модальностей. Для проверки гипотезы используется метод анализа образ-схем пространственных отношений, которые определяются в лексической и грамматической семантике языка и в расположении, движении и взаимодействии объектов в изображении. Материалом анализа являются 353 события, соотнесенные с клаузами в языке и с кадрами в изображении в двух англоязычных научно-популярных фильмах. Было установлено, что образ-схема ИСТОЧНИК-ПУТЬ-ЦЕЛЬ превалирует в обоих модусах, что не дает ей возможности значимо различать типы событий внутри и между модусами. Схемы ШКАЛА, ПРЯМО, БЛИЖЕ-ДАЛЬШЕ демонстрируют некоторые тенденции в разграничении типов событий, однако значимые различия не были обнаружены, что определенно обусловлено онтологической природой интерпретируемых событий, как и семиотическими характеристиками самих модусов. Также выявлено, что конструирование пространственных отношений в событиях языка и изображения может происходить параллельно (с использованием образ-схем ИСТОЧНИК-ПУТЬ-ЦЕЛЬ, КОНТАКТ, БЛИЖЕ-ДАЛЬШЕ) и комплементарно (с использованием образ-схем ЦЕНТР-ПЕРИФЕРИЯ, ШКАЛА, ВЕРХ-НИЗ, СПЕРЕДИ-ПОЗАДИ, ПРЯМО, СЛЕВА-СПРАВА), что отражает сложную природу межсемиотических связей в кинодискурсе. Выводы статьи способствуют более глубокому пониманию конструирования событий в мультимодальном дискурсе научно-документальных фильмов.

Ключевые слова: мультимодальность, семиотическая модальность, язык, изображение, образ-схема, конструирование пространственных отношений, научно-популярный кинодискурс, документальный фильм

Заявление о конфликте интересов: Авторы заявляют об отсутствии конфликта интересов.

Вклад авторов: Вклад авторов равнозначен на всех этапах исследования.

История статьи:
Дата поступления: 01.06.2024
Дата приема в печать: 15.06.2024

Для цитирования:
Ovagimian N.A., Kiose M.I. Event Construal through Spatial Relations in Science Documentaries: Language and Image // Вестник Российского университета дружбы народов. Серия: Теория языка. Семиотика. Семантика. 2024. Т. 15. № 3. С. 664–683. https://doi.org/10.22363/2313-2299-2024-15-3-664-683

Introduction

Multimodal studies of cinematic discourse which commonly address the functional structure of semiotic modes [1–3] have lately incorporated the methods of cognitive and discursive paradigms, which have allowed linguists explore the cognitive organisation of cinematic discourse [4–6]. One of the key directions in the cognitive studies is event construal analysis aimed at revealing the cognitive network of referents, their actions and states within the semiotic mode framework in cinematic discourse [7–9]. As shown in [4; 6; 10; 11], event construal is predominantly affected by the spatial relations of referents, which modulate mode alignment, or the synchronised use of semiotic modes. Meanwhile, we expect that spatial relations in cinematic discourse and their manifestations in the semiotic modes of language and image are maintained by different event types in cinematic discourse. Therefore, in this study, we address spatial relations in multimodal popular science documentaries displaying a complex event structure integrating several event types: 1) environmental events (Event Type 1); 2) human-environment interaction events (Event Type 2); and 3) interpersonal interaction events (Event Type 3). Presumably, these three event types display significant differences in presenting the spatial relations of referents, their actions and states in popular science documentaries. Additionally, we expect that their distribution in the semiotic modes of language and image will also be modulated by the event type. Methodologically, the research is grounded in image schema theories, which help explore the event construal process as a series of mental models displaying different degrees of embodiment [12; 13], and also in the discursive theories of spatial relations construal in different semiotic modes of cinematic discourse [4; 6].

The research data are taken from two scientific documentaries in English. The research questions which the paper advances, are primarily the following: 1) What spatial relations schemas are able to differentiate between different event types across both language and image modes? 2) How are spatial relations schemas in both modes related to each other? Following the studies which claim that semiotic modes display different discourse construal potential [14; 11], we hypothesise that i) there are significant differences between the image schemas of spatial relations in construing three types of events in (a) the language mode and (b) the image mode; ii) the language and image modes are related through both parallelism and complementation of image schemas of spatial relations.

The paper is structured as follows: First, we present the Theoretical Framework focusing on the background of analysing spatial relations applying to image schemas

as well as the studies treating inter-semiotic relations in multimodal discourse. Second, we introduce the Methods and Procedure of a multimodal analysis of spatial relations in cinematic discourse. Next, the Results and Discussion are presented, which specify 1) the most common image schemas of spatial relations construing the events in both modes, 2) the patterns and trends in the way image schemas differentiate between the three event types, and 3) the type (s) of relations between the two modes in the way they construe events with image schemas. In the Concluding Remarks section, we identify the research output and the prognostic prospects of its results.

Theoretical Framework

1. Image schemas of spatial relations

Since image schemas are seen as arising from our sensorimotor experience with the world, they are not restricted to any one perceptual mode, being "at once visual, auditory, kinesthetic, and tactile" [12. P. 349]. Image schemas may be viewed as generalisations over perceived similarities [13. P. 526]. Diachronic lexical-semantic changes of a word could be explained by the underlying image schemas [15, P. 107; 16; 12; 17], which direct the lexical-semantic development into extended meanings through metaphorical projections while simultaneously restricting it [17]. This tendency affects not only lexical units but also, importantly, grammatical features [18; 19. P. 325] whereby, for example, in the heart of any preposition lies a highly abstract conceptual structure [20; 21] which in many cases is based on spatial relations [10; 22].

Image schemas are commonly defined through their spatial characteristics and "consist of dynamic spatial patterns that underlie the spatial relations and movement found in actual concrete images." [12. P. 356]. Mandler and Cánovas [13] differentiate between three types of cognitive structure: spatial primitives, image schemas, and schematic integrations, where image schemas are "simple spatial stories" built from spatial primitives as conceptual building blocks and both are used to create concepts with non-spatial elements in schematic integrations [13. P. 510–511]. Image schemas are encoded by spatial representations which have a "geometric (or even quasi-topological)" character [23. P. 9], are not propositional and not limited to a certain point of view. This leads to their inherently relational character. It has been observed that most space schemas establish relations between figure and ground, thus creating such oppositions as near-far or up-down [24–26].

Talmy [26] speaks about material entities revealing three types of spatial properties both in themselves and in relation to other material entities: a) spatial properties of a single object or mass of material in itself; b) spatial properties of one material entity in relation to another; c) spatial properties within a set of material entities as an ensemble [26. P. 436–438]. The first set of properties could be expressed by image schemas of a basic kind, such as object, whereas the third

group of properties is closely associated with image schemas belonging to the group of unity-multiplicity, e.g. part-whole or count-mass. The second type of spatial properties is the one that comes closest to describing the relations construed by the image schemas of spatial relations. Clausner and Croft [27] list the following space image schemas: up-down, front-back, left-right, near-far, centre-periphery, and contact, while the schema path belongs to the group of scalarity [27. P. 15]. Among other schemas, Evans [29] includes straight [28] and verticality, while source-path-goal, in his view, belongs to the group of locomotion [29. P. 108]. Hurtienne and Blessing [30] take Clausner and Croft's list but introduce the schema path into it, as well as the schema scale.

Therefore, for the analysis in this paper, we consider the following most frequently addressed image schemas of spatial relations: source-path-goal, centre-periphery, contact, scale, near-far, up-down, front-back, straight, and left-right.

2.1 Image schemas in different semiotic modes

The multimodal character of image schemas allows for research into their characteristics within and across semiotic modes in a given multimodal discourse. Image schemas have been discussed in relation to visual [e.g. 7; 8], audiovisual discourses [e.g. 31; 32; 33; 5] with particular focus on spatial relations in [6; 4; 10]. The relations between language and image have been studied in terms of logico-semantic relations borrowed from Systemic Functional Grammar [34; 11]. Studying a multimodal discourse (language and static image) through the ideational, interpersonal and textual meta-functions, Royce [14] concludes that the relationship between the two modes is that of inter-semiotic complementarity, which is synergistic in nature [14. P. 27].

Particularly important for this paper is the analysis of multimodal scientific texts by Lemke in [1], where he relates the three meta-functions to the semiotic functions: presentation of a state of affairs (ideational meta-function), construction of an orientational stance towards a state of affairs (interpersonal meta-function), and construction of a system of organisational relations (textual meta-function) [ibid.]. The meanings derived from different semiotic modes in a scientific text may not only be multiplied but are essentially made by the "joint co-deployment" of these modes [ibid. P. 110]. These are partly shaped by the affordances of each semiotic mode, where image represents events in a way that manifests their topological and dynamic character, including their suitability to represent spatial relations, while speech is more appropriate when representing "sequential relations and the making of categorical distinctions." [11. P. 153]. Liu and O'Halloran [3] state that language and images complement each other across the expression, content and context planes. At the same time, the authors introduce two more cohesive mechanisms in a multimodal discourse: inter-semiotic parallelism and inter-semiotic polysemy, where the former "interconnects both language and images when the two semiotic components

share a similar form," and the latter "refers to the cohesive relation between verbal and visual components which share multiple related meanings in multi-semiotic texts." [ibid. P. 372]

The grammar-based approach to the study of image-language relations derives from the extrapolation of the relations between clauses in language to inter-semiotic relations, whereby their relative status may be equal or unequal [35]. In the case of equal status, language and image may be independent, where "there are no signs of one modifying the other," and complementary, where "an image and a text <…> modify one another." [ibid. P. 343] If language and image are unequal, one of them is dependent on the other, resulting in "image subordinate to text" and "text subordinate to image" relations [ibid.]. The discourse-based approach [3] discusses the inter-semiotic relations of comparison, addition, consequence (including consequence and contingency), and time. Martinec and Salway [35] also suggest employing the logico-semantic relations proposed in [34], including expansion (elaboration, extension, enhancement) and projection. Analysing different studies of semiotic characteristics of multimodal texts, Kiose [36. P. 70–71] distinguishes between the following types of alignment patterns between semiotic modes: (a) complementary alignment, where the two systems present different information; (b) parallel alignment, where both systems convey the same information; and/or (c) mutually reinforcing, where the characteristics of one mode increase the activity of the other. Since all the aforementioned studies refer to two distinct mode alignment patterns, parallelism and complementation, in this study we will adhere to this generalised view to explore the inter-semiotic relations between the language and image modes.

Multimodal study of spatial relations in film: Methods and procedure

To identify the distribution of image schemas of spatial relations in the construal of three types of events, we followed a series of steps. The research data were extracted from exposition scenes in two science documentaries, "Pop! The Science of Bubbles" (2013) and "Secrets of the Super Elements" (2017), comprising 194 clauses (286 autosemantic words, each of them tagged) manifesting 194 events, and 274 synchronised shots manifesting 159 events.

In Step 1, we identified the types of events in language and image based on the nature of entities and their interaction. Environmental events (Event Type 1) involve objects of the world, their relation and interaction as detached from an observer; 2) human-environment interaction events (Event Type 2) describe interactions between man as a cogniser and the environment as a cognised entity; and 3) interpersonal interaction events (Event Type 3) deal with events occurring in communicative situations between people. To identify the event type, we considered the components of each clause and each shot.

In Step 2, we explored event construal in the language mode, including an etymological analysis of each autosemantic word within a clause, which allowed us to identify the diachronically traced meanings of their composite parts (root and affixes). This approach is based on the assumption that the overall meaning is motivated by the sum of the meanings of the constituent parts as well as the emergent meaning. Step 3 involved an analysis of the relevant definitions from modern dictionaries as well as the contextual meanings of words, which allowed us to specify the image schemas.

In Step 4, we determined the image schemas of spatial relations in the image mode, where the main focus was placed on the layout of the object(s) on the screen, including the manner of their movement and interaction with each other.

In Step 5, the synchronised units from language and image (with the identified schemas) were put together to conduct statistical analyses, which allowed us to determine the tendencies in event construal and patterns of alignment between the two modes. The initial results were processed using the HETEROSTAT software [37] to provide quantitative data for further analysis. It should be pointed out that the quantitative data were subjected to analyses in both total co-occurrences (reflecting the overall number of instances) and in pairwise co-occurrences (marking only the presence/absence of a given image schema in a clause or shot) using Chi-square tests.

Below is an example of the procedure.

Step 1. The clause *'These little things are full of secrets and surprises'* (Pop! The Science of Bubbles, 2017) presents an environmental event and a human-environment interaction event in language, accompanied in image by a depiction of the presenter/scientist alternately looking at the object of her research (bubbles) and at the camera, thus exemplifying all three types of events: an environmental event (bubbles moving up in a liquid), an interpersonal interaction event (a scientist looking at the object of their research), and a human-environment interaction event (a presenter/scientist looking at the camera, thus addressing the viewer).

Steps 2 and 3. The events are construed through the following image schemas in the language mode:

1. The adjective ***little***, meaning 'not big; small; smaller than others', has its origins in OE *lȳtel*, related to *lȳt* 'little or few' from Proto-Germanic **lūti*; the word is ultimately related to IE **leud* 'to depress, reduce' thus making apparent the meaning of scalarity (image schema scale)[1], [2].
2. The noun ***thing*** is defined as 'an object whose name you do not use because you do not need to or want to, or because you do not know it; an object that is not alive in the way that people and plants are'[3]. In the definition

[1] Onions, C.T., Friedrichsen, G.W.S. & Burchfield, R.W. (Eds.). (1966). The Oxford dictionary of English etymology. Vol. 178. Oxford: Clarendon Press. P. 532.

[2] Barnhart, R.K. (Ed.). (1988). Barnhart Dictionary of Etymology: H.W. Wilson Co., 1988. P. 603.

[3] Oxford Learners' Dictionaries. URL: https://www.oxfordlearnersdictionaries.com/definition/english/thing?q=thing (accessed: 28.03.2024).

of the Cambridge Dictionary, the meaning of approximation is emphasised, as in 'used to refer in an approximate way to an object or to avoid naming it' or 'used to refer in an approximate way to an idea, subject, event, action, etc.'[4] Etymological entries for the word ***thing*** trace its roots to OE *þing* meaning 'meeting, assembly' (now obsolete), which developed through conceptual metonymy into 'entity, being, matter' ('subject of deliberation an assembly'), as well as 'act, deed, event'[5], [6],[7]. There have also been suggestions that the word may have meant 'day of assembly', from a base meaning 'stretch or extent of time', related to IE **tenk* 'to draw out, or draw together'[8], [9]. The image schemas evoked by the word *thing* are contact, centre-periphery, and source-path-goal.

3. OALD provides the following relevant definitions for the adjective ***full***: 'containing or holding as much or as many as possible; having no empty space; having or containing a large number or amount of something/somebody; to the highest level or greatest amount possible'[10]. Etymologically, the word derives from OE *full*, ultimately related to IE **pl̄nós* from IE **pol-, *pel-, *pl-* in words expressing fulness or abundance[11], [12], [13]. The image schemas to be distinguished here are centre-periphery and scale.
4. The Oxford Dictionary of English Etymology gives the following definitions of the preposition ***of***: "expressing removal, separation, derivation, origin, source, spring of action, point of departure in time, cause, agent, instrument, material."[14] The preposition comes from OE *of*, from CGerm adverb and preposition **ab (a)* deriving from IE **ap*, **apo* meaning 'away from, down from'[15]. In modern English, the first meaning attributed to this preposition

[4] Cambridge Dictionary. URL: https://dictionary.cambridge.org/dictionary/english/thing (accessed: 28.03.2024).

[5] Onions, C.T., Friedrichsen, G.W.S. & Burchfield, R.W. (Eds.). (1966). The Oxford dictionary of English etymology. Vol. 178. Oxford: Clarendon Press. P. 917.

[6] Barnhart, R.K. (Ed.). (1988). Barnhart Dictionary of Etymology: H.W. Wilson Co.P. 1134.

[7] Online Etymology Dictionary. URL: https://www.etymonline.com/search?q=thing (accessed: 28.03.2024).

[8] Barnhart, R.K. (Ed.). (1988). Barnhart Dictionary of Etymology: H.W. Wilson Co.P. 1134.

[9] Online Etymology Dictionary. URL: https://www.etymonline.com/search?q=thing (accessed: 28.03.2024).

[10] Oxford Learners' Dictionaries. URL: https://www.oxfordlearnersdictionaries.com/definition/english/full_1?q=full (accessed: 28.03.2024).

[11] Onions, C.T., Friedrichsen, G.W.S., & Burchfield, R.W. (Eds.). (1966). The Oxford dictionary of English etymology. Vol. 178. Oxford: Clarendon Press. P. 380.

[12] Barnhart, R.K. (Ed.). (1988). Barnhart Dictionary of Etymology: H.W. Wilson Co.P. 413.

[13] Online Etymology Dictionary. URL: https://www.etymonline.com/search?q=full (accessed: 28.03.2024).?

[14] Onions, C.T., Friedrichsen, G.W.S., & Burchfield, R.W. (Eds.). (1966). The Oxford dictionary of English etymology. Vol. 178. Oxford: Clarendon Press. P. 624.

[15] ibid.

is that of possession, belonging, or origin[16], [17]. Another meaning based on space relations is the following: "used to show the position of something/somebody in space or time"[18]; "used in expressions showing position"[19]. Here we identified the image schemas source-path-goal, centre-periphery, near-far.

5. The noun ***secret*** defined as 'known about by only a few people; kept hidden from others; used to describe actions and behaviour that you do not tell other people about'[20] comes from (O)F *secret*, further related to L. *sēcrētus* from *sē* 'without, apart' and *cernere* 'to separate', together meaning 'to separate, distinguish, secrete', 'set apart, hidden'[21], [22]. The word expresses the following space image schemas: source-path-goal, centre-periphery, near-far.
6. OALD gives the following relevant definition for the word ***surprise*** as a noun: 'an event, a piece of news, etc. that is unexpected or that happens suddenly'[23]. The word is derived from the noun use of the past participle of OF *surprendre* meaning 'to overtake' (from *sur* 'over' + prendre 'to take') from L. *super* ('over') + *prændere* contracted from *prehendere* 'to grasp, seize' from *prae-* 'before' + *-hendere* from PIE root ***ghend-* 'to seize, take'[24], [25],[26]. The schemas source-path-goal, centre-periphery, contact, and up-down lie at the base of the meaning of the word *surprise.*

Step 4. The clause is synchronised in the image mode with the images of the presenter/scientist, evoking the schemas source-path-goal (movement of the presenter/scientist in the foreground and the bubbles in the background), up-down (the bubbles moving up), front-back (the presenter/scientist looking back and forth), and straight (the straight path of the bubbles moving; the straight look of the presenter/scientist at the camera) (Fig. 1).

[16] Cambridge Dictionary. URL: https://dictionary.cambridge.org/dictionary/english/of (accessed: 28.03.2024).
[17] Oxford Learners' Dictionaries. URL: https://www.oxfordlearnersdictionaries.com/definition/english/of?q=of (accessed: 28.03.2024).
[18] ibid.
[19] Cambridge Dictionary. URL: https://dictionary.cambridge.org/dictionary/english/of (accessed: 28.03.2024).
[20] Oxford Learners' Dictionaries. URL: https://www.oxfordlearnersdictionaries.com/definition/english/secret_1?q=secret (accessed: 28.03.2024).
[21] Onions, C.T., Friedrichsen, G.W.S., & Burchfield, R.W. (Eds.). (1966). The Oxford dictionary of English etymology. Vol. 178. Oxford: Clarendon Press. P. 805.
[22] Barnhart, R.K. (Ed.). (1988). Barnhart Dictionary of Etymology, H.W. Wilson Co.P. 977.
[23] Oxford Learners' Dictionaries. URL: https://www.oxfordlearnersdictionaries.com/definition/english/surprise_1?q=surprise (accessed: 28.03.2024).
[24] Onions, C.T., Friedrichsen, G.W.S. & Burchfield, R.W. (Eds.). (1966). The Oxford dictionary of English etymology. Vol. 178. Oxford: Clarendon Press. P. 890.
[25] Barnhart, R.K. (Ed.). (1988). Barnhart Dictionary of Etymology: H.W. Wilson Co.P. 1096.
[26] Online Etymology Dictionary. URL: https://www.etymonline.com/search?q=surprise (accessed: 28.03.2024). Почему здесь эти сноски? Почему не в списке литературы?

Fig. 1. Fragment 1
Source: photos from the archiveof Nare A. Ovagimian & Maria I. Kiose. Prior to filming the participants signed the consent form.

Step 5. This example demonstrates how parallel relations are established with the use of the image schemas source-path-goal and up-down while the rest of the schemas lead us to attest the complementarity of image schemas in the two modes.

Results and Discussion

Results

The analysed multimodal corpus comprised 353 events which included 194 events presented in clauses in the language mode, each synchronising with events in one or more shots in the image mode with unique shorts numbering 159. The clauses and shots are established as units expressing one event at a time, i.e. if a clause presented two types of events at the same time, it was repeated twice for each event.

The distribution of spatial relations in three event types

Language

In Fig. 2, we present the data on the distribution of spatial relations in the language mode.

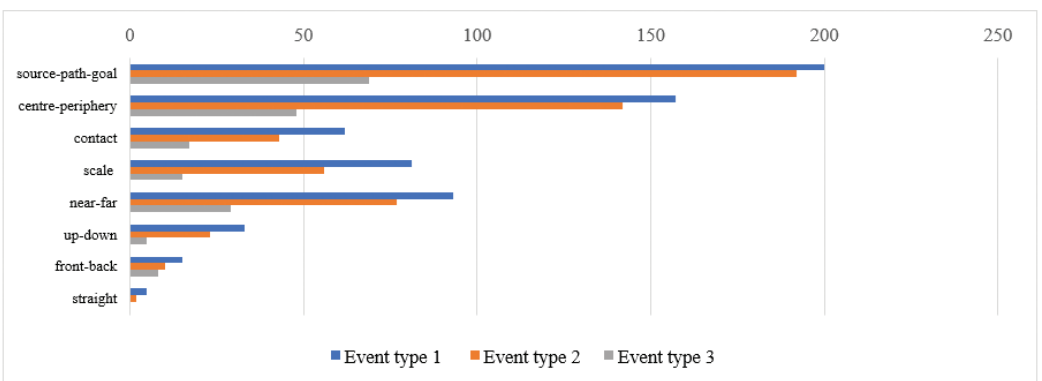

Fig. 2. Distribution of Event Types among Image Schemas in Language
Source: compiled by Nare A. Ovagimian & Maria I. Kiose.

As is shown in Fig. 2, across all the schemas of spatial relations, the most commonly construed type of event in the language mode is Event Type 1, with Event Type 2 coming next, while Event Type 3 is considerably less frequently construed with the given schemas. Another notable feature is that several schemas (e.g. source-path-goal, centre-periphery, etc.) are noticeably more commonly used than others (e.g. front-back, straight, etc.). Interestingly, no component of the analysed verbal material is used to express the schema left-right. In the following clauses, the two most commonly used schemas, i.e. source-path-goal and centre-periphery are simultaneously instantiated in the words *breaking, drag, times, of, welcome*: a) *… as breaking ways drag air under water.* (Event Type 1); b) *There's 300 times more gold in a kilo of smartphones than a kilo of gold ore.* (Event Type 1); c) *Welcome to my world…* (Event Type 3).

Image

In Fig. 3, we display the data on the distribution of spatial relations in the image mode.

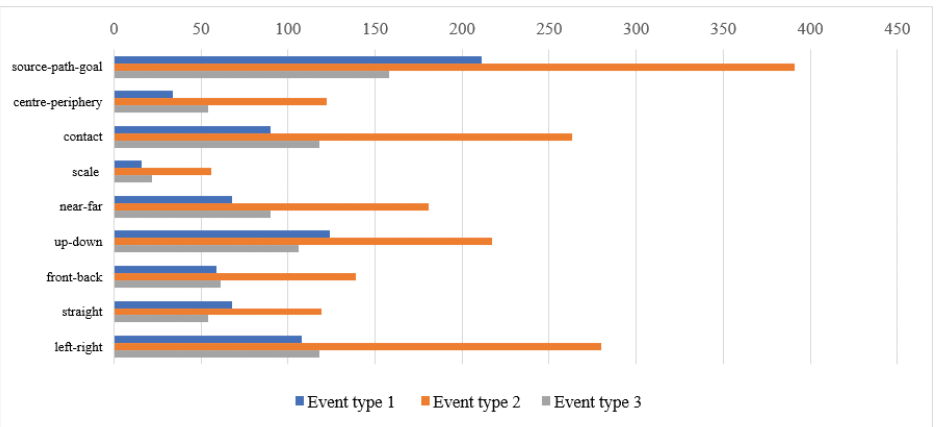

Fig. 3. Distribution of Event Types between Image Schemas in Image
Source: compiled by Nare A. Ovagimian & Maria I. Kiose.

In the image mode, the most commonly construed type of event is Event Type 2, while Event Type 1 comes second in the schemas source-path-goal, up-down, and straight, and Event Type 3 — in the rest of the schemas (Fig. 3). For instance, in Fragment 2 (Fig. 4) of an event pertaining to two event types, namely Event Type 2 and Event Type 3, the presenter/scientist interacts with a metal ball by holding (contact) and revolving (source-path-goal, left-right) it in his hands, then raising it (source-path-goal, up-down).

The construal of the three event types in language does not show any significant differences, i.e. no one image schema is more capable of significantly differentiating between event types than another. However, a certain tendency

to such distinction[27] can be observed in the image schema scale when differentiating between Event Type 1 and Event Type 3 (χ^2= 3.187, p=.075) and the image schema front-back when differentiating between Event Type 2 and Event Type 3 (χ^2=3.284, p=.070).

Fig. 4. Fragment 2
Source: photos from the archive of Nare A. Ovagimian & Maria I. Kiose. Prior to filming the participants signed the consent form.

Image schemas of spatial relations differentiating between event types

Total co-occurrences

A contrasting feature is observed in the image mode where the schemas source-path-goal, centre-periphery, contact have demonstrated their capability to differentiate between Event Type 1 and Event Type 2 as well as Event Type 1 and Event Type 3. The two illustrations in Fig. 5 show an instance of Event Type 2: a woman is holding a smartphone in one hand while she taps on it with the fingers of the other hand (source-path-goal, contact, near-far; Fig. 5, *a*), then moves her hand closer to herself (source-path-goal, centre-periphery, near-far, front-back; Fig. 5, *b*). In the layout of the shot, the movement traces the trajectory from left to right (left-right).

a *b*

Fig. 5. Fragment 3
Source: photos from the archive of Nare A. Ovagimian & Maria I. Kiose. Prior to filming the participants signed the consent form.

[27] With p-values not indicating the presence of significant differences but being sufficiently close to the threshold of significance equal to .05.

The other schemas either do not show any difference (scale, near-far, front-back, straight, left-right) or display it only in contrasting two event types (e.g. the image schema up-down in Event Type 1 and Event Type 2). Two schemas, straight and near-far, reveal a tendency towards differentiating between Event Type 1 and Event Type 2 (straight, χ^2=3.206, p=.074) and between Event Type 1 and Event Type 3 (near-far, χ^2= 3.315, p=.069). Interestingly, no schema has been found to be able to differentiate between Event Type 2 and Event Type 3.

Pairwise co-occurrences

When examining the occurrence of schemas only once per clause, we analysed how more likely a schema is to be used to construe one event rather than another. As in the previous analysis, no image schema has been found capable of differentiating between Event Type 1 and Event Type 2. The schema front-back does not show any difference between any of the pairs of events, although there is a trend towards it in Event Type 1 vs. Event Type 3 (χ^2=3.270, p=.071). Another tendency has been found in the image schema contact differentiating between Event Type 1 and Event Type 2 (χ^2=3.307, p=.070). However, all image schemas differ significantly in how they differentiate between Event Type 1 and Event Type 3. In another opposition, Event Type 2 and Event Type 3 are differentiated in all but two schemas (front-back and straight).

In the next three clauses exemplifying Event Type 1 in (a) and Event Type 3 in (b) and (c), we may see an overlap of image schemas: a) ¾ *of* (source-path-goal, centre-periphery, near-far) *the earth is covered* (source-path-goal, contact, up-down) *in* (source-path-goal, centre-periphery) *water* (Event Type 1); b) *Let* (source-path-goal) *me start* (source-path-goal, up-down) *with* (source-path-goal, centre-periphery, contact, near-far) *a simple question* (source-path-goal) (Event Type 3); c) *So, what are these superelements* (scale, up-down)? (Event Type 3). Interestingly, the construal of the event in the third clause does not include such a prevalent image schema as source-path-goal.

Multimodal alignment: Modes and image schemas of spatial relations

The analysis of the alignment of the two semiotic modes by means of image schemas of spatial relations with the primary focus on the language mode shows a similar pattern for most schemas, as shown in Fig. 6.

In every image schema in the language mode, the most commonly co-occurring schema in the image mode is source-path-goal (varying from 23 to 38% with the mode and median being 25%) followed by up-down and left-right (12–20%) as well as contact (mode and median: 14%). Taking into account the fact that image schemas very rarely occur alone, especially in the image mode, the following alignment models were identified: (a) the image schema source-path-goal is used simultaneously in both modes in 25% of co-occurrences; (b) a certain tendency towards expressing the same conceptual information is demonstrated by the schemas contact (14%) and

up-down (14%). In general, the main model of alignment between the two modes is that of complementation, with the range of 75–98% (median 91%, mode 86%). In other words, complementary relations between the language and image modes are more common than relations of parallelism. Three image schemas more commonly occur simultaneously in the analysed data: source-path-goal (86%), contact (78%), and near-far (75%), while scale and straight are used to construe events at the same time in 29 and 11% of cases respectively.

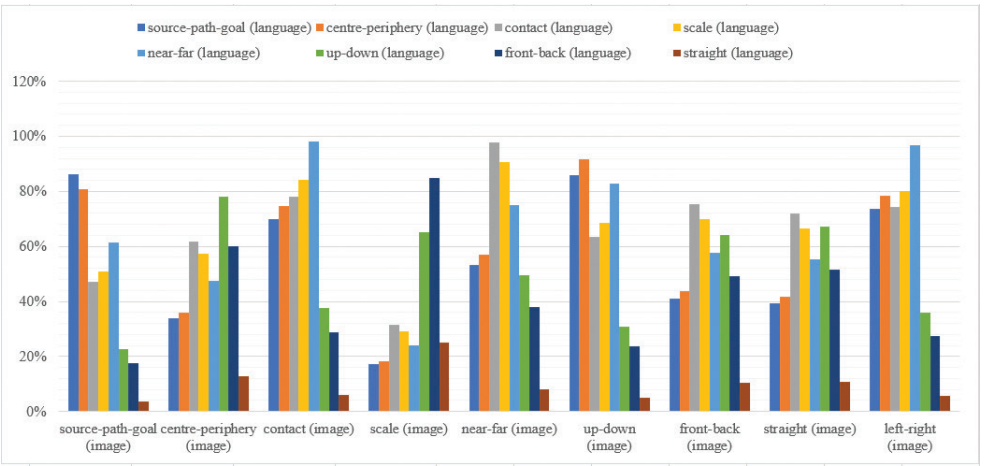

Fig. 6. Co-occurrence of image schemas of spatial relations in language and image, %
Source: compiled by Nare A. Ovagimian & Maria I. Kiose.

The clause below presents an example of Event Type 2 which is construed with the help of the following image schemas of spatial relations: *You can* (source-path-goal) *even* (scale) /pause/ *film* (source-path-goal) *a whole* (centre-periphery, scale) *BBC documentary on* (source-path-goal, contact, near-far) *them.* The clause is synchronised in the image mode with an instance of Event Type 3 (Fig. 7), where the presenter/scientist gets closer to the camera (source-path-goal, near-far) and moves it to his left to show the filming crew (source-path-goal, contact, left-right), which incidentally involves the camera shifting upwards (up-down).

Fig. 7. Fragment 4
Source: photos from the archive ofNare A. Ovagimian & Maria I. Kiose. Prior to filming the participants signed the consent form.

All in all, in a quarter of co-occurrences between the schemas in language and image, they are used in more than 75% of clauses, whereas in another quarter of cases, they occur together in fewer than 30% of clauses. On the other hand, the image schema straight is extremely rare in the analysed data, with only 7 instances in 194 clauses, while it is used 189 times in image. Still, the image schema scale is more commonly associated with the language mode, helping convey the idea of comparison, growth, etc., while in the image mode it is mainly used when zooming in or out or between the shots.

Discussion

In the language mode, across all the schemas of spatial relations, we can discern a hierarchy of frequency of construed event types, where the subject matter of a film (Event Type 1) takes centre stage while the interaction between a cogniser and the environment (Event Type 2) occupies the medium position between Event Type 1 and Event Type 3, which iconically represents the transitory character of the distinction between the three event types. The absence of the schema left-right in the language mode may be explained by the affordances of this semiotic mode, while in the image mode the left-right distinction could be considered part of its epistemological commitment [2].

A slightly different hierarchy found in the image mode with every schema more commonly construing Event Type 2 rather than the other event types, on the one hand, reflects the structural properties of the image mode, where the nature of the medium presenting a dynamic picture necessitates the change of position of an object on the screen, allowing for at least three schemas: source-path-goal, left-right, and up-down. On the other hand, the pervasive presence in the image mode of specialists, scientists as well as scientific tools and their interaction accounts for the frequency of Event Type 2 and it being construed by the schemas of spatial relations, especially contact and near-far.

In the language mode, the inability of any schemas to differentiate between the events' types (in total co-occurrences) could be accounted for by the fact that the order of preference/frequency for each schema to represent an event type is the same: environmental event → human-environment interaction event → interpersonal interaction event. The absence of any significant distinction between Event Type 2 and Event Type 3 construed by all schemas of spatial relations in image may be due to the fact that many events instantiating these two types are actually realised in the same shot, e.g. the presenter/scientist manipulating an object and at the same time directing his view at the viewer, which results in the events being construed in the same way. These results specify the role of the language and image modes [4; 6; 10] as contingent on event types.

The pairwise co-occurrences in the language mode showing that no schema is capable of differentiating between Event Type 1 and Event Type 2 prove that the environmental event and the human-environment interaction event are more closely

associated with each other than with the interpersonal interaction event, which could explain their frequent co-occurrence. The relations of parallelism between the two semiotic modes are most commonly established in the way they employ the schemas source-path-goal, contact, and near-far, which, incidentally, are realised 44 times in the language mode and 58 times in image. These findings confirm the significance of these schemas in event construal outlined in [24–27; 30]. In the predominant number of cases of language-image co-occurrences, the synchronised units construe events with different schemas, which is in accordance with the idea of inter-semiotic complementarity [1; 14].

These patterns suggest parallel and complementary alignment models [35; 36] between the two modes: they are parallel in the schemas naturally evoked by the subject matter of the films, where the main focus is on the interaction between objects, whether they both be in the realm of environmental events or in the realm of the other two types of events, thus necessitating conceptualisation through a certain trajectory and contact between the object(s) and the subject(s). The complementary relations between language and image reflect the way information is presented in the initial excerpts of popular science documentaries: most commonly, the type of objects and their interaction depicted in image are only vaguely related to the information evoked in the language mode. Usually, linguistically, the film addresses general questions it promises to cover (e.g. *What is a bubble?, Why do we need them* (superelements) *so badly?*), it gives observations about our daily lives (e.g. *We're learning that bubbles influence our world in all sorts of unexpected ways.*) referring to problems we (might) face (e.g. *We are reaching the limits of what our planet can provide.*). This unspecific, generic information cannot be presented in the image mode, which, owing to its epistemological commitment, is necessarily more concrete and specific: it predominantly shows a particular bubble, person, or lump of met al. This difference in presenting various types of referents is reflected in the way they are construed with image schemas of spatial relations, which results in the complementary relations between the two semiotic modes.

Concluding remarks

The research presents an analysis of image schemas of spatial relations as forming the basis of a multimodal construal of events belonging to three ontological types: environmental events, human-environment interaction events, and interpersonal interaction events. The study of popular scientific documentary films in English has enabled us to observe how schemas are realised in the language and image modes, including their ability to differentiate between the event types and determine inter-semiotic relations between the two modes. The schemas of spatial relations are necessarily invoked in image since a depicted object is always placed within the spatial constraints of the screen and

in the constraints imposed by other objects. In language, the schemas of spatial relations are at the core of the lexical-semantic development of many linguistic units (diachronic aspect) and may be evoked through the analysis of their definitions and contextual meanings (synchronic aspect). The study of inter-semiotic relations helped to determine how the conceptual structures of event construal in language and image were related to each other, with the main opposition being parallelism vs. complementarity. To answer these questions, a five-step procedure was developed, which included the identification of the event types and the schemas of spatial relations in both modes.

The results show that in scientific documentaries, the image schema source-path-goal is more pervasive than the others, whose (in) ability to differentiate between event types is determined by the mode and its affordances as well as by the types of events themselves. The most commonly used schemas (source-path-goal, contact, near-far) frequently co-occur, thus resulting in the relations of parallelism between the language and image modes, while the others (centre-periphery, scale, up-down, front-back, straight, left-right) do not appear in language and image at the same time as often, which suggests inter-semiotic complementarity. The results allow us to specify the way people construe information about different types of events, especially in such a complex type of discourse as that of popular science documentaries. Hopefully, the suggested procedure may be applicable in determining the characteristics of event construal with the means of different semiotic modes in other types of discourse.

References / Библиографический список

1. Lemke, J. (1998). Multiplying meaning: Visual and verbal semiotics in scientific text. In: Martin J.R., Veel R. (Eds.), *Reading science*: *Critical and functional perspectives on discourses of science*. London: Routledge. pp. 87–114.
2. Bezemer, J. & Kress, G. (2008). Writing in multimodal texts: A social semiotic account of designs for learning. *Written Communication*, 25(2), 166–195. https://doi.org/10.1177/0741088307313177
3. Liu, Y. & O'Halloran, K.L. (2009). Intersemiotic texture: Analyzing cohesive devices between language and images. *Social Semiotics*, 19(4), 367–388. https://doi.org/10.1080/10350330903361059
4. Forceville, C. & Jeulink, M. (2011). The flesh and blood of embodied understanding: The source-path-goal schema in animation film. *Pragmatics & Cognition*, 19(1), 37–59. https://doi.org/10.1075/pc.19.1.02for
5. Forceville, C.J. (2017). From image schema to metaphor in discourse: The force schemas in animation films. In: Hampe B. (Ed.) *Metaphor*: *Embodied Cognition and Discourse*. Cambridge: Cambridge University Press. pp. 239—256. https://doi.org/10.1017/9781108182324.014
6. Coëgnaerts, M. & Kravanja, P. (2012). The visual and multimodal representation of time in film, or: How time is metaphorically shaped in space. *Image and narrative online magazine of the visual narrative*, 13(3), 85–100.
7. Lou, A. (2017). Viewpoint and image schemas in multimodal political discourse. *Journal of Pragmatics*, 122, 77–90. https://doi.org/10.1016/j.pragma.2017.07.002

8. Khan, S. & Siddiqui, B. (2023). Space Image Schema in multimodal representation of Afghan migrants living in Pakistan: a cognitive linguistics approach to multimodal discourse analysis. *Journal of Positive School Psychology*, 7(4), 1920–1948.
9. Kiose, M.I. (2023). Image schemas as an instrument of multimodal research. *Germanistika-2022*: *nove et nova*. *Materials of the V International Scientific Conference, Moscow, 07–09 December 2022*, A.A. Klinovskaya (Ed.) [et al.]. Moscow: Moscow State Linguistic University. pp. 136–140. (In Russ.).
Киосе М.И. Теории образ-схем как инструмент полимодальных исследований // Германистика-2022: nove et nova: материалы V Международной научной конференции, Москва, 07–09 декабря 2022 года / редколл.: А.А. Клиновская [и др.]. М.: Московский государственный лингвистический университет, 2023. С. 136–140.
10. Deane, P.D. (2005). Multimodal spatial representation: On the semantic unity of over. *From perception to meaning*: *Image schemas in cognitive linguistics*, 29, 235–282. https://doi.org/10.1515/9783110197532.3.235
11. Unsworth, L. & Cléirigh, C. (2014). *Multimodality and reading*: *The construction of meaning through image-text interaction*. London: Routledge.
12. Gibbs, R. & Colston, H. (1995). The cognitive psychological reality of image schemas and their transformations. *Cognitive Linguistics*, 6(4), 347–378. https://doi.org/10.1515/cogl.1995.6.4.347
13. Mandler, J.M. & Cánovas, C.P. (2014). On defining image schemas. *Language and Cognition*, 6(4), 510–532. https://doi.org/10.1017/langcog.2014.14
14. Royce, T. (1998). Synergy on the page: Exploring intersemiotic complementarity in page-based multimodal text. *JASFL Occasional papers*, 1(1), 25–49.
15. Johnson, M. (1987). *The body in the mind*: *The bodily basis of meaning, imagination, and reason.* Chicago: University of Chicago Press. https://doi.org/10.7208/chicago/9780226177847.001.0001
16. Sweetser, E. (1990). *From etymology to pragmatics*: *metaphorical and cultural aspects of semantic structure.* New York: Cambridge University Press. https://doi.org/10.1017/cbo9780511620904
17. Ioannou, G. (2020). Image schemas as prototypes in the diachronic evolution of kámnō and eutheiázō in Greek: A behavioural-profile analysis. *Lingua*, 245(4), 102938. https://doi.org/10.1016/j.lingua.2020.102938
18. Langacker, R. (1987). *Foundations of cognitive grammar*: *Volume I*: *Theoretical prerequisites*. Stanford: Stanford University Press.
19. Langacker, R. (1991). *Concept, Image, and Symbol*: *The Cognitive Basis of Grammar*. Berlin; New York: Mouton de Gruyter.
20. Brugman, C. & Lakoff, G. (1988). Cognitive Topology and Lexical Networks. *Lexical Ambiguity Resolution*, 477–508. https://doi.org/10.1016/b978-0-08-051013-2.50022-7
21. Vandeloise, C. (1991). *Spatial prepositions*: *A case study from French*. Chicago: University of Chicago Press.
22. Talmy, L. (2005). The fundamental system of spatial schemas in language. *From perception to meaning*: *Image schemas in cognitive linguistics*, 29, 199–234. https://doi.org/10.1515/9783110197532.3.199
23. Jackendoff, R. (1996). The architecture of the linguistic-spatial interface. In: P. Bloom, M.A. Peterson, L. Nadel & M.F. Garrett (Eds.), *Language and Space*. Mass.: The MIT Press. pp. 1–30. https://doi.org/10.7551/mitpress/4107.003.0003
24. Levinson, S.C. (1996). Frames of References and Molyneux's Question: Crosslinguistic Evidence. In: P. Bloom, M.A. Peterson, L. Nadel & M.F. Garrett (Eds.), *Language and Space*. Mass.: The MIT Press. pp. 109–170. https://doi.org/10.7551/mitpress/4107.003.0006
25. Cienki, A. (1997). Some properties and groupings of image schemas. *Lexical and syntactical constructions and the construction of meaning*, 3–15. https://doi.org/10.1075/cilt.150.04cie

26. Talmy, L. (2000). *Toward a cognitive semantics. Vol. II*: *Typology and process in concept structuring*. Mass.: The MIT Press.
27. Clausner, T.C. & Croft, W. (1999). Domains and image schemas. *Cognitive Linguistics*, 10(1), 1–31. https://doi.org/10.1515/cogl.1999.001
28. Cienki, A. (1998). STRAIGHT: An image schema and its metaphorical extensions. *Cognitive Linguistics*, 9(2), 107–150. https://doi.org/10.1515/cogl.1998.9.2.107
29. Evans, V. (2007). *Glossary of cognitive linguistics*. Edinburgh: Edinburgh University Press.
30. Hurtienne, J. & Blessing, L. (2007). Design for Intuitive Use-Testing image schema theory for user interface design. In: *Proceedings of ICED 2007, the 16th International Conference on Engineering Design, Paris, France, 28–31.07. 2007*. Paris. pp. 829–830.
31. Zykova, I.V. (Ed.). (2021). *Linguistic creativity in different discourse types*: *Limitations and allowances*. Moscow: R. Valent.
Зыкова И.В. (ред.) Лингвокреативность в дискурсах разных типов: Пределы и возможности. М.: Р-Валент, 2021.
32. Kiose, M.I. (2022). Event Construal in Text and Image. Gaze Behavior and Discourse Interpretations. *RUDN Journal of Language Studies, Semiotics and Semantics*, 13(2), 396–415. https:// doi.org/10.22363/2313-2299-2022-13-2-396-415
33. Maksimenko, O.I. & Podlegaeva, E.P. (2023). *Animated videoverbal text*: *problems of translation and interpretation*. Moscow: Nauka. (In Russ.).
Максименко О.И., Подлегаева Е.П. Анимационный видеовербальный текст: проблемы перевода и интерпретации. М.: Наука, 2023.
34. Halliday, M. & Matthiessen, C. (2004). *An Introduction to Functional Grammar* (3rd ed.). London: Routledge. https://doi.org/10.4324/9780203783771
35. Martinec, R. & Salway, A. (2005). A system for image–text relations in new (and old) media. *Visual communication*, 4(3), 337–371. https://doi.org/10.1177/1470357205055928
36. Kiose, M.I. (2021). *Secrets of interpretation of text and image. Construal and oculographic experiment*. Moscow: R.Valent. (In Russ.).
Киосе М.И. Секреты интерпретации текста и изображения. Конструирование и окулографический эксперимент. М.: Р-Валент, 2021.
37. Kiose, M.I. & Efremov, A.A. (2020). HETEROSTAT. *Software for Comprehensive Parametrical Discourse Analysis*. Registration Date: 21.09.2020. Registration Number: 2020661240. 2020.

Information about the authors:

Nare A. Ovagimian, Senior Lecturer at the Department of English, Faculty of Translation and Interpreting, Moscow State Linguistic University (38, Ostozhenka str., Moscow, Russian Federation, 119034); *Research interests*: cognitive semantics, cognitive grammar, scientific documentary discourse, multimodal studies, English grammar; *e-mail*: n.ovagimian@yandex.ru
ORCID: 0009-0004-1981-0051; SPIN-code: 2035-7040; AuthorID: 926428.

Maria I. Kiose, Dr.Sc. (Philology) (Advanced Doctorate), Associate Professor, Chief Researcher of the Centre for Socio-Cognitive Studies of Moscow State Linguistic University (38, Ostozhenka str., Moscow, Russian Federation, 119034); Leading Researcher; Laboratory for multichannel communication, Institute of Linguistics, RAS (1, B. Kislovsky ln., Moscow, Russian Federation, 125009); *Research interests*: cognitive semantics, indirect naming in text, referential semantics, oculographic and corpus-based research of text inference and generation process; *e-mail*: maria_kiose@mail.ru
ORCID: 0000-0001-7215-0604; Scopus AuthorID: 56642747500; ResearcherID AAB-7989-2019; SPIN-code: 4419-0090; AuthorID: 341964.

Сведения об авторах:

Овагимян Наре Армени, старший преподаватель кафедры английского языка переводческого факультета, Московский государственный лингвистический университет (119034, Россия, г. Москва, ул. Остоженка, 38); *сфера научных интересов*: когнитивная семантика, когнитивная грамматика, научно-популярный кинодискурс, полимодальные исследования, грамматика английского языка; *e-mail*: n.ovagimian@yandex.ru
ORCID: 0009-0004-1981-0051; SPIN-код: 2035-7040; AuthorID: 926428.

Киосе Мария Ивановна, доктор филологических наук, доцент, главный научный сотрудник Центра Социокогнитивных исследований дискурса Московского государственного лингвистического университета (119034, Российская Федерация, г. Москва, ул. Остоженка, 38); ведущий научный сотрудник Лаборатории мультиканальной коммуникации Института языкознания РАН (125009, Российская Федерация, г. Москва, Б. Кисловский переулок, 1); *сфера научных интересов*: когнитивная семантика, непрямое наименование в тексте, референциальная семантика, окулографические и корпусные методы анализа восприятия и порождения текста; *e-mail*: maria_kiose@mail.ru
ORCID: 0000-0001-7215-0604; Scopus Author ID: 56642747500; ResearcherID AAB-7989-2019; SPIN-код: 4419-0090; AuthorID: 341964.