MULTIMODAL CONSTRUCTION GRAMMAR OF TRANSITIVITY IN RUSSIAN EXPOSITORY DISCOURSE: THE CASE OF FIRST-PERSON PERSPECTIVE

Maria I. Kiose

Moscow State Linguistic University, Institute of Linguistics RAS
(Moscow, Russia)
maria_kiose@mail.ru

Abstract: The study addresses the cognitive nature of event construal in multimodal discourse. Adopting a multimodal construction grammar approach, it explores transitive and intransitive discourse events with a view to identifying the cognitive grounds of speech and co-speech gesture recurrent alignment patterns. We hypothesize that these multimodal constructions are mediated by the degree of transitivity, which is manifested in different conceptual structures. The key research question is whether functional gestures serve to demarcate the degree of transitivity (high and low transitive, intransitive, and copular utterances) in first-person perspective Russian expository discourse. The research data are 20 recorded expository monologues detailing personal experiences of creativity in professional communication, containing 725 cases of first-person perspective utterances and 1959 cases of co-speech gesture use. The results show that a) copular utterances are aligned with pragmatic discourse representational, emphatic, and evaluative gestures, b) intransitive clauses co-occur with deictic pointing gestures, c) low transitive clauses employ more contact-establishing gestures, d) high transitive clauses are significantly more frequently aligned with representational gestures, particularly enacting, holding, and molding. Overall, multimodal transitivity is expressed in the higher frequency of representational gestures, which supports prior results, but specifies them in viewing transitivity as a scaled phenomenon. Transitivity then acts as an event construal category in multimodal expository discourse, displaying actionability and objectivity in speech and in gesture; while intransitivity displays prevalence to foregrounding the role of the cognizing and acting agent. The results contribute to developing a multimodal construction grammar framework in exploring universal categories of transitivity, actionality, and perspective in discourse.

Keywords: multimodal construction grammar, transitivity, co-speech gesture, event construal, first-person perspective, expository discourse, professional communication

How to cite this article: Kiose, M.I. (2025). Multimodal Construction Grammar of Transitivity in Russian Expository Discourse: The Case of First-Person Perspective. *Professional Discourse & Communication, 7*(4), 38–55. https://doi.org/10.24833/2687-0126-2025-7-4-38-55

1. INTRODUCTION

In this paper, we explore the multimodal discourse construal of transitivity in speech and gesture. While substantial research on transitivity was carried out in typological studies and construction grammar (Hopper & Thompson, 1980; Testelec, 1998; de Swart, 2007; Haspelmath, 2015) and in cognitive grammar in attribution to language (Talmy, 2000; Goldberg, 2006; Langacker, 2008), its multimodal accounts are still scarce, and in attribution to the multimodal behavior within Russian language culture they are absent. Therefore, this study aims to reveal whether transitivity acts as a multimodal phenomenon in Russian language culture and to identify the co-speech gesture patterns, if there are any. To proceed, we adopt the multimodal construction grammar framework (Cienki, 2017; Ziem, 2017; Wu & Cienki, 2019; Hoffmann, 2021; Iriskhanova & Alieva, 2024), which views multimodal constructions in discourse as recurrent speech and co-speech gesture alignment patterns.

To examine the multimodal construal of transitivity, we address the cognitive linguistic accounts of force dynamics (Talmy, 1985; 2000) in identifying the cognitive grounds of transitive and intransitive events. Following the prior studies on multimodal construal of transitivity in English and German language culture, which specify the role of representational gestures in transitive events construal (Bressem & Müller, 2014; Wu, 2019; Wu & Cienki, 2019; Chan & Kelly, 2021), we hypothesize that i) high and low transitive, intransitive and copular utterances contribute to different multimodal constructions in Russian language culture; ii) the rate of representational gestures serves as a demarcation line between the multimodal constructions under consideration.

The paper is structured as follows. The Theoretical Framework section presents multimodal construction grammar and its cognitive postulates in attribution to exploring transitivity both in language and multimodally in speech and co-speech gesture. The Materials and Methods section introduces the compiled corpus and analysis procedure. The Study and Results section presents the data describing transitivity in the speech accounts and co-speech gestures, and finally outlines the multimodal constructions expressing transitivity with a first-person perspective. The Discussion section identifies the contribution of the study results as contrasted with prior studies in the area of multimodal construction grammar and the studies on transitivity. The Limitation section lists major constraints of the study. In the Conclusions section, we specify the study implementation for the field of multimodal, semiotic, and cognitive research.

2. THEORETICAL FRAMEWORK

In attribution to language, construction grammar is a "usage-based framework which represents grammatical knowledge in such a way that it can interface transparently with theories of processing, acquisition, and historical change" (Goldberg, 2006, pp. 215). Within this framework, constructions are viewed as displaying motivation for form-meaning correspondence, explaining why the construction makes sense. Meanwhile, the use of constructions is not restricted to language; at present, they are examined in multimodal systems, for instance, in language and prosody or language and gesture. In one of the earlier papers on multimodal construction grammar, Steen and Turner (2013) claim that "the construction of meaning across modalities is not mechanically additive; rather, meanings emerge as crossmodal blends that rapidly synthesize selected features of the information into new wholes" (Steen & Turner, 2013, p. 272). In his paper, Cienki (2017) resorts to multimodal constructions as lying at the heart of spoken language analysis while dis-

tinguishing between the surface construction of the utterance (both verbal and nonverbal) and its deep structure. The term 'crossmodal collostruction' was further proposed which manifests strong associations between semiotically different constructions (Uhrig, 2022) and which presupposes the existence of a nonverbal feature that may or may not have any independent meaning, but that "must be an integral part of the utterance construction even though it might not surface in every instance of it" (Lehmann, 2023, pp. 177). In examining multimodal constructions, the researchers adopt either of two procedural decisions (Bressem & Müller, 2014; Cienki, 2017; Zima & Bergs, 2017; Hoffmann, 2021; Lehmann, 2022). They either start with the linguistic phenomena and explore their co-occurrence with nonverbal units, or look into nonverbal units first and identify their co-occurrence with linguistic phenomena. Since in this paper we address transitivity, which is much explored in language studies, its language manifestations will serve as a starting point.

In language, transitivity is viewed in Hopper & Thompson (1980) as "a global property of an entire clause, such that an activity is 'carried-over' or 'transferred' from an agent to a patient" (Hopper & Thompson, 1980, pp. 251). In Haspelmath (2015), the notion of transitivity prominence was introduced, by which Haspelmath means "the extent to which the languages make use of transitive encoding" (Haspelmath, 2015, pp. 131). Following a number of studies where the same verbs (e.g., *help* and *follow*) are either encoded transitively (i.e., employ an argument in the Accusative) or not, Haspelmath observes that languages display "different proclivities for transitive encoding" (Ibid., pp. 136). Therefore, transitivity prominence is a feature attributed to a language that is identified via contrastive analysis of the transitivity patterns in languages. As opposed to this binary view, in Hopper & Thompson (1980), transitivity is viewed as a scalar property displaying a higher and a lower degree of intensity. Hopper & Thompson (1980) identify a list of parameters of transitivity, where each parameter variance contributes to achieving high and low transitivity. The list involves the following parameters (in brackets two poles of the scale "high – low transitivity" are presented): participants (2 or more participants vs. 1 participant), kinesis (action vs. non-action), aspect (telic vs. atelic), punctuality (punctual vs. non-punctual), volitionality (volitional vs. non-volitional), affirmation (affirmative vs. negative), mode (realis vs. irrealis), agency (high in potency vs. low in potency), affectedness of Object (totally affected vs. not affected), individuation of object (highly individuated vs. non-individuated). Therefore, transitivity is regarded as a continuum, and each language under analysis manifests a certain distribution of high and low transitivity features (Hopper & Thompson (1980) consider their relative values). This idea is partly concomitant with the one expressed in Testelec (1999), who specifies what might count as core transitives; these are the ones that require two semantic (thematic) roles, Agent and Patient. Revealing the degrees of transitivity/intransitivity in languages has become one of the key issues in linguistic typology stimulating the studies in valency and object marking (see the review in (de Swart, 2007)) in attribution to different phenomena connected with the idea of transitivity in languages, for instance the idea of split transitivity which describes two types of intransitive verbs functioning either as transitive or manifesting specific features depending on the presence of semantic roles and aspect (see the review in Arkadiev (2004)).

In their scaled view of transitivity, Hopper & Thompson (1980) hypothesize that transitivity can be manifested either morphosyntactically or semantically (Ibid., pp. 255). The second approach is adopted and further developed in cognitive grammars. For instance, Langacker (2008) distinguishes between high and low transitivity considering the role of arguments; he differentiates between Active Transitive (in *I opened the door*), Middle (in *The door opened easily*), Absolute Intransitive (in *The door opened*), and Passive (*The door was opened*) basing on the type of agent-theme interaction (Langacker, 2008, pp. 385). Langacker claims that the form of a clause is not

enough to make it transitive, since transitivity depends on how a situation or event is construed. For instance, in force dynamics construal (Talmy 2000), the intransitive verb *contain* can act as transitive in *The crowd was contained by the security guards;* similarly, *approach* is transitive when construed in relation to a social interaction in *She was approached by a strange*r (Langacker, 2008, pp. 388). Langacker then deduces that while agentive interaction is the prototype for transitive clauses, a thematic process is prototypical for intransitives (Ibid., pp. 396). Developing the force dynamics approach proposed by Talmy (1985), Pozdniakova (1997) specifies different roles of Agent within the construal patterns involving Figure, Ground, Move, and Path, determined by the use of first-person perspective. For instance, the author distinguishes between two types of subjects, Agent and Figure, where the first initiates the action while the second is construed as moving along the Ground. Therefore, the first-person perspective contributes to the construal of actionality, which might appear in the distribution of transitivity events in discourse.

In this study, to explore transitivity as a multimodal phenomenon in discourse, we adopt its cognitive view developed by Langacker (2008), and therefore identify the high, low transitivity, and intransitive constructions, considering the role of arguments within a discourse event. Following force dynamics framework (Talmy, 2000) we presume that high transitivity in discourse is observed when the Agent (Agonist) displays a tendency towards action (not towards rest); consequently, the object undergoes some change expressed for instance in the change of location, change of state, physical properties (shape, substance, color, etc.), motion type. These changes can be attributed to 13 palpability-related parameters specified by Talmy (Ibid.), which include palpability, clarity, intensity, ostension, objectivity, localizability, identifiability, content-structure, type-of-geometry, accessibility to consciousness, certainty, actionability, and stimulus dependence (Ibid., pp. 141-143), overall specifying different forms of ception (Ibid., pp. 159). When no such change occurs, a lesser degree of transitivity is identified. Intransitive construction is then found when an event does not have an object that is subjected to any force, while copular constructions are those that do not have an agent.

Meanwhile, we expect to reveal transitivity in co-speech gesture as part of multimodal constructions in the least studied expository discourse type, which explains or develops a topic and maintains a focus on the relations between various phenomena (Longacre, 1983; Berman, Nir-Sagiv, 2007). Therefore, following Ziem (2017) and Hoffmann (2021), we view multimodal constructions as conventional pairings of a complex form, comprising a verbal and a kinetic element and relating a specific meaning or function. While several constructions (e.g., bounded and unbounded, copular, transitive, and constructions of negation, comparison, and opposition) have already been explored in multimodal aspects (Müller, 2004; Bressem & Müller, 2014; Cienki & Iriskhanova, 2018; Wu, 2019; Iriskhanova & Alieva, 2024), the discourse event properties, for instance, types of agents, are also reported to affect the variance in the use of speech and gestures (Wu & Cienki, 2019; Iriskhanova et al., 2023). As is known, transitive, intransitive, and copular utterances (Langacker, 2008) mediate the use of representational gestures (in English-language culture) (Wu, 2019). Additionally, Wu & Cienki (2019) show that in the case of a human agent (as event property), the rate of representational gestures is 1.5 times higher than in the case of a natural force in both transitive (71%) and intransitive (76%) utterances. Chan & Kelly (2021) also report that the rate of gestures with iconic function (in Wu & Cienki (2019) these are representational gestures) increases in displaying an agentic relationship to actions and events. Meanwhile, Zlatev et al. (2020) claim that with a first-person perspective, the rate of representational gestures (pantomimes) is higher than with a third-person perspective. Therefore, presumably, in first-person perspective events, the observed differences between the high and low transitive, intransitive,

copular utterances will be high. This assumption explains why, in this paper, we addressed only the examples of the first-person perspective in compiling our corpus. Following these observations, we hypothesize that i) high and low transitive, intransitive, and copular utterances contribute to different multimodal constructions in Russian language culture; ii) this will be the rate of representational gestures that serves as a demarcation line between the multimodal constructions under consideration.

3. MATERIALS AND METHODS

The study employed the method of multimodal construction grammar in determining the recurrent patterns of utterance types in speech and co-speech functional gestures. To explore the distribution of gestures mediated by the utterance type: high transitive, low transitive, intransitive, and copular, we addressed the corpus of multimodal behavior of young people answering a pre-prepared set of questions related to their personal experience of creativity in professional communication and creative performance. The questions stimulated the speaker's recollection of his act of being creative, his manifestations of creative behavior, his forms of creativity in different spheres of life, the contribution of personality traits to developing creative skills – overall, these questions provoked a discourse of a truly personal type, which allowed to observe multiple cases of first-person perspective use. During the experiment, the participant's speech and gestures were videorecorded; the speech was further automatically transcribed using the Whisper Tool, and the gestures were manually annotated. Overall, the compiled multimodal corpus was 111 minutes long; 20 participants (average age – 20) were videorecorded, and their interview duration varied from 2:04 to 12:14 minutes. To annotate co-speech manual gestures, we adhered to the functional classification of gestures developed by Müller (2014) and specified for discourse analysis in Iriskhanova & Cienki (2018), Iriskhanova et al. (2023). Within this classification, four gesture groups are differentiated: deictic gestures (further classified as pointing, touching, directing), representational gestures (holding, molding, enacting, embodying, enlining), pragmatic gestures (discourse structuring, discourse representational, emphatic, contact-establishing, evaluative, negation, word search), adaptors (self-adaptors, object-adaptors).

The research procedure involved several steps.

At Step 1, the compiled corpus was annotated in the ELAN software[1] using the gesture classification presented above. Annotation was carried out by a group of 10 annotators, each working on 2 files; next, the annotations were checked within each pair of annotators and cross-checked by two more experienced coders systematically and entirely; total agreement was thus achieved. To give examples of gesture groups, consider their manifestations in Figure 1.

In Figure 1, all presented gestures are two-hand gestures. In Figure 1a, the speaker is using both hands to point at an imaginary object with his index fingers. In Figure 1b, the speaker is enacting with both hands her playing the piano while speaking about it. Figure 1c shows the speaker performing a pragmatic discourse representational gesture while presenting his idea as if it were part of the discourse. In Figure 1d, the speaker is using a self-adaptor in promoting her stream of thought. In several cases, gestures were identified as fulfilling several functions at a time. Two possible variants were observed when these gestures were produced by the same hands (Figure 2a) or different hands (Figure 2b).

---

[1]   ELAN (Version 6.9) [Computer software]. 2024. Nijmegen: Max Planck Institute for Psycholinguistics, The Language Archive: https://archive.mpi.nl/tla/elan

Fig. 1a. *Deictic gesture (pointing)*

Fig. 1b. *Representational gesture (enacting)*

Fig. 1c. *Pragmatic gesture (discourse representational)*

Fig. 1d. *Adaptor (self-adaptor)*

Fig. 2a. *Representational and pragmatic gestures produced by the same hands*

Fig. 2b. *Pragmatic gestures and adaptors produced by different hands*

In Figure 2a, the speaker uses a manual gesture that is both representational (holding) and pragmatic (discourse structuring). Both hands are employed simultaneously in performing these two gestures with the clause *потому что я не могу придумать что-то новое* (since I cannot think of something new)[2]. In Figure 2b, the speaker performs a pragmatic gesture (which is discourse representational and additionally contact-establishing) with his right hand; at the same time, we observe his use of an adaptor (self-adaptor) made with his left hand synchronized with the clause *у меня не получалось* (I did not succeed). Meanwhile, in most cases, different gesture types were produced by both hands simultaneously.

At Step 2 in the participants' speech, the examples with discourse markers of first-person perspective were selected and subjected to further multimodal analysis. As a unit for analysis, we appealed to the clauses containing the discourse markers of first-person perspective. We considered only the cases with explicit first-person perspective expressed in the discourse markers within the clauses, disregarding the false starts, hesitations, and interruptions. The cases included the clauses with i) first-person singular personal pronoun *I,* e.g., in *потому что в целом **я** сейчас больше такой аранжировщик композитор* (since I am now overall more of a conductor composer), ii) verbs in first-person singular without a first-person singular pronoun, e.g., in *в целом **хочу** как бы научиться ее исполнять хорошо* (overall [I] would like to learn to perform it well), iii) first-person singular possessive pronouns, e.g., in *в два часа ночью у **меня** пришла мелодия в голову* (at

---

<p>[2]   Here and below, the translation is provided by the author.</p>

two o'clock of the night a melody came into my head). We additionally included in the compiled subcorpus the cases of speech disfluencies containing first-person singular discourse markers, e.g., in *но я все равно приходилось доставать телефон* (but I still was necessary to use the phone). If the clause was incomplete but contained the proposition, we still included it into the subcorpus, e.g., in *поэтому это мне было более…* (therefore it was for me more of…) which ended with a prolonged pause.

At Step 3, the clauses were annotated to identify the discourse markers of transitivity and the type of utterance: high and low transitive, intransitive, and copular. The type of utterance was determined following the three-point scale suggested by Langacker (2008), i.e., intransitive, low and high transitivity, where high transitivity was identified in case of a transitive verb with the direct object present where the agent changed the state of the object, or as it is considered in force dynamics, forced it to move along the ground (whether imaginary or not). Low transitivity cases were observed with transitive verbs and direct objects where the agent did not change the state of the object, e.g. in *я выбрала танцы* (I chose dancing lessons), or where the verb is used in its indirect meaning, e.g., in *если касаемо моего направления* (if it concerns my area of interest), or when the object was not explicitly present in the clause but inferred from a prior clause like in *я учила танцы дома / я специально зеркалила себе видео с телека / и учила* (I learnt the dances at home / I purposefully mirrored the video from the TV screen / and learnt), where in the third clause we observe a case of low transitivity where direct object *танцы* (dances) is inferred being used in the first clause. Direct object could also be expressed via a deictic word, e.g., in *которую я написал* (which I wrote), where the direct object (*песня* (song)) of the verb *написал* (wrote) was found in the previous clause.

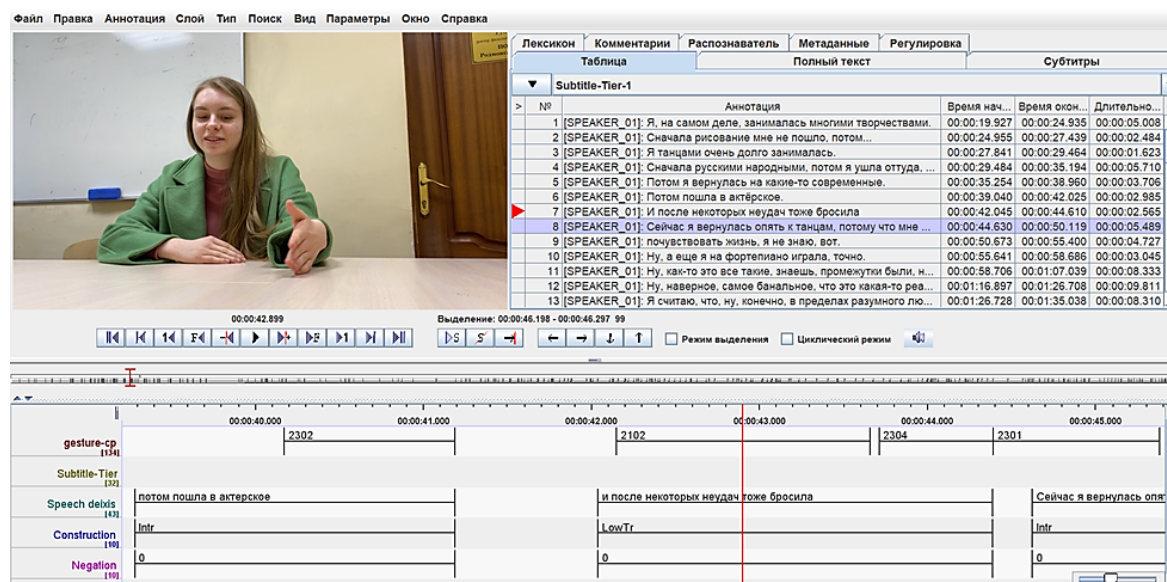In Figure 3, an example of completed annotation in the ELAN software is presented.

Fig. 3. *ELAN annotation of speech and gesture constructions of transitivity*

The tier "Speech deixis" introduces the clauses with a first-person perspective; here, three clauses of this type are presented. As noted above, as an example of a low transitive construction in speech, we considered the clauses where a transitive verb was not followed by a direct object, but the object was inferred from prior clauses, here that was in *потом пошла в актерское / и после некоторых неудач тоже бросила* (first I went to the actor's school / but following some failures I left). In the second clause, a transitive verb *бросила* (left) occurs, which is not followed

by a direct object; meanwhile, this object (*актерское* / the actor's school) is found in the previous clause; therefore, the second clause is an elliptical manifestation of a low transitive construction in speech. While pronouncing the first clause *потом пошла в актерское* (first I went to the actor's school), the speaker uses a pragmatic discourse representational gesture coded 2302 in the tier "gesture-cp". The second clause *и после некоторых неудач тоже бросила* (but following some failures I left) is accompanied by two consecutive gestures, deictic (touching) coded 2102 and next pragmatic (contact-establishing) coded 2304. At the moment presented in Figure 3, the participant is using a deictic gesture.

At Step 4, to explore the co-speech gesture synchronization, we adhered to the following regulations. First, in the study, we submitted to analysis all the gestures appearing as synchronized with the clause, irrespective of their co-use with any of the clause components. This decision was prompted by the fact that the predicative meaning of transitivity, which is under study, is formed by both verbal and nominal components of the proposition. Second, we did not count the number of gestures commonly equal to the number of gesture strokes, but considered the number of instances of functional gesture co-speech use with the clauses. This decision is explained by the functionality of the gesture types employed for the analysis, since multiple gestures do not manifest a clear stroke but are performed with a certain hand tension, especially representational gestures and pragmatic discourse representational gestures. Additionally, gesture strokes commonly occur with an accentuated word in the clause, which is frequently an intransitive verb, rarely a transitive verb in case it is followed by an object, and almost never a copular verb. For this reason, the gesture stroke criterion is not considered in the study.

At Step 5, the obtained quantitative data displaying speech and co-speech gesture distribution were presented in absolute values and in ratio and were subjected to further statistical processing to identify whether there exist recurrent patterns of co-speech gesture use with high transitive, low-transitive, and intransitive utterances, which might serve as multimodal constructions.

4. STUDY AND RESULTS

In 4.1, we present the results of the distribution of utterance types in the participants' speech in the clauses with a first-person perspective. In 4.2, we proceed to the distribution of gesture types used with the same clauses. Finally, in 4.3, we address the co-speech gesture alignment in attribution to three types of utterances to explore the possible multimodal constructions.
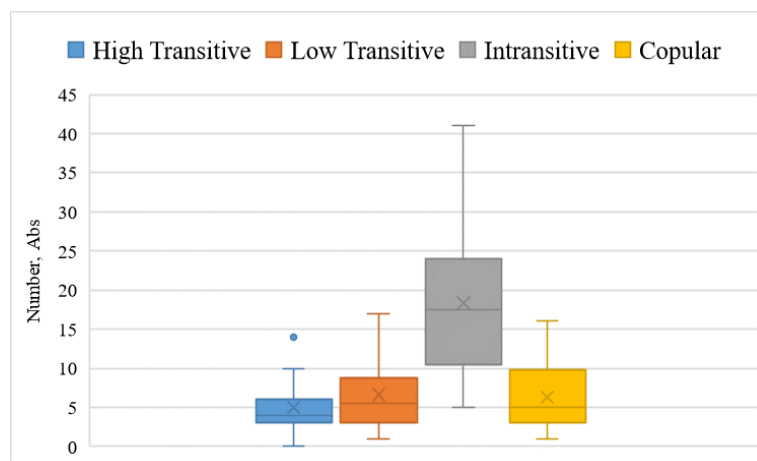

Fig. 4. *The distribution of utterances in participants' speech*

4.1. Transitivity in the speech accounts

Overall, 725 cases of first-person perspective utterances were subjected to analysis. In Figure 4, we present the distribution of utterance types in the participants' first-person perspective clauses, considering the speech of each participant.

Figure 4 shows that the prevalent utterance type is intransitive (Range=36, SD=10.0, Mean=18.35, Median=17.5) which appears in multiple cases manifesting different forms of ception (Talmy 2000), including actionability in *я пошла на рисование* (I went to drawing classes), *я играю в группе* (I play in a group), accessibility to consciousness in *как я считаю* (as I think) and *что-то я соскучился по нему* (I am somehow missing him), *я сейчас так не вспомню* (I cannot recollect it), identifiability in *ну у меня такой появился всплеск* (I had a splash), ostension in *то есть если я вижу* (that is if I see it), etc. Other utterance types are less frequent. Low transitive utterances (Range=16, SD=4.42, Mean=6.6, Median=5.5) appear in the cases when the changes in several types of ception occur, for instance, in identifiability in *потому что как я это объясню* (because how can I explain it) and *я всегда привожу этот пример с этой песней* (I always give this example with this song), accessibility to consciousness in *потому что в то время я переживал очень сильный период в своей жизни* (because at that time I was passing through a very harsh period in my life), *я уже не вспомню текст* (I won't recollect this text), objectivity in *я брошу все свои дела* (I will give up all my work), content-structure in *пока не достиг этого* (up to now I haven't reached it), etc. Copular structures display similar frequency (Range=15, SD=4.35, Mean=6.3, Median=5); these utterances reveal ostension in *и вот для меня креативность это смотреть на мир не напрямую* (and well creativity for me is looking at the world indirectly), actionability and intensity in *у меня был творческий порыв* (I had a flash of creativity), content-structure and intensity in *у меня было долгое время перегорание* (I had a burnout), identifiability in *На самом деле у меня довольно много хобби* (As things stand I have several hobbies), etc., overall, a variety of ception types. Least frequent are high transitive utterances (Range=14, SD=3.39, Mean=5, Median=4) which display the following changes in ception: actionability, objectivity in all cases, and type-of-geometry in *и начал рисовать там структуру человека* (and started to create a contour of a man), and content-structure in *и я захотела писать книгу* (and I wished to write a book) where high transitive utterance also displays intransitivity due to the presence of *захотела* (wished), and localizability in *и когда я ставила какие-то постановки* (and when I staged some plays), and identifiability in *которая во многом наверно сформировала мой музыкальный вкус* (which mostly contributed to shaping my sense of music). Meanwhile, in many cases, other ception types can be identified in the examples presented above due to the fact that different types of ception commonly co-occur. In rare cases, a veritable type of ception can be identified in the participants' speech since the speech frequently discloses the type of object under consideration but seldom reveals the details on how the participants conceived this object.

To explore individual variance in the use of utterance types, we subjected the data to non-parametric repeated measures ANOVA (The Friedman's test) since the data in low transitive and copular utterance types were not normally distributed (Shapiro-Wilk's p=0.031 and 0.049, respectively). ANOVA test revealed high variance in the participants' use of utterance types, with $\chi2$ (3) =35.4 at p<0.001, which means that the distribution of utterance types is highly dependent on individuated narrative structure, although the questions answered were the same for all participants.

The distribution of different types of ception exposes several regulations which help differentiate the observed force dynamics conceptual structures underlying these surface language structures. Since high transitive utterances in all cases construe the event, focusing its actionability and

objectivity, we presume that adopting the framework of force dynamics, these exemplify the cases where the Agent initiates / forces the Figure to move against the Ground. Low transitive utterances do not manifest the action of the Agent producing this move; therefore, the Agent acts as the conceptualizer of Figure, making or not making its move against the Ground. Intransitive utterances construe information as focusing actionability, accessibility to consciousness, ostension, etc.; however, Agent is construed as Figure, which is making or not making a move against the Ground. In copular utterances, the Figure is immobile, while the Agent is also present (in our case, with first-person perspective).

4.2. Transitivity expressed in gestures

Next, we present the distribution of gestures irrespective of the utterance type, to give a view of general regulations attributed to the first-person perspective appearing in all utterances. Since we observed multiple cases of gesture use within the same clause where one gesture type could be used repeatedly, being co-occurrent with other gesture types (see examples above in Fig. 2), the decision was adopted to count not the overall number of gestures but the number of their co-occurrences with utterances, here the utterances with first-person deixis. Overall, 1959 cases of co-speech gesture use were subjected to analysis against 5603 co-speech gestures found in 20 recorded expository monologues. In Figure 5, we show the total distribution of all gesture types.
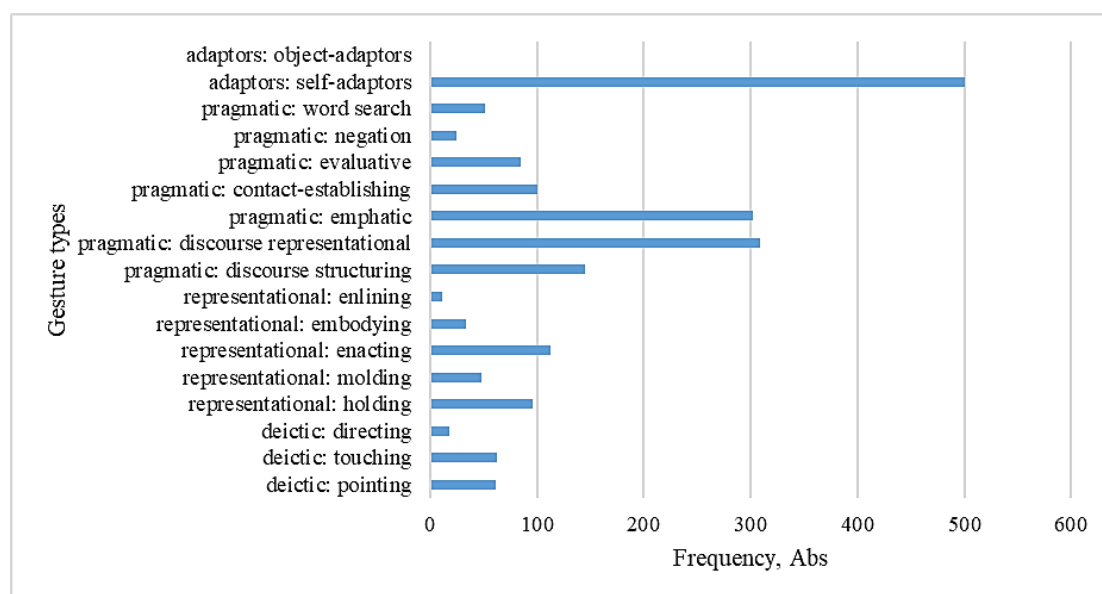

Fig. 5. *Overall distribution of co-occurrent gestures in the clauses*

The results show that while several functional gesture types are infrequent or absent, others predominate. The most frequent are self-adaptors (500, 39% of all self-adaptors (1280) in the corpus of monologues), which is highly expected due to the fact that we observed the cases of spontaneous speech related to the matters of the participants' personalities. Next frequent are pragmatic gestures expressing discourse representational potential (309, 33.3% of all discourse representational gestures (928) in the corpus) and emphaticity (302, 40.1% of all emphatic gestures (753) in the corpus), which is not unexpected either since in answering the interviewer's questions the speakers had to construe the object of reference in presenting their point and commonly expressed their opinion. While representational and deictic gestures may seem less common in contrast to pragmatic gestures and adaptors, their ratio is still significant; this assertion mostly refers to rep-

resentational enacting (112, 45.2% of all enacting gestures (248) in the corpus) and holding (96, 28.7% of all holding gestures (334) in the corpus) gestures and deictic touching (62, 36.5% of all deictic touching gestures (170) in the corpus) and pointing (61, 35.3% of all deictic pointing gestures (173) in the corpus) gestures. Consequently, the object of reference "appears" both in representational gestures and in highly frequent pragmatic discourse representational gestures. Interestingly, we observe an overall increase in the relative frequency of representational enacting gestures, which can be attributed to the construal of objects of reference in the first-person perspective (see Figure 6).
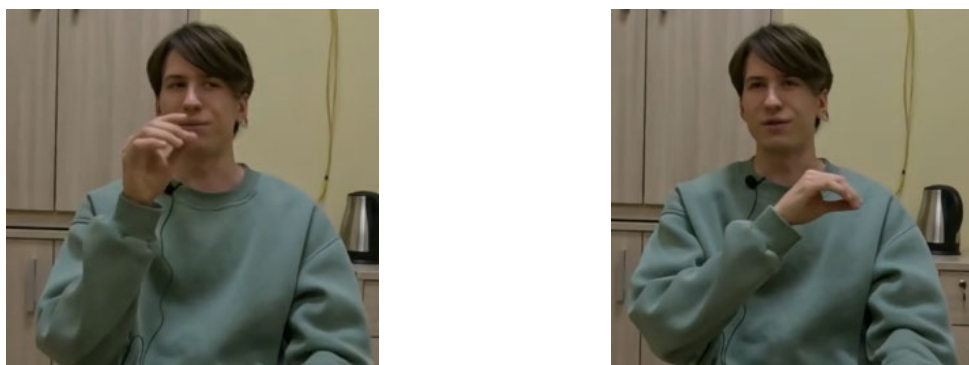

Fig. 6. *Representational enacting gesture*

Figure 6 shows the case of a manual gesture performed as if the participant were holding a paint brush and were drawing something with it in the air; this gesture co-occurs with the active transitive verb копировал (was copying) in the clause То есть я как бы копировал стиль (that is I was somehow copying the style). Presumably, the first-person perspective stimulates higher enactment; however, this result can also stem from the redistribution of actional verbs in the overall corpus, which we did not consider in this study.

4.3. Multimodal constructions expressing transitivity with first-person perspective

Finally, we address the distribution of gesture types with utterance types in the compiled corpus with first-person perspective clauses. To proceed, we identify the absolute frequency of gesture type occurrences; next, to contrast the distribution of gestures with different utterance types manifesting a different frequency, we adhere to their relative percentage values calculated as a ratio from the total number of each co-utterance type instances. In Table 1, we show the relative percentage distribution of gestures with four types of utterances, copular, intransitive, low transitive, and high transitive, manifesting a scaled transitivity potential. Additionally, we consider the accumulated relative percentage value of gestures used with both high and low transitivity cases to contrast the results with copular and intransitive cases.

Table 1. *Gesture types co-occurrences with utterance types*

	Copular	Intransitive	Low transitive	High transitive	Transitive
deictic: pointing	2.357	**4.070**	2.667	2.674	2.669
deictic: touching	2.357	2.762	3.333	4.278	**3.696**
deictic: directing	1.010	0.727	1.000	1.070	1.027
representational: holding	4.377	3.343	5.000	**8.021**	**6.160**
representational: molding	2.020	1.453	2.000	**5.348**	**3.285**
representational: enacting	3.704	3.052	5.667	**12.299**	**8.214**

representational: embodying	2.020	2.180	1.000	1.604	1.232
representational: enlining	0.337	0.872	0.333	0.535	0.411
pragmatic: discourse structuring	**5.387**	8.576	6.667	8.021	7.187
pragmatic: discourse representational	**17.508**	16.424	15.667	13.369	14.784
pragmatic: emphatic	**17.508**	15.988	17.000	**10.160**	14.374
pragmatic: contact-establishing	4.377	5.378	**7.000**	**2.139**	5.133
pragmatic: evaluative	**5.724**	4.070	3.667	4.813	4.107
pragmatic: negation	1.684	1.163	1.667	0.535	1.232
pragmatic: word search	2.020	2.180	3.333	2.674	3.080
adaptors: self-adaptors	27.609	27.616	24.000	**22.460**	**23.409**
adaptors: object-adaptors	0.000	0.145	0.000	0.000	0.000

Table 1 shows (the percentage values displaying major differences are given in bold) that copular utterances less frequently co-occur with pragmatic discourse structuring gestures, while the clauses with both stative and active verbs require the use of gesture in structuring the discourse flow. Meanwhile, these clauses are aligned with pragmatic discourse representational, emphatic, and evaluative gestures. Intransitive clauses, as opposed to all other clause types, co-occur with deictic pointing gestures; presumably, the latter are attributed to spatial localization expressed more frequently in intransitive clauses. Low transitive clauses employ more contact-establishing gestures, which may be due to the fact that they display different types of changes in ception without actionalizing this change; it is possible that information transfer in this case substitutes for this actionalizing, which, as a result, produces contact-establishing gestures. This assumption is strengthened by the fact that these are contact-establishing gestures, which are the least frequent with high transitive utterances. Additionally, emphatic gestures appear rarely with this utterance type. High transitive clauses are significantly more frequently aligned with representational gestures, particularly enacting, holding, and molding, which suffices to presume that in this case, this is the action itself and the object undergoing the change, which become focalized and not the attitude towards this object or the speaker's position towards the interviewer. Overall, transitive utterances display alignment with representational gestures (namely, enacting, holding, and molding) and are less frequently found with self-adaptors than copular and intransitive clauses.
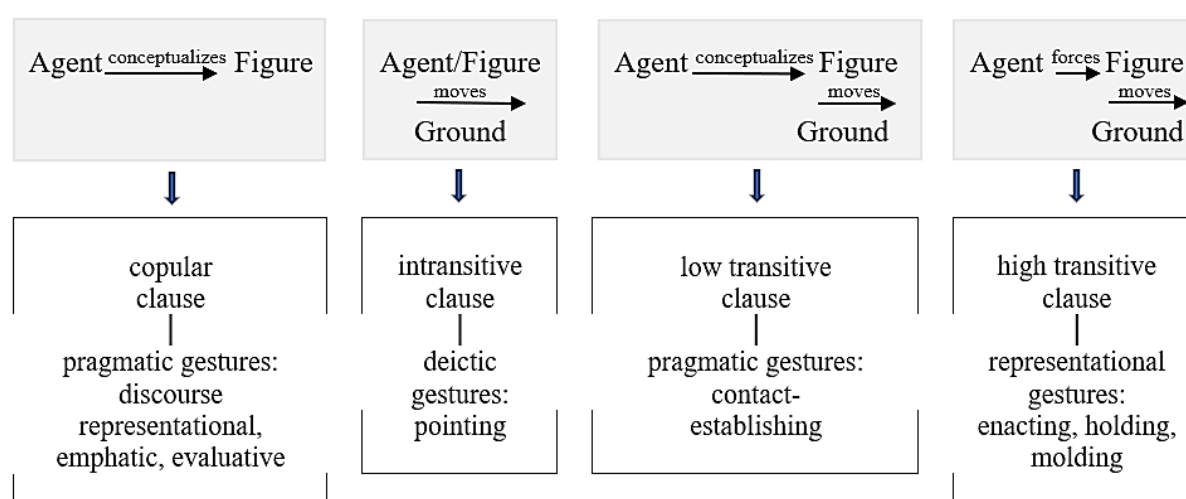

Fig. 7. *Multimodal constructions of scaled transitivity in first-person perspective expository discourse*

Nonparametric repeated measures ANOVA (The Friedman's test) following the normality tests of co-speech gesture distribution with four utterance types (Shapiro-Wilk's p<0.001, 0.011, <0.001, and <0.001, accordingly) revealed significant differences in gesture distribution with utterances types, with $\chi2$ (3) =32.5 at p<0.001, which proves that the utterances display specificity in gesture co-occurrences. Overall, the study reveals several multimodal constructions in speech and gesture mediated by transitivity in the utterances with a first-person perspective in expository discourse (Fig. 7).

Figure 7 shows the conceptual structures that underlie the multimodal constructions manifested in speech and gesture. As described above, four conceptual structures display differences in the roles of Agent (either a conceptualizer or the one who initiates the action), and Figure (either moving or stative), as well as in their interaction (either two concepts or one). These structures serve to construe events multimodally via four multimodal constructions. Figure 8 presents several specific alignment patterns of speech and gesture co-occurrences that can be related to multimodal constructions.

Fig. 8a. *Representational molding gesture used with a high transitive utterance*

Fig. 8b. *Deictic pointing gesture used with an intransitive utterance*

In Figure 8a, the speaker reports her action (here придумала is used in the meaning 'devised') in *даже придумала обложку к книге* (I even devised a book cover), which produced a result, a new object, which is therefore construed both in speech and gesture. The conceptual structure underlying it is the one where Agent (the Speaker) forces the Figure (a book cover) to move against the Ground (the book). In speech, we observe the use of a high transitive clause, while in gestural behavior, a representational molding gesture showing the appearing cover on the book is used. In Figure 8b, the speaker gives an account of a non-resultant action which focalizes its actor, the speaker herself, in *вот я например еще фотографирую* (As for me, I also take photos), which brings out a self-pointing gesture. The conceptual structure underlying it is the one where Agent (the Speaker) is simultaneously the moving Figure (the Ground is not explicitly present in the clause). In speech, this structure appears in the intransitive clause, while in gestural behavior, the speaker focalizes her agentive role with a deictic pointing gesture.

The revealed constructions display a considerable degree of flexibility, which means that the aligned gesture types are also used with other utterance types, and vice versa, that the considered utterance types are used with other gesture types as well. This observation reveals the specific character of multimodal constructions, which results from the discourse multifunctionality of gestures; the construction stability can consequently be mediated by discourse type, the type of action, the object of reference, the type of perspective, and even individual preferences. Meanwhile, the revealed regularities evidence in favor of existing constructions which can shed light onto how speech and gesture attribute to the construal of transitivity as a multimodal discourse phenomenon.

5. DISCUSSION

The obtained results specify the multimodal nature of transitivity within Russian language culture as well as contribute to the overall development of the multimodal construction grammar framework as a component of multimodal discourse studies.

First, the results evidence that viewing transitivity as a scalar property following Hopper & Thompson (1980) helps identify the degrees of its manifestation, which appear in different co-occurrent gestures in expository discourse. The study shows that since gestures differentiate between the utterances with different transitivity types (here, low and high transitivity after Langacker (2008)) as opposed to intransitive and copular utterances, these patterns are attributed to different ways of multimodal construal. Therefore, apart from the typological research of transitivity in language (Testelec, 1999; Arkadiev, 2004; Haspelmath, 2015), its multimodal research in speech and gesture can allow us to identify the distribution of valency and object marking (de Swart, 2007) expressed multimodally to form the idea of transitivity in thought, rather than in language only. What is more, the study proves that the cognitive idea of transitivity proposed by Langacker (2008), who claims that transitivity depends on how a situation is construed, is highly relevant for examining this phenomenon as multimodal. Developing his idea of agentive interaction being the prototype for transitive clauses, and a thematic process being prototypical for intransitives, we found that these two utterance types manifest significantly different multimodal construal patterns. In the first case, transitivity in language is coupled with act and object representation in gesture, which overall reveals the idea of agentive interaction. In the second case, intransitivity in language is coupled with deictic pointing at the actor himself, revealing the idea of a thematic process accomplished by the speaking actor in this study. Presumably, in this case, the Agent who is the speaker himself acts as the Figure or the locus of action, making self-reference via a deictic gesture, which seems a highly economical and logical multimodal strategy of discourse construal. Additionally, the framework of force dynamics and the system of palpability-related parameters specified by Talmy (2000) allow us to distinguish between the conceptual structures underlying these multimodal constructions. While intransitive utterances in speech display actionability, accessibility to consciousness, identifiability, ostension, etc., the prevalence of deictic gestures aligned with these utterances emphasizes the role of the cognizing and acting agent. Since transitive utterances in speech all display actionability and objectivity, these very features are emphasized in the use of representational gestures. Therefore, deictic gestures in intransitive multimodal constructions function complementarily with speech, whereas representational gestures in transitive multimodal constructions act to enhance speech. The study also revealed the regulations in the distribution of contact-establishing gestures, which prevailed with low-transitive utterances, presumably due to multimodally enacting the information transfer to the interlocutor as a proxy for the absent transfer of force to a patient in the absence of a direct, affected object[3]. Consequently, our first hypothesis, claiming that high and low transitive, intransitive, and copular utterances contribute to different multimodal constructions in Russian language culture was majorly verified.

Second, the study helped identify the multimodal constructions that underlie each of the utterance types in the Russian language spontaneous expository discourse of first-person perspective. Importantly, the obtained results are in line with the results reported of transitivity in its multi-

---

[3]   The author is grateful to the anonymous reviewer who contributed to the discussion of the ontological nature of multimodal constructions.

modal construal in other cultures. As noted above, transitive and intransitive utterances mediate the use of representational gestures in English-language culture (Wu, 2019); similar evidence was revealed for Russian language culture in this study. These were representational gestures of three types, enacting, holding, and molding, which served as a borderline of transitive and intransitive construal, being more frequent with transitive patterns. Consequently, our second hypothesis stating that this will be the rate of representational gestures that serves as a demarcation line between the multimodal constructions under consideration is also verified. Meanwhile, we additionally observed the redistribution of deictic gestures with intransitive clauses and pragmatic gestures with low transitive clauses, which significantly specifies the results. Presumably, the high rate of representational gestures is also attributed to the first-person perspective of the clauses in expository discourse. Therefore, the results support the idea expressed in Wu & Cienki (2019) and Chan & Kelly (2021), who show that in the case of a human agent, the rate of representational gestures is higher.

Overall, the results support the multimodal construction framework (Cienki, 2017; Ziem, 2017; Hoffmann, 2021; Lehmann, 2023; Uhrig, 2022), which allows for transplanting the cognitive constructionist view of language (Goldberg, 2006; Langacker, 2008) onto a multimodal discourse ground. The study showed that the adopted approach, which takes linguistic phenomena as a starting point for revealing the speech and gesture alignment, is relevant and efficient in disclosing these constructions in discourse. Meanwhile, a different approach, namely taking a specific gesture type and exploring its alignment with utterance types, will probably be less productive for the objectives of this study, since it will help scale the constructions as attributed only to one gesture type. In the case of transitivity, where several types of gestures contributed to differentiating the utterances, appealing to language in the first place produced a more accomplished picture. Hopefully, further research in the area of multimodal construction grammar will contribute to multimodal typology of languages as well as to multimodal typology of discourse, at least in attribution to such essential categories as transitivity, actionality, and perspective.

6. LIMITATIONS

The primary limitation of this study is the small and age-specific sample size, which may have affected the generalizability of the findings determining the distribution of multimodal constructions of transitivity. Future research should include a larger and more diverse sample to validate the findings. Meanwhile, performing on a homogeneous sample helped attain a more focused result. The second limitation is the obligatory necessity to consider all the co-speech functional gestures synchronized with the clause components as contributing to multimodal transitivity. While gesture strokes are commonly synchronized with accentuated words, the functional gestures shape the propositional meaning of transitivity, not restricted to a verb, which is a key research interest of this paper. Meanwhile, including all co-speech gestures means that other speech patterns may have stimulated them, which is a clear shortcoming. This limitation can be balanced in future studies by exploring the multimodal patterns containing only accentuated utterance verbs.

7. CONCLUSIONS

The findings of this study offer a multimodal construction grammar framework for dealing with the problem of transitive and intransitive event construal in discourse. This framework provides valuable insights into the nature of multimodal cognition and communication. The results

show that the multimodal constructions manifested in speech and gestures are mediated by different conceptual structures that describe the patterns of force-dynamic event construal. Therefore, appealing to event construal patterns allows us to reveal the allowances and constraints of human cognition and communication via multimodal behavior.

The study has several major contributions.

First, it revealed that transitivity in discourse is largely multimodal, expressed not only via speech but also via gesture. The results obtained for Russian language culture add to the data earlier found in other cultures, mainly English and German-speaking. This is evidence of the importance of further incorporating a gestural component into transitivity studies carried out in application to language, since the multimodal account of transitivity can contribute to specifying the roles of agent and object in discourse construal. This could help promote a typological multimodal research framework to identify the discourse patterns of multimodal construal of universal categories of time, space, agency, and reification, among them.

Second, the observed regularities specify the input of both communicative modalities, speech and gesture, in the multimodal construal of transitivity. The results show that while in intransitive utterances the deictic gestures which shape this construction specificity act to complement speech, in transitive utterances the representational gestures serve to enhance it. This assumption, although needing further verification, allows us to consider multimodal constructions within a semiotic framework, assigning different functions to different modalities. The findings are therefore important for multimodal discourse studies since they contribute to shaping the view of multimodal expository discourse, which is the least explored discourse type in terms of its multimodal status.

Third, the results help develop the frameworks of multimodal cognition research with its applications in neurocognitive, educational, cultural, and ontological studies. Multimodal construction grammar, with its revealed regularities, calculable distributions, and solid linguistic ground, serves as a reliable instrument for exploring how we construe discourse information multimodally.

Conflict of interest
The author declares that there is no conflict of interest.

REFERENCES

1. Arkadiev, P.M. (2004). *Rasshcheplennaya neperekhodnost': problemy teorii i tipologii [Split intransitivity. Semantics and typology]*. Report in Institute for Information Transmission Problems (Kharkevich Institute) RAS. 17.09.2004. https://inslav.ru/images/stories/people/arkadiev/Arkadiev_unaccus_spb-handout2007.pdf (in Russian).

2. Berman, R.A., & Nir-Sagiv, B. (2007). Comparing narrative and expository text construction across adolescence: A developmental paradox. *Discourse Processes, 43*(2), 79-120. https://doi.org/10.1207/s15326950dp4302_1

3. Bressem, J., & Müller, C. (2014). The "Negative-Assessment-Construction" – A multimodal pattern based on a recurrent gesture? *Linguistics Vanguard, 3*(1), 20160053. https://doi.org/10.1515/lingvan-2016-0053

4. Chan, D.M., & Kelly, S. (2021). Construing events first-hand: Gesture viewpoints interact with speech to shape the attribution and memory of agency. *Memory and Cognition, 49,* 884-894. https://doi.org/10.3758/s13421-020-01135-0

5. Cienki, A. (2017). Utterance Construction Grammar (UCxG) and the variable multimodality of constructions. *Linguistics Vanguard, 3*(s1), 20160048. https://doi.org/10.1515/lingvan-2016-0048

6. Cienki, A. & Iriskhanova, O. (Eds.). (2018). *Aspectuality across languages: Event construal in speech and gesture.* Amsterdam: John Benjamins.

7. de Swart, P. (2007). *Cross-linguistic Variation in Object Marking* [Ph. D. thesis, Radboud University Nijmegen]. The Netherlands: LOT. https://www.lotpublications.nl/Documents/168_fulltext.pdf

8. Goldberg, A.E. (2006). *Constructions at work. The nature of generalization in language.* Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199268511.001.0001

9. Haspelmath, M. (2015). Transitivity prominence. In A. Malchukov, B. Comrie (Eds.), *Introducing the Framework, and Case Studies from Africa and Eurasia* (pp. 131-148). Berlin, Boston: De Gruyter Mouton. https://doi.org/10.1515/9783110338812-008

10. Hoffmann, Th. (2021). Multimodal Construction Grammar: From multimodal constructs to multimodal constructions. In: Xu Wen & R. John (Eds.), *The Routledge Handbook of Cognitive Linguistics* (pp. 78–92). New York: Routledge.

11. Hopper, P.J., & Thompson, S.A. (1980). Transitivity in grammar and discourse. *Language, 56*(2), 251-299.

12. Iriskhanova, O.K., Kiose, M.I., Leonteva, A.V., & Agafonova, O.V. (2023). Multimodal languaging: Reification profiles in language and gesture. *Linguistic Frontiers, 6*(2), 78-91. https://doi.org/10.2478/lf-2023-0014

13. Iriskhanova, O.K., & Alieva, M.R. (2024). Polimodal'naya grammatika konstruktsii: protivopostavlenie i zhesty v ustnom eksplanatornom diskurse [Multimodal construction grammar: Opposition and gestures in spoken explanatory discourse]. *Cognitive Studies of Language, 5*(61), 95-105 (in Russian).

14. Langacker, R.W. (2008). *Cognitive Grammar: A basic introduction.* Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195331967.001.0001

15. Lehmann, C. (2023). As if that wasn't enough: English as if clauses as multimodal utterance constructions. *English Language and Linguistics, 27*(1), 175-202. https://doi.org/10.1017/S1360674322000326

16. Longacre, R.E. (1983). *The grammar of discourse.* New York: Plenum Press.

17. Müller, C. (2004). The Palm-Up-Open-Hand: A case of a gesture family? In C. Müller & R. Posner (Eds.), *The semantics and pragmatics of everyday gesture: The Berlin conference* (pp. 233–256). Berlin: Weidler Verlag.

18. Pozdniakova, E.M. (1997). Yazykovye sredstva realizatsii semanticheskoi voli agensa v pozitsii sub'ekta [Language means of expressing the semantic role of agent in subject position]. *Vestnik of Tambov University. The Humanities, 1*(5), 98-102 (in Russian).

19. Steen, G., & Turner, M. (2013). Multimodal Construction Grammar. In: M. Borkent, B. Dancygier, and J. Hinnell (Eds.), *Language and the Creative Mind* (pp. 255-274). Stanford, CA: CSLI Publications. https://doi.org/10.2139/ssrn.2168035

20. Talmy, L. (1985). Lexicalization patterns. Semantic structure in lexical forms. In T. Shopen (Ed.), *Language Typology and Syntactic Description, vol. 3: Grammatical categories and the lexicon* (pp. 57-149). Cambridge: Cambridge University Press.

21. Talmy, L. (2000). *Towards a cognitive semantics. Vol 1. Concept structuring systems.* Cambridge, Massachusetts: The MIT Press. https://doi.org/10.7551/mitpress/6847.001.0001

22. Testelec, Ya. (1998). On two parameters of transitivity. In: L. Kulikov, H. Vater (Eds.), *Typology of Verbal Categories. Papers presented to Vladimir Nedjalkov on the occasion of his 70th birthday* (pp. 29-46). Tübingen: Niemeyer.

23. Uhrig, P. (2022). Hand gestures with verbs of throwing: Collostructions, style and metaphor. *Yearbook of the German Cognitive Linguistics Association, 10*(1), 99-120.

24. Wu, S. (2019). Multimodality of transitive, intransitive and copular constructions in spoken language. *Linguistics and Literature Studies, 7*(6), 241-249. https://doi.org/10.13189/lls.2019.070601

25. Wu, S., & Cienki, A. (2019). Transitivity, events, and gesture: The Case of the causative-inchoative alternation. *Open Linguistics, 5,* 311-331. https://doi.org/10.1515/opli-2019-0017

26. Ziem, A. (2017). Do we really need a multimodal construction grammar? *Linguistics Vanguard, 3*(s1), 20160095. https://doi.org/10.1515/lingvan-2016-0095

27. Zima, E., & Bergs, A. (2017). Multimodality and construction grammar. *Linguistics Vanguard, 3*(s1), 20161006. https://doi.org/10.1515/lingvan-2016-1006

28. Zlatev, J., Żywiczyński, P., & Wacewicz, S. (2020). Pantomime as the original human-specific communicative system. *Journal of Language Evolution, 5*(2), 156-174. https://doi.org/10.1093/jole/lzaa006

About the author:

Maria I. Kiose is Doctor Habil. of Philology, Associate Professor, and Chief Researcher at the Centre for Socio-Cognitive Studies (Moscow State Linguistic University, Moscow, Russia) and at the Multi-Channel Communication Laboratory (Institute of Linguistics, Russian Academy of Sciences). Her research interests include cognitive semantics and multimodal communication. ORCID: 0000-0001-7215-0604.