PENSOFT.

La Mo | **Languages and Modalities**

Aesthetic multimodality of speech and gesture: Towards its functional framework

Maria Kiose[1,2], Anna Leonteva[1,2], Olga Agafonova[1]

1   *Moscow State Linguistic University, Moscow, Russia*
2   *Institute of Linguistics, Russian Academy of Sciences, Moscow, Russia*

Corresponding author: Anna Leonteva (lentevanja27@gmail.com)

Abstract

The study develops a functional multimodal approach to speech and gesture behavior to explore aestheticism in more and less staged discourse of cinema and interview. We hypothesize that cinema and interview employ the same communicative functions; however, these functions constitute different frameworks which contribute to the higher aesthetic potential of cinema. This approach allows to study the aesthetic via communicative functions frameworks in multimodal discourse.

To establish the function frameworks in cinema and interview, we apply a contrastive functional analysis of speech and gesture in the highly ranked actors' argumentative and descriptive monologues. With the help of variance and regression analysis, we explore the distribution of pragmatic and discourse-structuring functions (with sub-functions) in speech as contingent on pragmatic, deictic, representational and adaptive functions of gestures. The study confirms that cinematic discourse exploits fewer deictic, representational and adaptive gesture functions, whereas pragmatic gesture functions (especially emphatic ones) appear more frequently and are contingent on several pragmatic and discourse-structuring functions of argumentative and descriptive speech. Interview function frameworks display lower predictability, which shows higher spontaneity of gestures; however, there are specific gestures typical of interview (self-adaptors) which may serve as indicators of pragmatic functions of argumentation. The study also manifests individual variations within the function frameworks. Overall, cinema and interview display variance in replication and regularity of speech and gesture functions, which presumably helps create higher and lower aesthetic effects.

Key Words

1. Introduction

Until recently, aesthetic semiotics has been the only influential approach exploring aestheticism in multimodal discourse, for instance, in cinema. Aesthetic semiotic studies initiated by G. Shpet and R. Jakobson (Shpet, 1922 (2007), Jakobson, 1960) single out the poetic (or aesthetic according to Shpet) function of semiotic unities and develop the notion of their structure in terms of its form and content. As Jakobson puts it, "poetic function is not the sole function of [verbal] art, but only its dominant, determining function" (Jakobson, 1960: 356), it foregrounds

the inherent features of poetic discourse, totally re-evaluating it. The major principle of poetic discourse is "the projective principle of equivalence from the axis of selection onto the axis of combination" (Ibid: 358), which presupposes that these are the combinations of semiotic constituents (signs with their codes according to Jakobson), and not the constituents themselves that stimulate the poetic potential of discourse, however being so poetic function depends on other communicative functions (referential, emotive, phatic, metalingual (metalinguistic) and conative) within the functions hierarchy. Jakobson's "projective principle" and contrastive analysis as a semiotic

instrument for exploring aesthetic function of artwork was later much debated in semiotic studies. For instance, R. Goodrich in his highly cited paper (1997) claims that both selection and combination are presumed to be characteristics of a semiotic system (language in the works of Jakobson and Goodrich) as a whole, besides these are not solely the relations of similarity and contiguity which according to Jakobson guide selection and combination, but other relations produced by "deductive reasoning" (Goodrich, 1997: 63) especially in causal connections. Goodrich even states that "the purported connection between the concepts of selection and similarity is hardly a necessary or universal one" (Ibid), which undermines the functions hierarchy. Nevertheless, he admits that poetic discourse does display the "clusters of concomitant features" functioning as its criteria with poetry in general being one of the "open-ended concepts" (Ibid 65).

Multimodal approach to aesthetic experience can offer its own solution to the problem of aesthetic function in discourse. Following Goodrich, we assume that aesthetic (poetic) discourses can hardly possess their own unique aesthetic function represented by some unique markers of aesthetic discourse; however, they definitely possess the communicative functions common for each discourse type, although their aesthetic potential may be different. Therefore, developing a multimodal approach which considers the communicative functions of modalities can become more efficient since it can scale the function frameworks of more and less aesthetic discourse.

The idea of studying higher-order hierarchy of functions (however, termed differently in different works) to explore the dynamicity and modifications of semiotic systems has received theoretical approbation (Bertalanffy, 1968; Thelen & Smith, 1996, among many) and has recently been incorporated into applied research, for instance in psychology and medicine (Pincus, 2019; Hayes & Andrews, 2020). According to the dynamic systems approach, the semiotic systems evolve over time and under different conditions and self-organize into higher-order functional units. In the present study, the multimodal system of speech and gesture is viewed as self-organizing into more and less aesthetic discourse of cinema and interview. This system operates on the communicative functions of speech and gesture which will presumably display variance in cinema and interview. Contrasting the use of communicative functions in more and less aesthetic discourses, here in cinema and interview, we reveal the prevailing functions (for instance, in speech and gesture) as well as their function frameworks. This approach may provide a structure for conceptualizing and studying varied multimodal patterns in cinematic and interview discourse which apart from speech and gesture integrate other modalities (images, camera, sound, etc.).

Overall, the contributions of this study include: (i) introducing the functional multimodal approach to the study of speech and gesture complexes as specific aesthetic (poetic) means, (ii) revealing the dominant functions of speech and gesture in cinematic discourse as opposed to less aesthetic interview discourse, (iii) disclosing individual variation in multimodal function frameworks that stimulate cinematic aestheticism.

2. Aesthetic multimodality of speech and gesture

Aesthetic semiotics offers different semiotic instruments to explore the aesthetic (poetic) function in artwork focusing on multiple dimensions of the aesthetic. For instance, in Yu. Stepanov's semiotic school (Moscow, Russia) the category of the aesthetic appears in constants (invariants) and variations. The semiotic constants (Stepanov, 2004) are the generalized characteristics serving to foreground the poetic (aesthetic) on the background of non-poetic semiotic formats. It is noticeable that these constants are contrasted with anti-constants (described as anti-concepts), which help reveal the essence of constants. The basic semiotic constant representing the aesthetic potential of artwork is the constant of creativity (and its anti-concept – stereotypicality), which is further expressed in the concepts of dynamicity, individualization, figurativity, synesthesia, deformation, experimental originality, autorepresentation, etc. (Feschenko & Koval, 2014). Thus, we may explore a piece of artwork in terms of its creative and non-creative (stereotypical, conventional) potential, and single out less conventional patterns as being more creative and consequently more aesthetic (Zykova & Kiose, 2020). In the semiotic approaches that develop a phenomenological perspective of the aesthetic, special attention is paid to the role of perception in aesthetic experience. In (2009) P. Bundgaard formulates the principles of meaning making (in artwork) exploring the category of Aesthetic object. Following R. Ingarden (1985), Bundgaard claims that "aesthetic objects are intentionally shifted objects: their qualities do not specify a thing, but presentify a represented object" (Bundgaard, 2009: 50). Here the qualities of objects perceived in an aesthetic mode do not determine it but participate in its presentation via transforming and modifying reality. Therefore, we can notice that aesthetic semiotics still considers aesthetic function via other functions, for instance, via creativity (expressed in dynamicity, figurativity, etc.), embodiment and objectivity.

In aesthetic multimodal discourse, for example, in cinema, exploring these functions in different modalities can become more difficult since their markers will display considerable variation. Besides, multimodality in cinematic discourse is often regarded in two dimensions, the first incorporates the actor's behavior which is treated on the whole as "Gesture" (Agamben, 2000) and the other is mise en scène or "Image". "Gesture" is represented as a complex of verbal and non-verbal communication means such as eye-contact, gestures, posture, facial expressions, intonation, proximity, etc. Multiple works explore "Gesture" and "Image" to disclose the ways of stimulating the

poetics of cinema (Eisenstein, 1964–1968; Ivanov, 1976; Deleuze, 1983, 1985; Auerbach, 2007; Noys, 2014; Harbord, 2015, Zykova, 2020), among the cases revealing the synchronization patterns of "Gesture" and "Image" we could name the study of kinesic behavior and cinematic freeze-frame shot as well as the shot or mise en scène repeated appearance in the film (Mulvey, 2006), the work on speech patterns and metacinematic gestures in films (Ciccognani, 2018), the study of gestures and narration stages (Chare & Watkins, 2020). In the current study, we will consider only speech and gestures as part of "Gesture" complex. Under the term "gesture" we understand movements of the human body parts, e.g., hand or head movements, which convey certain messages as they have meaning (Kendon, 2004). Speech and gestures form a single entity and their co-existence and influence contribute to a better understanding of the communication process (Pease, 1981). According to the theory of growth points, speech and co-occurring gestures appear from the same semantic intent and together they form and organize our discourse (McNeill & Duncan, 2000; Calbris, 2011). There are different types of information which can be expressed by co-speech gestures. Gestures and their use might be regarded not only as a way of communication but also as a marker of creativity of a speaker (Cienki & Mittelberg, 2013). This ability makes them an integral part of cinematic discourse. It is noteworthy that cinematic studies of speech and gesture in terms of their poetic or aesthetic potential are still rare. Interestingly, that in the first gesture studies developed within the frame of multimodal analysis of American kinesic school (and mostly in the works of R. Birdwhistell) the aesthetic function was largely neglected; what mattered was the communicative function of gestures. However, both R. Birdwhistell and S. Eisenstein develop the notion of a kineme, providing two different definitions of it. In Birdwhistell's works kineme is used to represent the structure of kinetic code system resembling the language code system, and the selection and combination of single body motion elements (kines) is guided by the communicative situation, for instance that of a game, dance, theatre plays (Birdwhistell, 1963). Therefore, this view is more communicatively rather than aesthetically oriented. In Eisenstein's works kineme is used as a type (more and less concrete) of art systems in literature and cinema which are reproduced in new artistic forms, still speech and gesture patterns are not explored with regard to their synchronization. J. Kristeva considers "gesture communication as a semiotic text in the process of its production which is not hampered by language structures" (Kristeva, 1969 (2013): 44). In her theory on semanalysis, kinesics becomes a part of "trans-linguistics" (Ibid: 45) where the bodily drives into language may suffice to explore the poetic dimension of language. In recent semiotic works exploring speech and gesture patterns in cinematic discourse, for instance in the works of K. O'Halloran (2004) and A. Lavender (2021) the synchronization patterns of speech and gesture do become the focus of attention, although their aesthetic potential is still explored as

their inherent feature and not as a feature that is not typical of less poetic or artistic forms.

In this study, we explore the aesthetic potential of speech and gesture in cinema and interview via their communicative function frameworks. To reveal them, we 1) assess the activity (frequency) of communicative functions in speech and gesture in two discourse types, 2) reveal the function frameworks of cinematic and interview discourse (applying the methods of variance and regression analysis), 3) explore the individual variance in function frameworks.

3. Methodology. Aesthetic multimodality: A functional framework

In recent years, multimodal (as opposed to semiotic) approach to speech and gesture in cinematic discourse has been applied in different works, for instance in corpus studies (Grishina & Savchuk, 2008). However, they do not aim to reveal the aesthetic potential of discourse. What communicative functions of speech and gesture may serve to explore it?

In terms of speech, we will address the communicative functions outlined in the discourse pragmatic theories; nevertheless, we need to frame these functions on the same grounds. Following T. van Dijk (van Dijk, 1990) who develops the theory of Functional Discourse Analysis modelling functional relations in discourse, we will consider two main discourse functions which are established at the first stage of analysis, the Functional Text Analysis (the second stage is Functional Text-Context Analysis). The first function, semantic and pragmatic function (here termed pragmatic function) describes the discursive nature of microevents explored in (Austin, 1962; Searle & Vanderveken, 1985; Green, 2000, among others); and the other, rhetorical (Mann & Thompson, 1988; Kibrik & Podlesskaya, 2009) or discourse-structuring function (Holler et al., 2020) describes the sequences of microevents (here it is termed discourse-structuring). To frame these functions, we will introduce a common discourse unit of analysis corresponding to a microevent (linguistically expressed by a proposition or its modal frame) to describe argumentation and description and to study the discourse functions of each unit and the discourse functions which help integrate each two units into discourse.

In terms of gesture, we will address the communicative functions of gesture manifested in gesture types. The functional approach to gesture analysis initiated in (Cienki, 2005; Müller, 2005; Cienki & Mittelberg, 2013) was further developed by O. Iriskhanova and A. Cienki (2018) who introduce a functional typology of gestures distinguishing the functions of pragmatic transparency, iconicity, indexicality, symbolism, conventionality, awareness, autonomy, salience, metaphoricity, arbitrariness, semanticity, and recurrence (Ibid: 31). The application of gesture

functions to studying poetic discourse is not a novel idea. For instance, in (2019) O. Iriskhanova explores the aesthetic (poetic) functioning of gestures accompanying speech where the functions of metaphoricity, iconicity and indexicality play the major role. In our study, four gesture functions describing four basic gesture types will be considered, pragmatic, representational, deictic and adaptive, which appear in all previously mentioned typologies irrespective of the discourse type considered.

We will apply the method of contrastive analysis to reveal the communicative functions of speech and hand gesture in more and less aesthetic discourse, here in cinematic discourse and the discourse of interview. We expect to detect the specificity of speech and gesture functions in cinematic discourse which will allow to establish the function framework stimulating aesthetic discourse potential. To avoid the possible clines in individual multimodal behavior, we select the samples of cinematic and interview discourse with the same highly ranked Russian actors performing in monologues. Cinematic discourse in contrast to the interview is a more staged discourse, employing rehearsed multimodal behavior patterns aimed at enhancing the aesthetic potential. Still, when performing in interviews the actors will hardly avoid applying the familiar patterns, however we expect they will be much influenced by non-staged discourse format. Contrastive analysis of function frameworks in cinematic discourse and in interview will help identify 1) variance in the functions of gesture and speech, 2) variance in their combinations, 3) variance in the individual input in terms of both speech and gesture functions and in their combinations. The research has several constraints, first, the selected fragments of cinematic and interview discourse display different combination of genres, here argumentation and description (although we selected the samples very similar in genre), second, the actors performing will appear in the samples in different age. These constraints will definitely affect the function distribution; however, we consider that the study will only benefit if we manage to identify the steady function frameworks despite the variations which are not systemic.

4. Procedure

4.1. Data and taxonomy of communicative functions in speech and gesture

The research data are the film samples (fragments) with 5 male actors performing monologues in very popular Russian films: «Звонят, откройте дверь», А. Митта ("The doorbell rings, open the door" directed by A. Mitta), 1961, «Они сражались за Родину», С. Бондарчук ("They fought for the Country" directed by S. Bondarchuk), 1975, «Судьба человека», С. Бондарчук ("Man's fate" directed by S. Bondarchuk), 1959, «Доживем до понедельника», С. Ростоцкий ("Let us live till Monday" directed by S. Rostotskiy), 1968, «Москва слезам не верит», В. Меньшов ("Moscow does not believe in

tears" directed by V. Menshov), 1980. They feature the monologues of R. Bykov, Yu. Nikulin, S. Bondarchuk, V. Tikhonov and A. Batalov. To contrast the same actors' function frameworks in less staged discourse, we selected 5 interviews taken in different environments, with the interviews with V. Bykov, Yu. Nikulin, and V. Tikhonov taken in the studio, the interview with S. Bondarchuk taken in his study, the interview with A. Batalov taken on the river embankment. All the monologues display realistic character; they are both descriptive and argumentative.

Since the research data were the samples of argumentation and description, to compile a list of functions we addressed the studies which looked into general discourse structuring and pragmatic aspects as well as the studies featuring argumentative and descriptive discourse strategies. Regarding argumentation, the studies account for its ability to express opinions and beliefs (Amossy, 2009), subjectivity and intersubjectivity (Willard, 1989), as well as argumentation schemes like Example, Cause to Effect and Effect to Cause, Practical Reasoning, Inconsistency (Cabrio et al., 2013) employing more fine-grained classification. All these features and schemes display pragmatic functions. Following these classifications and limiting their number to contrast the pragmatic functions in the selected cinematic and interview monologues, we explored 11 pragmatic functions (Opinion, (Emotional) assessment, Stating reasons, consequences, conditions, Contrast, Accusation, Agreement / Disagreement, Appeal to action, Promise, Threat, Comparison, Appeal to power). The studies of description mostly appeal to the topics which are considered in discourse and the way they are presented (Merlo & Mansur, 2004) as well as to the discourse components and types of discourse events including referent types and space construal specifics (Von Stutterheim & Klein, 1989; Longacre, 1996). Hence, we appealed to the pragmatic function of foregrounding the description event components and included their 9 types with three components specifying the type of action performed (Achievement, Process, and State) and six other components (Subject, Object, Action or State, Characteristics, Time, and Place). In terms of the second function, discourse structuring, we addressed the discourse construal theories where the patterns of information foregrounding were presented (Wårwik, 2004; Verhagen, 2007; Iriskhanova, 2014), the pragmatic theories of speech acts and performativity (Austin, 1962; Searle & Vanderveken, 1985; Green, 2000; Kearns, 2006, among others), but mostly to the theories of Rhetorical acts (Mann & Thompson, 1988; Kibrik & Podlesskaya, 2009; Holler et al., 2020). They name the patterns of activating information which are suitable for both argumentation and description, for this reason they are in most cases similar for the discourse stretches. The common functions are Emphasizing opinion or assessment / Emphasizing a discourse component, Self-correcting, Specification, Generalization, Intersubjectivity, Appeal to attention, Chain of arguments / Chain of events, Self-quote, Quoting others, and Figurativity; Rhetorical communication and Initializing communication are applicable only to argumentation.

In Table 1 we present the taxonomy of speech functions further employed for studying speech behavior in cinematic and interview monologues. We applied the coded annotation of speech and gesture functions using the codes 101–123 and 201–219 to annotate speech patterns.

Table 1. The coded taxonomy of speech functions (with codes).

Argumentation		Description	
Pragmatic functions		Pragmatic functions	
Opinion	101	Achievement	201
(Emotional) assessment	102	Process	202
Stating reasons, consequences, conditions	103	State	203
Contrast	104	Accentuated subject	204
Accusation	105	Accentuated object	205
Agreement / Disagreement	106	Accentuated action or state	206
Appeal to action	107	Accentuated characteristics	207
Promise	108	Accentuated time	208
Threat	109	Accentuated place	209
Comparison	110		
Appeal to power	111		
Discourse-structuring functions		Discourse-structuring functions	
Emphasizing opinion or assessment	112	Emphasizing discourse component	210
Self-correcting	113	Self-correcting	211
Specification	114	Specification	212
Generalization	115	Generalization	213
Intersubjectivity	116	Chain of events	214
Appeal to attention	117	New event	215
Rhetorical communication	118	Appeal to attention	216
Initializing communication	119	Self-quote	217
Chain of arguments	120	Quoting others	218
Self-quote	121	Figurativity	219
Quoting others	122		
Figurativity	123		

To explore the gesture functions in the current work in both cinema and interview discourse, we will address their four basic communicative functions (with further specification): pragmatic, representational, deictic and adaptive (Kendon, 1995; Bressem, 2012; Cienki, 2017). Pragmatic gestures have various functions as these gestures are context dependent rather than form dependent, i.e., the meaning might change when their co-occur with different words, although the form can be the same. The main functions of pragmatic gestures are Discourse emphatic, Discourse structuring, Discourse representational, Expressing attitude/evaluation and Contact establishing (McNeill, 1992; Kreidlin, 2002; Calbris, 2011). As it is illustrated in the example below (Figure 1), pragmatic gesture with the function of expressing attitude is used in order to highlight the intensity of the laughter in the described event.

Representational gestures, also known as iconic gestures, are based on the idea of similarity between a hand form and / or its movement and the process or object which it refers to (Streek, 2008). There are several modes of representation which are distinguished in this study: Holding, Molding, Acting, Embodying and Tracing (Müller, 2014). In the example taken from the interview (Figure 2) with R. Bykov, we can see a Tracing gesture used to show the form of a small table in order to highlight its properties.

Another type, deictic gestures, are used to refer to people, objects, notions, places, events, etc. by creating axis in space which connects the speaker and the target of speech (Clark, 2003; McNeil, 2003; Cienki et al., 2014).

Немцы / грохнулись все
The Germans / roared with laughter (Yu. Nikulin, interview discourse)
Figure 1. Pragmatic gesture with the function of expressing attitude.
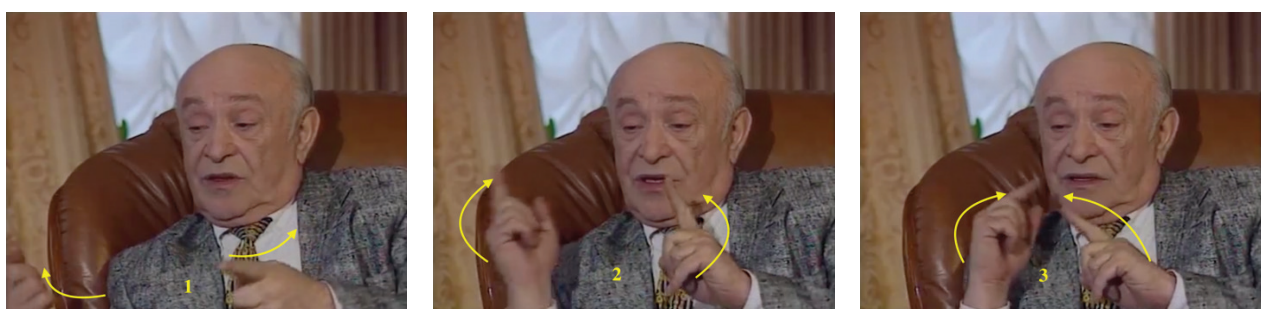
Вот на таком кругленьком столике
On a small table like this (R. Bykov, interview discourse)
Figure 2. Representational tracing gesture.

Мы там бегали / значит / около моста
We were running there / so / near the bridge (A. Batalov, interview discourse)

Figure 3. Deictic pointing gesture.

We highlight two categories of deictic gestures: Pointing and Touching. The example of a deictic gesture (Figure 3) demonstrates its Pointing function. The gesture has a vector, it creates a line which links the speaker and the referent, the place where the action took place, thus attracting and guiding the attention of the listener.

The last type, adaptors, represent some movements, which can be self-oriented (Self-adaptors) such as rubbing one's nose, adjusting glasses, fidgeting one's fingers, etc., or they can be object-oriented (Object adaptors): touching the table in front of the speaker, moving a glass of water, trifling with a pen, etc. These gestures can be used in order to gain control of the situation when the speaker is in the state of distress (Ekman, 2004). As in the example given below (Figure 4), the speaker uses this gesture (touching his upper lip with his index and middle fingers repeatedly) while reflecting on the subject of writing and literature of the past.

И только потому / что эти писатели сумели
And only because / the writers managed to (S. Bondarchuk, interview discourse)

Figure 4. Self-adaptor.

In Table 2 we present the taxonomy of gesture functions further employed for studying gesture behavior in cinematic and interview monologues with the codes 301–314.

In modelling and processing the data, we will apply the method of variance and regression analysis in discourse profiles construal. The notion of discourse profiles

Table 2. The coded taxonomy of gesture functions (with codes).

Gesture functions			
Pragmatic functions		Representational functions	
Discourse structuring	308	Holding	303
Discourse representational	309	Molding	304
Discourse emphatic	310	Acting	305
Expressing attitude/evaluation	311	Embodying	306
Contact establishing	312	Tracing	307
Deictic functions		Adaptive functions	
Pointing	301	Self-adaptors	313
Touching	302	Object-adaptors	314

suggested in construction grammar and structure building frameworks (Ariel, 2004), developmental studies (Singer, 2013) and since recently in multimodal studies (Iriskhanova & Cienki, 2018) and discourse studies (Kiose, 2021) is used here to assess the relative activity of speech and gesture functions in communication.

4.2. Data annotation and procession

The study develops a two-stage procedure. First, it detects the variance in the speech and gesture function frameworks in more aesthetic cinematic discourse and less aesthetic interview discourse. Next, we proceed to contrastive analysis of individual specificity of each actor in stimulating this aestheticism, hypothesizing that despite the individual variance these function frameworks will display a steady character.

The main procedural questions were the selection of the unit of analysis and the annotation format. To select the procedural unit for analyzing speech and gesture complexes, we adopt the view that this unit should be able to manifest (and describe) both description and argumentation. Since the smallest unit capable of manifesting description is a word combination displaying predicate or attribute relations and the smallest unit capable of manifesting argumentation must necessarily be a proposition or its modal frame with either predicate or performative relations, we select the unit with a higher information potential which is the proposition or its modal frame.

For instance, Example (1) has 5 propositions which display different argumentation and description potential:

(1) *Я никогда не был пионером // Но у нас во дворе был форпост // Но так как я никогда не был первым пионером // то я вам расскажу не про себя а про горниста // который жил у нас во дворе*

I have never been a pioneer // But we had a fort-post // But since I have never been a first pioneer // I will tell you not about myself but about a bugler // who lived in our courtyard (R. Bykov, cinematic discourse)

Each of the units was annotated following the same procedure. In case we faced difficulties separating propositions in speech (for instance, when there were hesitations, hedges, interruptions) we adopted the following principle: a unit must necessarily involve either a proposition or a modal frame, therefore all the fragments which do not constitute a proposition or a modal frame are incorporated into the proposition or modal frame that was previously started and was not yet terminated. Example (2) illustrates the described case.

(2) *Спасибо вам // что вы именно сюда приехали // потому что как... здесь снимались «Журавли» // Вот ... эээ ... ну я то ... тут финал снимался ... вот на этом месте буквально*

Thank you // that you came right here // because it was … here… the "Cranes" was filmed // Right here … er… and I … here the final episode was made … at this very place (A. Batalov, interview discourse)

Units 3 (*потому что как... здесь снимались «Журавли»*) and 4 (*Вот ... эээ ... ну я то ... тут финал снимался ... вот на этом месте буквально*) may have combined several propositions, however they are not completed in oral speech.

Annotation was performed in ELAN software, created by Max Plank Institute and used to annotate gestures (https://archive.mpi.nl/tla/elan). We chose it since it allowed to annotate the cinematic and interview shots considering their dynamicity. In Figure 5 we give the annotation example with the annotation tiers showing the role of speech and gesture.

The decision on the gesture function was adopted following the analysis of their form (e.g., the form of the hand: palm up / down, fist, finger extended, etc.; its movement: straight line, circle, wave, etc.; then the direction of it, as well as the space on which it occurred: horizontal or vertical axis, away or towards the speakers, etc. (see Bressem, 2013)) consistent with communicative functions. After that the attention was paid to the semantics of the gesture, which determines its type and corresponding functions in speech, as mere form analysis cannot be used to determine their role in speech due to the polysemantic nature of gestures (Calbris, 2011). The process of annotation included several steps. After uploading a video to ELAN, we created different tiers that represented the parameters that we analyze (speech functions, gesture functions). Figure 5 demonstrates the analysis of the interview with S. Bondarchuk, where the actor discusses the role of cinema and what cinema should look like. In the fragment, we annotated the following proposition in Example (3):

(3) *Это слишком серьезное занятие*

It is a very serious occupation (S. Bondarchuk, interview discourse).

First, we annotated speech functions in terms of argumentation and description. In the above given proposition, we can see that the actor uses Generalization as
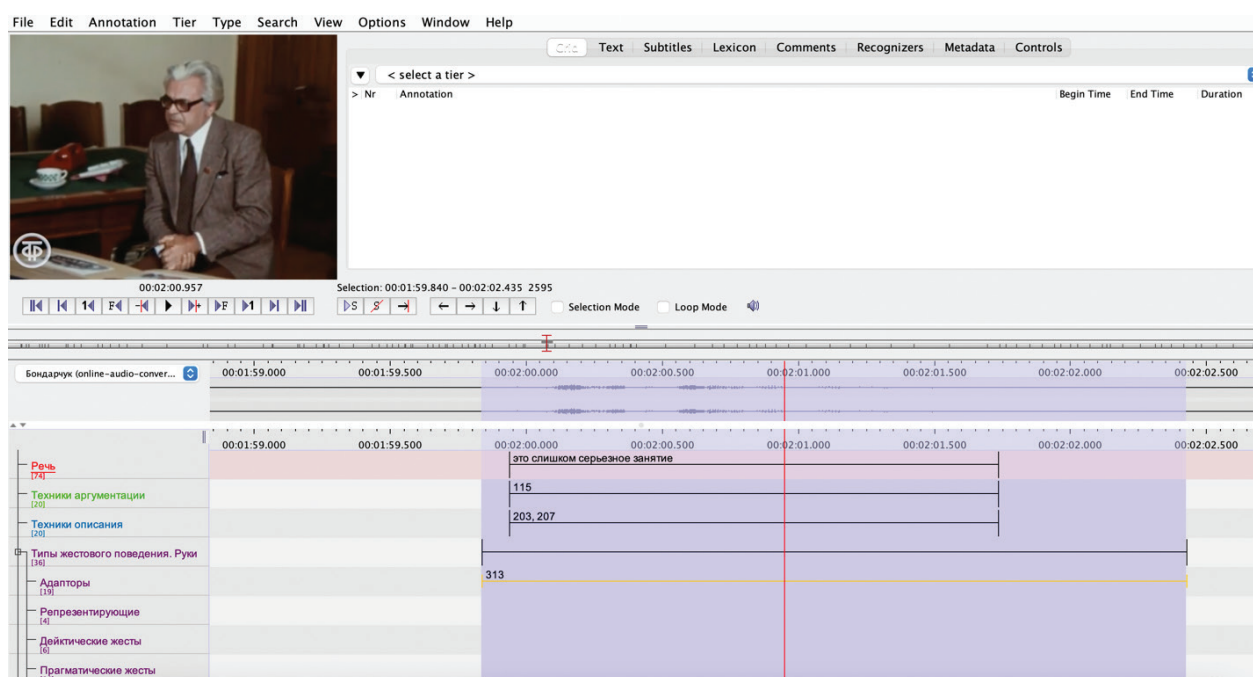

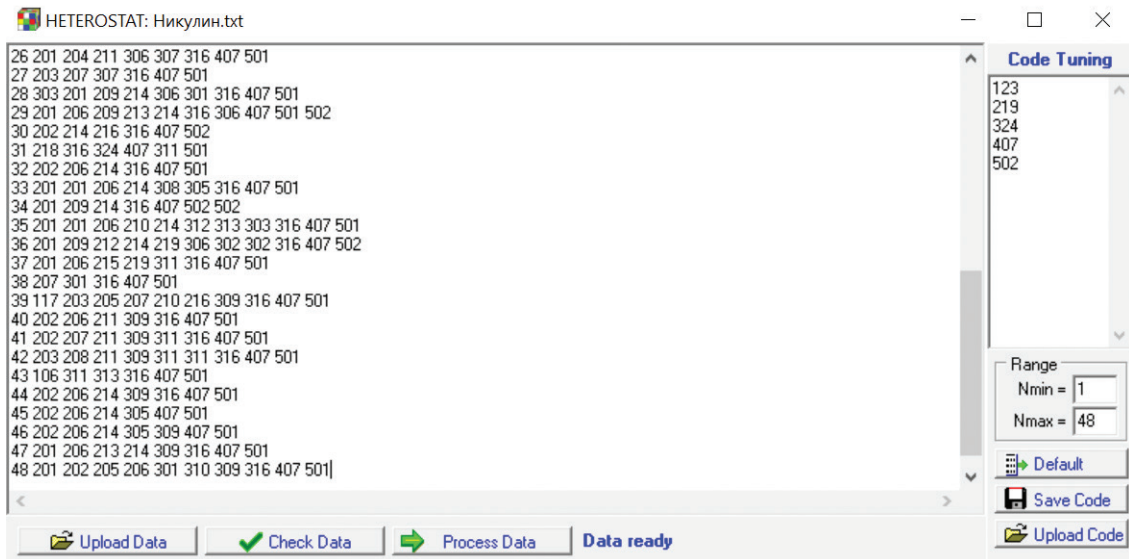**Figure 5.** The annotation process of the interview with S. Bondarchuk.

Figure 6. Window of HETEROSTAT software processing the data. **Note:** Apart from annotating speech and gesture, we also annotated head movements and shot types which will not be considered here.

a pragmatic function of argumentation. S. Bondarchuk discusses the problems that exist in the cinema and summarizes the point by stating that cinema is a serious type of occupation. We pointed out two pragmatic functions of description: State (since this is a description of a state) and Accentuated characteristics (since the modifier *serious* is foregrounded). Next, we addressed the co-speech gestures, and specified the gesture function depending on the performed hand movements and speech. In the example given the actor is rubbing his hands which could indicate that he is using Self-adaptors.

We also employed the speech scripts with full annotations in txt-format. To process the data, we applied HETEROSTAT software (Kiose & Efremov, 2020) which allows to identify the annotated functions activity as well as their contingency.

In Figure 6 we show the HETEROSTAT window processing the data.

As it can be seen, the software checks the data for its consistency with the coded taxonomy, allows to select single or all tiers for further processing. To perform further processing to check the contingency of communication functions of speech and gesture in two discourse types and in individual discourse, we applied JAMOVI software (https://www.jamovi.org).

5. Results

5.1. Speech and gesture functions in cinematic and interview discourse

The analysis of functions in speech and gesture was performed with 10 samples, with 5 of them representing cinematic discourse and 5 representing interview discourse. Each sample lasted approximately 2–3 minutes (min 2:12, max 3:36). The number of annotation units (propositions and modal frames) varied significantly, with min

22, max 66 in cinematic discourse, and min 24, max 71 in interview discourse. The samples displayed finalized communicative events, for instance, in his monologue A. Batalov describes the way the film was made in the very place the interview is taken and presents his arguments on why the film has achieved great success. The total number of the units of analysis in cinematic discourse is 203, and 205 in the interview discourse respectfully, so the data are compatible.

The annotation procedure was carried out by three annotators with two annotators working with the interview discourse, and one annotator working with the cinematic discourse. Then the annotated samples were subjected to crosscheck and Cohen's Cappa statistical coefficient (Landis & Koch, 1977) was applied (https://idostatistics.com/cohen-kappa-free-calculator/#risultati) to evaluate the agreement between the annotators to verify the validity of results.

We processed the Cohen's Cappa separately for two discourse types. Since 56 functions of speech and gesture were annotated and the number of units was 203 and 205 correspondingly, we received a total number of annotation responses equal to 11,368 in cinematic discourse, and 11,480 in interview discourse. In terms of cinematic discourse, both groups of judges agreed to decide 1402 cases in the positive and 9866 cases in the negative with 90 cases decided in the positive by the first annotator and 439 by the other annotator group. The agreement coefficient is 95.52%, and Cohen's k = 0.82, which is almost perfect agreement. In terms of interview discourse, both groups of judges agreed to decide 1478 cases in the positive and 9720 cases in the negative with 211 cases decided in the positive by the first annotator and 387 by the other annotator group. The agreement coefficient is 94.93%, and Cohen's k = 0.8, which is also almost perfect agreement. We then voted for including the functions (since there were three annotators) and the final results are the following: the total activity of speech and gesture

functions is 863 in cinematic discourse (728 in speech and 135 in gestures), and 1118 in interview discourse (826 in speech and 292 in gestures).

In Tables 3, 4 we give the results of contrastive function activity in speech and gesture in two discourse types.

The difference in all functions of speech and gesture seems significant but it is not statistically verified. With $F_{(1, 14)} = 0.085$ at $p = 0.775$, we cannot reject the null hypothesis that there is no difference between the functions in cinematic and interview discourse. However, the variance in single functions (56 functions) is significant ($F_{(1, 110)} = 4.92$, $p = 0.027$), which proves that specifying functions was an effective solution. It is noticeable that significant differences are observed in gesture only if we consider speech and gesture separately. With $F_{(1, 26)} = 7.14$, $p = 0.008$ in gesture and $F_{(1, 26)} = 0.947$, $p = 0.33$ in speech, we can claim that gesture distribution is of higher importance in more and less aesthetic discourse multimodal construal.

Tables 3. Function activity in speech.

Discourse types / Functions	Speech			
	Pragmatic		Discourse-structuring	
	Argumentation	Description	Argumentation	Description
Cinematic	155	402	105	66
Interview	167	461	73	125

Tables 4. Function activity in gesture.

Discourse types / Functions	Gesture			
	Deictic	Representational	Pragmatic	Adaptors
Cinematic	17	16	89	13
Interview	36	76	127	53

Therefore, we now move on to discussing single functions of gesture in two discourse types in more detail. Figure 7 gives the mapped gesture profiles of the gesture functions 301–314 which represent 4 basic functions.

Since these functions might be contingent on the speech functions, we performed regression modelling to reveal the predicting gesture functions in multimodal discourse. Regression modelling is an efficient method to cope with the problem of mixed effects of functions which is a typical case of construal in speech. However, we frequently deal with a problem of aliased coefficients. In the current study, however, these effects were surprisingly scarce which means that functional instrument works well for the needs of multimodal discourse analysis. We will present the model performance summary statistics for the most active functions, first in cinematic discourse, next for the interview discourse.
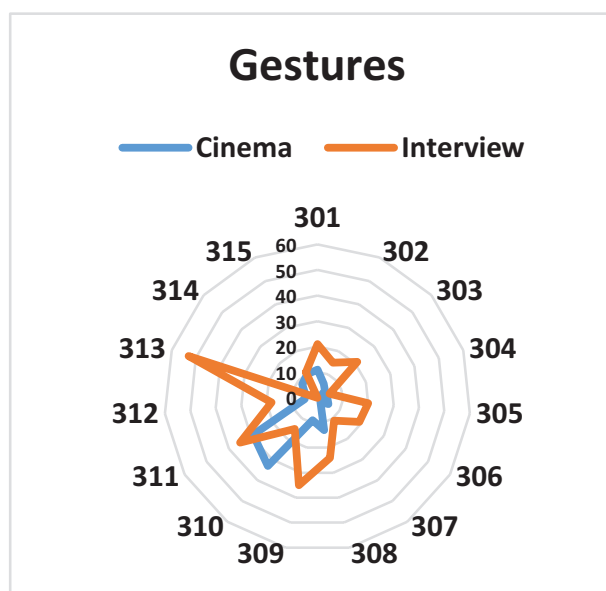

Figure 7. Mapping gesture profiles in cinematic and interview discourse

In cinematic discourse, two gesture functions with the highest activity are the pragmatic functions, Discourse emphatic (310), and Expressing attitude / evaluation (311). Their model performance statistics is given in Table 5. There were 36 non-aliased functions, in the table we give the statistics on the best predictors only.

As seen from Table 5, the best predictors of Discourse emphatic gestures are Specification, Accentuated time and Emphasizing discourse component which relate to description. In terms of Expressing attitude gestures, the best predictors are (Emotional) assessment, Threat, Stating reasons, consequences, conditions which relate to argumentation, and Achievement relating to description. The results show that the predictability of Regression model is quite high ($R^2 = 0.341$ and 0.359).

In the following example (Figure 8) the actor cites Ivan Karamazov from F. Dostoevsky's "The Brothers Karamazov", who denied accepting the fact that the language of arms and guns was the only means by which the dialogue with the Tsar could take place. Figure 8 illustrates two propositions introduced by the actor. In proposition (1) *Все не верил* ((He) still didn't believe) the actor emphasizes the discourse component in description, whereas in proposition (2) the actor specifies the unwillingness to perform the action (*Не хотел верить* (He) **didn't want** to believe). It is to point out that while introducing both propositions the

Table 5. Regression Model predicting Pragmatic gestures, Discourse emphatic and Expressing attitude.

Predictor	310 Discourse emphatic $R^2 = 0.341$				311 Expressing attitude/evaluation $R^2 = 0.359$			
	Estimate	SE	t	p	Estimate	SE	t	P
Intercept	0.077	0.091	0.853	0.395	-0.133	0.078	-1.7	0.091
(Emotional) assessment	-	-	-	-	0.267	0.076	3.5	<.001
Stating reasons, consequences, conditions	-	-	-	-	0.181	0.072	2.507	0.013
Threat	-	-	-	-	-1.535	0.424	-3.623	<.001
Achievement	-	-	-	-	0.198	0.08	2.484	0.014
Accentuated time	0.168	0.078	2.144	0.034	-	-	-	-
Emphasizing discourse component	0.358	0.122	2.938	0.004	-	-	-	-
Specification	1.779	0.423	4.2	<.001	-	-	-	-

(1) *Все не верил*
(He) still didn't believe
(V. Tikhonov, cinematic discourse)

(2) *Не хотел верить*
(He) didn't want to believe
(V. Tikhonov, cinematic discourse)

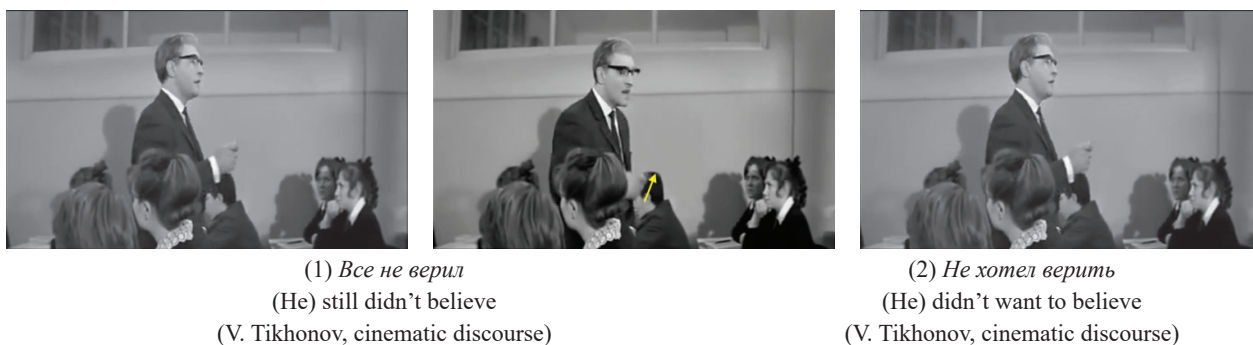**Figure 8.** Discourse emphatic gestures with Discourse structuring functions of Description (Emphasizing discourse component and Specification).

Не дай Бог / ребята услышали бы
God forbid / the fellows would hear (Yu. Nikulin, cinematic discourse)

Figure 9. Expressing attitude / evaluation gestures with the Pragmatic functions of argumentation and description.

actor used Discourse emphatic gestures to stress the importance of the ideas he was delivering to the pupils.

The best predictors for Expressing attitude / evaluation gestures are the pragmatic functions of argumentation and only one function of description. Figure 9 demonstrates an interesting example of the way the character (performed by Yu. Nikulin) introduces the phrase *Не дай Бог ребята услышали бы* (God forbid the fellows would hear (Yu. Nikulin, cinematic discourse), which has multiple pragmatic functions of argumentation (Emotional assessment and Stating reasons, consequences, conditions) and two pragmatic functions of description (Achievement and Accentuated action or state). The proposition co-occurs with two types of gestures: Expressing attitude / evaluation and Pointing, which on the one hand intensify the attitude of the character towards the message sent in the proposition, and, on the other hand, the character makes a reference to those who took part in the action.

The received data may provide evidence in favor of a specific functional feature of aesthetic multimodal discourse, which is the rigid (perhaps, because much practiced and rehearsed) correspondences of different gesture functions with different types of speech functions in argumentation and description. We will check this assumption contrasting the results with the model performance statistics of the interview discourse.

Three gesture functions with the highest activity are two pragmatic functions, Discourse representational and Expressing attitude / evaluation; and also one Adaptive function, Self-adaptors. We will present the results on only two functions, Expressing attitude / evaluation (311), which is the same with the cinematic discourse, and Self-adaptors (313), which is of a different function type not active in the cinematic discourse. Their model performance statistics is given in Table 6. There were 39 non-aliased functions, in the table we give the statistics on the best predictors only.

As seen from Table 6, results show that the predictability of Regression model is lower in both cases than the predictability in cinematic discourse ($R^2 = 0.282$ and 0.256). The best predictors are fewer; the best predictor for Expressing attitude/evaluation in gestures in interview discourse is Self-correcting, see Figure 10:

Table 6. Regression Model predicting Pragmatic gestures, Discourse emphatic and Expressing attitude.

Predictor	311 Expressing attitude/ evaluation $R^2 = 0.282$				313 Self-adaptors $R^2 = 0.256$			
	Estimate	SE	t	p	Estimate	SE	t	P
Intercept	0.056	0.096	0.588	0.558	0.308	0.114	2.71	0.007
Opinion					0.437	0.145	3.009	0.003
Intersubjectivity					1.366	0.616	2.218	0.028
Self-correcting	0.55	0.187	2.943	0.004				

Там наступали вроде бы // разведка боем

(They) were allegedly attacking // conducting reconnaissance in force (Yu. Nikulin, interview discourse)

Figure 10. Expressing the attitude/evaluation gestures with the Pragmatic function of argumentation (Self-correcting).

The sample given in Figure 10 demonstrates that the actor (Yu. Nikulin) expresses some level of uncertainty while proposing the argument that the troops were attacking and self-corrects the argumentation saying that it was the reconnaissance in force. While expressing the idea the actor uses Expressing the evaluation gesture which could symbolize his evaluation of the actions taking place, basing on the information he had (but this piece of information possibly requires verification). This is perhaps the least expected choice since it does not manifest high speech pragmatism to comply with the gesture pragmatism.

There are no other predictors, which makes us think that the interview discourse being less staged allows much more freedom in multimodal function frameworks. In terms of Self-adaptors, the situation is less peculiar. The best predictors are Opinion and Intersubjectivity, which are argumentation Pragmatic and Discourse-structuring functions, and the appearance of Self-adaptors in such situations was most expected.

The example shown in Figure 11 illustrates the gestural behavior of the speaker (S. Bondarchuk) expressing opinion and intersubjectivity.

As far as S. Bondarchuk introduces his own opinion, he is gaining control over the situation by using Self-adaptors.

The contrastive results presented here may suffice to deduce several structural types of function frameworks.

REPLICATION – SINGULARITY. This structure integrates several functions of the same type. It can have three types. Type 1 describes the framework of intensifying the same function in both modalities. In the present study we observed Type 1 with pragmatic function which can be demonstrated in both speech and gesture. Interestingly, in cinematic discourse the gesture functions (at least the ones we explored in terms of predictability) seem to comply with similar functions in speech as if to enhance them additionally, whereas in the interview we did not observe this effect. Therefore, cinematic (and more aesthetic) discourse mostly exploits REPLICATION function structure. Type 2 describes the framework of in-

(Кинематограф) не должен превращаться // нууу // есть такое грубое слово / киношку

(Cinema) should not transform into // errr // there is a rude word / into a fleapit (S. Bondarchuk, interview discourse)

Figure 11. Self-adaptors expressing Opinion and Intersubjectivity.

tensifying the same function within one modality. This happens when the same function is activated several times within one proposition or modal frame. Surprisingly, cinematic discourse shows less preference for multiple pragmatic function expression in description, with very scarce pragmatism of the subject (in contrast with the interview discourse). This result might be explained by the spontaneous nature of the interview discourse and multiple pragmatic patterns chosen by the speakers to describe the events and their participants. Type 3 describes the framework of intensifying multiple sub-functions of one function. The data have shown that this function framework is frequent in both discourse types. Therefore, cinematic discourse mostly exploits REPLICATION function framework, whereas interview discourse shows preference for SINGULARITY.

REINFORCEMENT – SELF-SUFFICIENCY. This structure integrates several contingent functions in multimodal discourse. As we have shown above, it has 2 types of realization. Type 1 describes the framework of reinforcing the function of one modality with the functions of the other modality which display rigid contingency. REINFORCEMENT is more frequently present in cinematic discourse, whereas interview discourse has fewer contingent functions. Interview discourse spontaneity may also account for it. Type 2 describes the framework of reinforcing the functions within one modality. We may notice here that neither discourse exhibits this type of REINFORCEMENT. The functions do not display alignment effects, therefore their individual input into the process of multimodal discourse construal needs no other evidence: these functions are self-sufficient in determining the communicative potential of the modalities considered.

LIMITATION – EXTENSION. This structure is of a different type; it describes the function framework of constraining and widening the number of functions which are manifested in multimodal discourse. Therefore, this structure can be detected only in contrastive analysis. Type 1 appears when the number of functions is significantly smaller in one of the modalities. Unexpectedly, this type is more typical of cinematic discourse which employs far fewer gestures of specific types, for instance, of Self-adaptors. They are mostly present in interview, since they are the most natural and uncontrolled movements, used in our speech when the speaker might feel more stressed and in need of exercising control over the situation (Ekman, 2004) or they can be markers of turn taking (Żywiczyński et al., 2017). Type 2 appears when this structure is manifested in both modalities. In this study, we revealed that cinematic discourse uses fewer pragmatic speech functions of description and fewer pragmatic functions of Expressing attitude / evaluation.

Next, we will find out whether these function frameworks appear in individual multimodal discourse of the 5 actors. This analysis might also reveal some other function frameworks relevant for assessing the aestheticism of cinematic discourse.

5.2. Individual variations of speech and gesture in cinematic and interview discourse

At the second stage we turn to the individual differences in multimodal discourse which may be found if we contrast speech and gesture in two discourse types, cinematic and interview, performed by the same actors. This procedure allows to detect the individual variations which may or may not fall within the function frameworks of either cinematic or interview discourse and consequently, may help specify them and show them in more detail. To proceed, we contrast the multimodal discourse profiles within each discourse. The idea that we entertain is that the profile differences will display some similar tendencies which will suffice to claim that they are more typical of more aesthetic discourse.

At the first step, we identify whether the differences in the functions distribution by each actor are significant and therefore dependent on individual multimodal discourse. In cinematic discourse, variance analysis in three sets, speech functions in argumentation, speech functions in description, functions in gesture revealed the following results: for argumentation (101–123) $F(4, 22) = 19.2$ at $p < .001$, for description (201–219) $F(4, 18) = 23.9$ at $p < .001$, for gestures (301–314) $F(4, 13) = 15.7$ at $p = 0.003$. In interview discourse, variance analysis revealed the following results: for argumentation (101–123) $F(4, 22) = 17.9$ at $p = 0.001$, for description (201–219) $F(4, 18) = 36.4$ at $p < .001$, for gestures (301–314) $F(4, 13) = 19.9$ at $p < .001$. The results suffice to claim that in both discourse types, the functions of speech and gesture display significant variance; interestingly, the highest variance values appeared in the functions of description, especially in the interview discourse. The lowest variance was in the functions of gestures in cinematic discourse, which specifies our earlier findings in the way that gesture functions are not only synchronized with similar speech functions but are also more restricted and allow fewer alternatives in their selection.

At the second step, we find out whether the speech and gesture functions distribution is similar with the actors. To find the answer, we contrasted the function distribution in three data sets (functions in argumentation, functions in description, and functions in gesture) in both discourse types, cinematic and interview. To do this, we introduced a grouping variable, 0 and 1 (for cinematic and interview discourse). In Tables 7, 8 we present the results for each of the three datasets and for each actor.

In terms of functions in argumentation, there is one actor, V. Tikhonov, whose speech behavior displays significant variance. This may happen only because his mono-

Table 7. Variance in functions in argumentation in two discourse types.

Actor	F	df1	df1	p
A. Batalov	0.43	1	43.8	0.515
S. Bondarchuk	1.783	1	44	0.189
R. Bykov	1.948	1	27.6	0.174
Yu. Nikulin	2.348	1	43.5	0.133
V. Tikhonov	6.801	1	26.9	0.015

Table 8. Variance in functions in description in two discourse types.

Actor	F	df1	df1	p
A. Batalov	0.4352	1	34.6	0.514
S. Bondarchuk	0.0429	1	33.5	0.837
R. Bykov	3.7269	1	26.8	0.064
Yu. Nikulin	0.4049	1	35.8	0.529
V. Tikhonov	0.9128	1	27	0.348

Table 9. Variance in gesture functions in two discourse types.

Actor	F	df1	df1	p
A. Batalov	4.099	1	23.3	0.055
S. Bondarchuk	3.081	1	16	0.098
R. Bykov	9.252	1	22.8	0.006
Yu. Nikulin	10.045	1	20.2	0.005
V. Tikhonov	0.421	1	26	0.522

Table 10. Variance in gesture functions with the actors in two discourse types.

Gesture function	F	df1	df1	p
Pointing	0.935	1	7.42	0.364
Touching	1.514	1	4.56	0.278
Holding	6.881	1	4.72	0.05
Molding	2.667	1	5.54	0.158
Acting	2.359	1	4.27	0.195
Embodying	1.13	1	4.24	0.345
Tracing	3.6	1	4.57	0.122
Discourse structuring	1.561	1	5.68	0.26
Discourse representational	4.971	1	7.08	0.061
Discourse emphatic	2.331	1	6.95	0.171
Expressing attitude/evaluation	0.196	1	6.41	0.672
Contact establishing	6.377	1	7.86	0.036
Self-adaptors	11.157	1	4.41	0.025
Object-adaptors	NaN	1	NaN	NaN

logues display a different communicative character which is true since his speech patterns clearly display fewer opinion and emotional assessment in interview. However, in terms of description functions, all 5 actors displayed similar speech behavior. The most interesting results appear with the distribution of gesture functions. Three actors out of 5 displayed significant differences in gesture functions' selection in cinematic and interview discourse. The highest variance is attributed to Yu. Nikulin and R. Bykov. The results point out that these are mostly gestures which manifest variance in multimodal discourse of actors in more and less aesthetic discourse and not the speech functions. In a way it supports our previous findings that the inventory of gesture functions in cinematic discourse is more restricted, and there exists higher alignment between gesture and speech functions in cinematic discourse.

Therefore, it is possible that these are particular / specific gesture functions that may serve as indicators of aestheticism of cinematic discourse. To find out which functions may fulfill this role and appear in gesture behavior of all the 5 actors, we now turn to analyzing the gesture functions in the actors' discourse. To perform it, we again introduce a grouping variable, 0 and 1 (for cinematic and interview discourse). In Table 10 we present the variance results for each gesture function.

The results show that there are three gesture functions that have statistically different distribution with 5 actors and in two discourse types, they are Holding gestures, Contact establishing gestures and Self-adaptors; Discourse representational gestures also play a role. It means that other gesture functions work similarly in both discourse types, and the aesthetic specificity of gesture in cinematic discourse mostly lies within these four types.

The example given in Figure 12 illustrates gestural behavior of the actor introducing one proposition that has multiple pragmatic functions of description.

At the beginning of the sentence (1) while describing the situation, Yu. Nikulin uses Contact establishing gestures to attract attention of the audience to the case being described. Afterwards while giving the description of the fear (2) perceived by him (the German soldier) the actor uses Self-adaptors. At the end of the proposition the actor uses the Holding gesture to show the silence (3), in which the described situation took place. This example demonstrates the variety of co-speech gestures functions used in one proposition.

We selected a fragment employing two propositions with multiple pragmatic functions of description in cinematic discourse with the same actor (Figure 12) to contrast it with the situation given in Figure 11.

1)	2)	3)
Дело такое получилось, что	видно от страха он	в этой тишине

Дело такое получилось // что видно от страха / он в этой тишине громко пукнул
Such a thing happened // that maybe because of fear / he farted loudly in this silence (Yu. Nikulin, interview discourse)

Figure 12. Co-speech gestures used by Yu. Nikulin in interview discourse.

(1) *Сотни полторы в одной избе набилось*
Around fifteen hundred people crammed into one house
(Yu. Nikulin, cinema discourse)

(2) *Спали и валетами, и сидя, и по-всякому*
We slept top and tail, seated, differently
(Yu. Nikulin, cinema discourse)

Figure 13. Co-speech gestures used by Yu. Nikulin in cinematic discourse.

In comparison with interview discourse, Yu. Nikulin does not use many gestures while giving the description in cinematic discourse. In proposition (1) no gestures were identified, whereas in proposition (2) the actor uses one gesture with the function of expressing attitude / evaluation while describing the way a lot of people were forced to sleep in one house.

It is also noticeable that Holding gestures, Contact establishing gestures, Self-adaptors and Discourse representational gestures irregularly appear in cinematic discourse, being common only for the discourse of interview. At the same time, other gestures appear frequently in both discourse types, for instance, Discourse emphatic gestures. Considering this fact, we distinguish one more function framework structure revealed in contrastive analysis of individual difference – that is REGULARITY – IRREGULARITY. It describes the function frameworks with regular or irregular function distribution within the multiple samples of multimodal discourse of the same or contrasting types. At this stage we cannot account for providing a sufficient description of this structure since the research data need to be much extended to become relevant. Nevertheless, even with these data we still managed to detect the regular patterns of function distribution which may evidence in favor of this structure. The function distribution in cinematic discourse has shown more regularity since it was better predicted. As we have revealed, the lowest variance was observed with the functions of gestures in cinematic discourse and this variance displayed regularity among all 5 male actors.

6. Concluding remarks

The study develops a methodological approach of contrasting multimodal discourse as displaying higher or lower aestheticism, here manifested in speech and gesture in cinematic and interview discourse. Following R. Goodrich (1997), we intended to prove that aesthetic discourse does show the "clusters of concomitant features", although it does not have its own poetic (or aesthetic) function markers. As opposed to the aesthetic semiotic approach specifying the aesthetic functions of "Gesture" and "Image" (Agamben, 2000; Auerbach, 2007; Noys, 2014, among others), we explore the distribution of communicative functions in speech and gesture employed by the actor, as part of multimodal discourse. Aesthetic multimodality analysis is built on the functional framework which is frequently used in analyzing speech (for example, in Mann & Thompson, 1988; van Dijk, 1990; Kibrik & Podlesskaya, 2009), however it has only recently been integrated into gesture studies (Cienki, 2005; Müller, 2005; Cienki & Mittelberg, 2013; Iriskhanova & Cienki, 2018).

We hypothesized that higher and lower aestheticism can be established via communicative functions in multimodal discourse of speech and gesture, and therefore addressed more and less aesthetic discourse types, the cinematic and the interview discourse pursuing an idea of elaborating a suitable instrument for studying their variance. The contrastive study revealed different activity of pragmatic and discourse-structuring functions in speech, and pragmatic, deictic, representational and adaptive functions in gesture. At the same, we found out that the function activity cannot be viewed as a reliable criterion for distinguishing the discourse types in terms of the multimodal behavior, since the activity of all functions does not display significant variance. For this reason, we explored the function framework structure (Bertalanffy, 1968; Thelen & Smith, 1996, among many) to contrast the multimodal discourse types. The study has revealed 4 function frameworks REPLICATION – SINGULARITY, REINFORCEMENT – SELF-SUFFICIENCY, LIMITATION – EXTENSION, REGULARITY – IRREGULARITY. They allow to scale multimodal behavior in the cinematic and interview discourse. The aestheticism of cinematic discourse manifests itself in higher replication (the gesture functions (at least the ones we explored in terms of predictability) seem to comply with similar functions in speech), higher reinforcement (it has more contingent functions than interview dis-

course), limitation (fewer gesture functions are exploited), regularity (the activity of several functions is contingent on the discourse type with all the actors considered).

The devised function framework may be applicable to contrastive discourse studies exploring multimodal resources and can help describe other social or cultural functions besides the aesthetic function. Additionally, it may reveal significantly more regularities and specifics of the aesthetic function implemented in multiple aesthetic discourse formats. The obtained results may be applicable to studying multimodal behavior as part of different discourses, for instance the cinematic discourse employing not only "Gesture" (speech and gesture of actors) but also "Image" (cinema shots).

Acknowledgements

References

Agamben, G. (2000). Notes on Gesture. In V. Binetti, C. Casarino (Eds.) *Means without end: Notes on politics* (pp. 49–59). Minneapolis-London, University of Minnesota Press.

Amossy, R. (2009). Argumentation in Discourse: A Socio-discursive Approach to Arguments. *Informal Logic*, 29(3), 252–267. https://doi.org/10.22329/il.v29i3.2843

Ariel, M. (2004). Accessibility marking: discourse functions, discourse profiles, and processing cues. *Accessibility in Text and Discourse Processing: A Special Issue of Discourse Processing*, 37(2), 91–116. https://doi.org/10.1207/s15326950dp3702_2

Auerbach, J. (2007). *Body shots: Early cinema's incarnations*. Berkeley-London, University of California Press. https://doi.org/10.1525/9780520941199

Austin, J. L. (1962). *How to do Things with Words*, 2ⁿᵈ ed. Cambridge, MA, Harvard University Press.

Bertalanffy, L. (1968). *General System Theory. Foundations, Development, Applications*. N.Y.: University of Alberta.

Bressem, J. (2012). Repetitions in gesture: Structures, functions, and cognitive aspects. Ph.D. dissertation, Faculty of Social and Cultural Sciences, European University Viadrina, Frankfurt (Oder).

Bressem, J., Ladewig, S. H., & Müller, C. (2013). Linguistic annotation system for gestures. Berlin: De Gruyter Mouton. https://doi.org/10.1515/9783110261318.1098

Bundgaard, P. F. (2009). Toward a Cognitive Semiotics of the visual artwork — Elements of a grammar of intuition. *Cognitive Semiotics*, 5, 42–65. https://doi.org/10.1515/cogsem.2009.5.fall2009.42

Birdwhistell, R. (1963). *Conceptual Bases and Applications of the Communicational Sciences*. The University of California.

Blom, N. (2020). Multimodal hermeneutics: Aesthetic response to literature in the English language Arts classroom. Doctoral Thesis. Columbia University.

Cabrio, E., Tonelli, S., & Villata S. (2013). From Discourse Analysis to Argumentation Schemes and back: Relations and differences. In J. Leite et al. (Eds.) *CLIMA XIV* (pp. 1–17). Springer-Verlag Berlin Heidelberg. https://doi.org/10.1007/978-3-642-40624-9_1

Calbris, G. (2011). *Elements of Meaning in Gesture*. John Benjamins Publishing Company. https://doi.org/10.1075/gs.5

Ciccognani, M. (2018). Metacinematic gestures an investigation of the productionist aspect of self-reflexive films. Doctoral Thesis. Leicester, University of Leicester.

Cienki, A. (2005). Image schemas and gesture. In J. Grady (Ed.) *From Perception to Meaning: Image Schemas in Cognitive Linguistics* (pp. 421–442). New York, Mouton de Gruyter. https://doi.org/10.1515/9783110197532.5.421

Cienki, A. (2017). Ten lectures on Spoken language and Gesture from Perspective of Cognitive Linguistics. *Issues of Dynamicity and Multimodality* (pp. 24–44). Leiden, Boston, Brill. https://doi.org/10.1163/9789004336230

Cienki, A., Mittelberg, I. (2013). Creativity in the forms and functions of spontaneous gesture with speech. In T. Veale, K. Feyaerts, C. Forceville (Eds.) *The agile mind: A multi-disciplinary study of a multi-faceted phenomenon* (pp. 231–252). Berlin, De Gruyter Mouton. https://doi.org/10.1515/9783110295290.231

Cienki, A. Bietti, L. & Kok, K. (2014). Multimodal alignment during collaborative remembering. *Memory Studies*, 7, 354–369. https://doi.org/10.1177/1750698014530624

Chare, N., Watkins, L. (Eds.) (2020). *Gesture and Film. Signalling New Critical Perspectives*. London, Routledge.

Clark, H. (2003). Pointing and placing. In S. Kita (Ed.) *Pointing: Where Language, Culture and Cognition Meet* (pp. 242–268). Lawrence Erlbaum Associates, Publishers. Mahwah, New Jersey, London.

Eco, U. (1984). *Semiotics and the Philosophy of Language*. Bloomington, Indiana University Press. https://doi.org/10.1007/978-1-349-17338-9

Eisenstein, S. (1964–1968). *Selected Papers*. Vol. 1–6. Moscow, Iskusstvo.

Ekman, P. (2004). Emotional and conversational nonverbal signals. In J. M. Larrazabal and L.A Perez Miranda (Eds.) *Language, knowledge, and representation* (pp. 39–50). Alphen aan den Rijn, Kluwer. https://doi.org/10.1007/978-1-4020-2783-3_3

Feschenko, V. V., Koval, O.V. (2014). *The Creation of a Sign. Essays on Linguoaesthetics and the Semiotics of Art*. Moscow, Languages of Slavic Culture.

Goodrich, R. (1997). On Poetic Function: Jakobson's Revised 'Prague' Thesis. Literature & Aesthetics, 7, 54–66.

Green, M. (2000). Illocutionary force and semantic content. *Linguistics and Philosophy*, 23, 435–473. https://doi.org/10.1023/A:1005642421177

Grishina, E. A. & Savchuk S. O. (2008). Corpus of Spoken Language in National Corpus of the Russian Language. Project. Computational Linguistics and Intellectual Technologies, 7(14), 125–132.

Halloran, K. (2004). Visual semiosis in film. In K. Halloran (Ed.) *Multimodal Discourse Analysis. Systemic Functional Perspectives* (pp. 109–130). London, New York, Continuum.

Harbord, J. (2015). Agamben's cinema: Psychology versus an ethical form of life. *NECSUS European Journal of Media Studies,* 4(2), 13–30. https://doi.org/10.5117/NECSUS2015.2.HARB

Hayes, A. M., Andrews, L. A. (2020). A complex systems approach to the study of change in psychotherapy. *BMC Medicine*, 18, Article number 197. https://doi.org/10.1186/s12916-020-01662-2

Holler, A., Suckow, K. & de la Fuente, I. (Eds.) (2020) Information Structuring in Discourse. Boston: Brill. https://doi.org/10.1163/9789004436725

Ingarden, R. (1985). *Selected Papers in Aesthetics*. Munich: Philosophia Verlag.

Iriskhanova, O. K. (2014). *Focus Games in Language: Semantics, Syntax and Pragmatics of Defocusing*. Moscow, The Languages of Slavic Culture.

Iriskhanova, O. K. (2019). Cognitive poetics of gestures. *Cognitive Studies of Language*, 36, 72–86.

Iriskhanova, O. K. & Cienki, A. (2018). The semiotics of gestures in cognitive linguistics: Contributions and challenges. *Issues in Cognitive Linguistics*, 4, 25–36. https://doi.org/10.20916/1812-3228-2018-4-25-36

Ivanov, V. V. (1976). *Essays on the History of Semiotics in the USSR*. Moscow, Nauka.

Jakobson, R. O. (1960). Closing Statement Linguistics and Poetics. In T. A. Sebeok (Ed.) *Style in Language* (pp. 350–377). Cambridge, Massachusetts.

Kearns, J. (2006). Conditional assertion, denial, and supposition as illocutionary acts. *Linguistics and Philosophy*, 29, 455–485. https://doi.org/10.1007/s10988-006-0007-y

Kendon, A. (1995). Gestures as illocutionary and dis- course structure markers in Southern Italian conversational. *Journal of Pragmatics*, 23, 247–279. https://doi.org/10.1016/0378-2166(94)00037-F

Kendon, A. (2004). *Gesture: Visible Action as Utterance*. Cambridge University Press. https://doi.org/10.1017/CBO9780511807572

Kibrik, A. A. & Podlesskaya, V. I. (2009). Night Dream Stories: A Corpus Study of spoken Russian discourse. Moscow: Languages of Slavonic Culture.

Kiose, M. I. (2021). Linguistic creativity and discourse profiles of English language children's novels. *Russian Journal of Linguistics*, 25(1), 147–164. https://doi.org/10.22363/2687-0088-2021-25-1-147-164

Kiose, M. I. & Efremov A. A. (2020). HETEROSTAT. Program for discourse parameters variance calculation. Registered 21.09.2020. Registration Number 2020661240.

Kostoulas, Th., Muszyński, M., Chaspari, Th., Amelidis, P. (2020). Multimodal affect and aesthetic experience. *ICMI '20: Proceedings of the 2020 International Conference on Multimodal Interaction* (October 2020), 888–889. https://doi.org/10.1145/3382507.3420055

Kreidlin, G. E. (2002). *Nonverbal semiotics. Body language and natural language*. Moscow, Novoe literaturnoe obozrenie.

Kristeva, J. (1969). *Recherches Pour Une Sémanalyse*. Éditions du Seuil. Reprinted in Kristeva, J. (2013) Semiotics. The Studies on Semanalysis. Moscow, Akademicheskii proekt.

Landis, J. R. &. Koch, G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics* 33, 159–174. https://doi.org/10.2307/2529310

Lavender, A. (2020). Multimodal acting and performing. *Beyond Media Borders*, 1, 113–140. https://doi.org/10.1007/978-3-030-49679-1_3

Longacre, R. (1996). *The Grammar of Discourse*. New York and London: Plenum Press. https://doi.org/10.1007/978-1-4899-0162-0

Mann, W.C. & Thompson S.A. (1988). Rhetorical structure theory: Toward a functional theory of text organization. *Text*, 8(3), 243–281. https://doi.org/10.1515/text.1.1988.8.3.243

McNeill, D. (1992). *Hand and Mind: What Gestures reveal about th*ought. Chicago, University Press.

McNeill, D. (2003). Pointing and Morality in Chicago. In S. Kita (Ed.) *Pointing: Where Language, Culture and Cognition Meet* (pp. 1–16). Mahwah, NJ: Erlbaum.

McNeill, D. & Duncan, S. D. (2000). Growth points in thinking-for-speaking. In D. McNeill (Ed.), *Language and gesture* (pp.141–161). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511620850.010

Merlo, S., Mansur, L. L. (2004). Descriptive discourse: topic familiarity and disfluencies. *Journal of Communication Disorders*, 37(6), 489–503. https://doi.org/10.1016/j.jcomdis.2004.03.002

Mulvey, L. (2006). *Death 24 x a Second*. London, Reaktion Books.

Müller, C. (2005). *Metaphors, Dead and Alive, Sleeping and Walking. A Cognitive Approach to Metaphors in Language Use*. Berlin, Freie Universitat.

Müller, C. (2014). Gestural Modes of Representation as techniques of depiction. *Body – Language – Communication: An international Handbook on Multimodality in Human Interaction*. (Handbooks of Linguistics and Communication Science (38.2)) (pp.1687–1701). Publisher: De Gruyter Mouton. https://doi.org/10.1515/9783110302028.1687

Noys, B. (2014). Film-of-Life: Agamben's profanation of the Image. In H. Gustafsson and A. Grønstad (Eds.) *Cinema and Agamben: Ethics, Biopolitics and the Moving Image* (pp. 89–102). New York-London-New Delhi-Sydney, Bloomsbury. https://doi.org/10.5040/9781628929911.ch-004

Pease, A. (1981). *Body language: How to read others' thoughts by their gestures*. Sydney: Camel Publishing Company.

Pease, A. & Pease, B. (2017). *The Definitive Book of Body Language*. Harlequin.

Pincus, D. (2019). Clinical psychology at the crossroads: An introduction to the special issue on nonlinear dynamical systems. *Nonlinear Dynamical Psychological Life Science*, 23(1), 1–15.

Searle, J. & Vanderveken, D. (1985). *Foundations of Illocutionary Logic*. Cambridge, Cambridge University Press. https://doi.org/10.1007/1-4020-3167-X_5

Shpet, G. G. (1922) Theatre as art. *The Art of the Theatre*, 1. Reprinted in (2007) in T. G. Schedrina (Ed.) *Art as a Kind of Knowledge. Selected Papers on the Philosophy of Culture* (pp. 19–39). Moscow, Russian Political Encyclopedia.

Singer, M. (2013). Profiles of discourse recognition. *Discourse Processes*, 50 (6), 407–429. https://doi.org/10.1080/0163853X.2013.822297

Stepanov, Yu. S. (2004). *Constants. The Dictionary of Russian Culture*. 3rd ed. Moscow, Akademicheskii proekt, 2004.

Streeck, J. (2008). Depicting by gesture. Gesture, 8(3), 285–301. https://doi.org/10.1075/gest.8.3.02str

Thelen, E. & Smith, L. B. (1996). *A dynamic systems approach to the development of cognition and action*. Cambridge, MIT Press.

Tønnessen, S. & Forsgren F. (Eds.). (2020). *Multimodality and Aesthetics* (Routledge Studies in Multimodality). New York and London, Routledge.

Trifonas, P. P. (2015). Crafting the literature of semiotic possibility: From the metaphysical to the detective story in The Name of the Rose. In P. P. Trifonas *International Handbook of Semiotics* (pp. 239–264). Dordrecht, Springer. https://doi.org/10.1007/978-94-017-9404-6_9

Van Dijk, T. (1990). Issues in Functional Discourse Analysis. *Unity in Diversity* (pp. 27–46). The Netherlands, Foris Publications. https://doi.org/10.1515/9783110847420.27

Verhagen, A. (2007). Construal and perspectivization. In: D. Geeraerts, H. Cuyckens (Eds.) *The Oxford handbook of cognitive linguistics* (pp. 48–81). Oxford, Oxford University Press.

Von Stutterheim, C., & Klein, W. (1989). Referential movement in descriptive and narrative discourse. In R. Dietrich, & C. F. Graumann (Eds.), *Language processing in social context* (pp. 39–76). Amsterdam: Elsevier. https://doi.org/10.1016/B978-0-444-87144-2.50005-7

Wårwik, B. (2004). What is Foregrounded in Narratives? Hypothesis for the Cognitive Basis of Foregrounding. In T. Virtanen (Ed.) *Approaches to Cognition through Text and Discourse* (pp. 99–122), New York, Mouton de Gruyter.

Willard, C. A. (1989). *A theory of argumentation*. Tuscaloosa, AL: University of Alabama Press.

Zykova, I. (2020). Verbal sources of cinematic metaphors: From cinematic performativity to linguistic creativity. *Lege artis. Language yesterday, today, tomorrow*, V(1), 499–532.

Zykova, I., & Kiose, M. (2020). Linguistic creativity parametrization in contrasting discourse types: Cinematic discourse vs. Discourse of children's literature. *Issues in Cognitive Linguistics*, 2, 26–40. https://doi.org/10.20916/1812-3228-2020-2-26-40

Żywiczyński, P., Wacewicz, S. & Orzechowski, S. (2017). Adaptors and the turn-taking mechanism: The distribution of adaptors relative to turn borders in dyadic conversation. Interaction Studies. https://doi.org/10.1075/is.18.2.07zyw