

Evidential components in multimodal communication

Julia Nikolaeva^a, Evgeniya Budennaya^{b,c} and Alexandra Evdokimova^d

^a Lomonosov Moscow State University, GSP-1, Leninskie Gory, Moscow, 119991, Russia

^b National Research University Higher School of Economics, Myasnitskaya Ulitsa, 20, Moscow, 101000, Russia

^c Institute of Linguistics, Russian Academy of Sciences, Bolshoy Kislovsky Ln, 1 bld 1, Moscow, 125009, Russia

^d Russian State University for the Humanities, Miusskaya Ploshchad', 6, Moscow, 125993, Russia

Abstract

The paper explores head and hand movements as markers of direct and indirect evidentiality, along with some lexemes in Russian. While there are numerous examples of evidential markers in speech, especially in languages where the category is grammaticalized, much less is known about non-verbal evidential markers. We claim that although there are no systemic rules for coding evidentiality, some polysemic hand and head gestures, such as palm up open hand gesture and head turns to the source of information can be regarded as indirect evidentials in line with the lexeme *vidimo* ('apparently'). More interestingly, character viewpoint hand gestures and their combinations of *vidimo* with representative gestures can be considered as direct evidential means, despite the fact that Russian lacks obligatory evidential marking.

Keywords

evidentiality, hand gesture, head gesture, multimodal communication

1. Introduction. Studies on evidentiality in verbal and non-verbal communication

Evidentiality as a category marking the information source of a statement has received a lot of attention in linguistic studies (see Section 2). How evidentiality is encoded in language has been widely discussed cross-linguistically from both a grammatical (separate affix / part of the tense system / modal morpheme, see Figure 1) and semantic (direct / indirect access to the information) point of view. According to [1], indirect evidentiality is marked in more than 50% of languages (267 out of a sample of 418), sometimes combined with direct evidentials:

Proceedings of the Linguistic Forum 2020: Language and Artificial Intelligence, November 12-14, 2020, Moscow, Russia

EMAIL: julianikk@gmail.com; jane.sdrv@gmail.com; arochka@gmail.com

ORCID: 0000-0001-8753-5945; 0000-0002-1502-2750; 0000-0002-6557-2064



© 2020 Copyright for this paper by its authors.

Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)

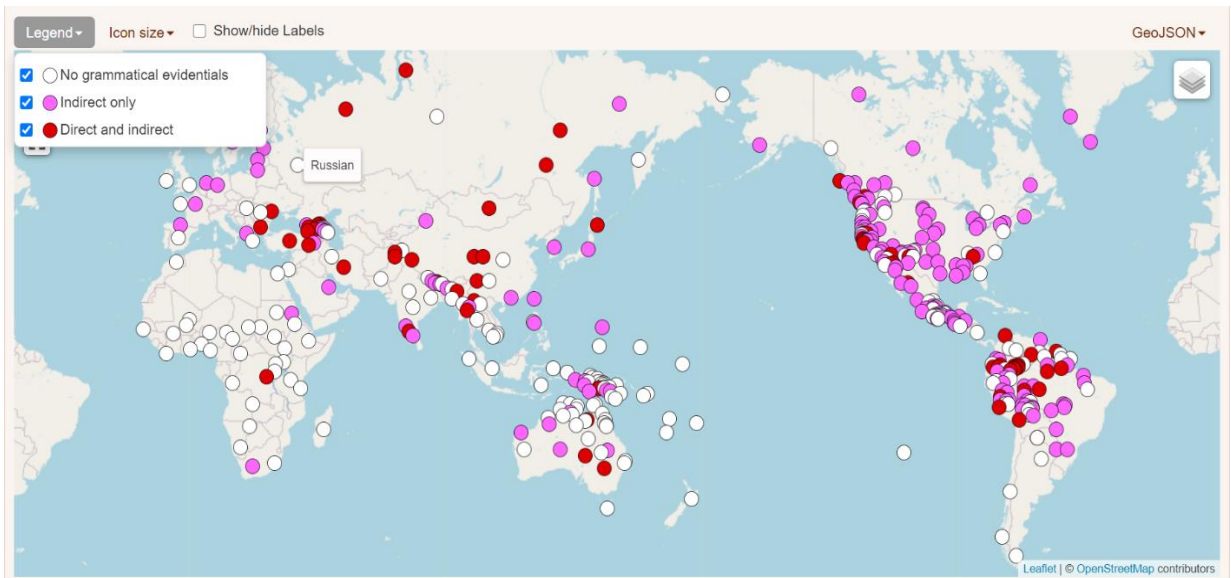


Figure 1: Coding of Evidentiality across the languages [1]

Semantically indirect evidentials indicate a situation the speaker has not witnessed himself but learned afterwards. They fall into two main subcategories, inferentives and reportatives. Consider Example (1) from Khalkha Mongolian [2: 129] which marks inferred information with a special evidential particle (INFER) and Example (2) from Estonian [3: 33] which has a special verbal form for reportatives (REP):

(1) Khalkha Mongolian:

ter irsen biz
 3SG come INFER
 ‘He must have come’

(2) Estonian:

ta olevat aus mees
 3SG be.REP.PRS honest man
 ‘He is said to be an honest man’

Russian is a language which lacks grammatical evidentiality (see Figure 1) and expresses the same information only through certain lexical markers, such as *vidimo* ‘apparently’, *kažets’a* ‘it seems that’, *okazyvaets’a* ‘it turns out that’, cf. Examples (3-4).

(3) Russian:

On vidimo doma [we see the lights in the window]
 3SG.M evidently at.home
 ‘Evidently, he is at home’

(4) *Kažets’a, on v komnat-e*
 It.seems 3SG.M in room-SG.LOC
 ‘It seems that he is in the room’

However, Paducheva [4] demonstrated that evidentiality in Russian is fused with some syntax structures, such as Negative Existential Sentences (NES). In this case, evidential meaning is expressed with the so-called “Genitive of Negation”, see Examples (5a) with a Nominative subject and (5b) with a Genitive. In (5a) Nominative marks simple absence the bottle, while in (5b) Genitive marks the presence of the observer who witnessed the situation. Subsequently, Genitive of Negation is fused with a direct evidential.

(5)

a. Butylk-a ne by-l-a v xolodil’nik-e
 bottle-NOM not be.PST-3.SG.F in fridge-SG.LOC

<i>b. Butylk-i</i>	<i>ne</i>	<i>by-l-o</i>	<i>v</i>	<i>xolodil'nike</i>
bottle-GEN	not	be.PST-3.SG.N	in	fridge-SG.LOC

‘The bottle was not in the fridge’

It should be noted that such an opposition is allowed for a limited list of verbs and only for referential and quantified subjects. However, given that previous studies on evidentials primarily dealt with regular grammatical markers, this finding broadens the scope of potential evidential components for languages which lack regular grammatical evidentiality. The development of multimodal corpora contributed to exploring how the message is conveyed through the channels other than verbal [5, 6]. Recent studies demonstrate that kinetic behavior conveys a range of language aspects previously discussed only through the verbal channel (the referent’s animacy and protagonism; foreground and background information, and so forth [7, 8, 9]). In this regard, non-verbal communication is a huge field for exploring potential evidential components.

This article addresses the evidential component on a Russian discourse level within the framework of multichannel communication, with a special focus on the interaction between verbal channel, kinetic behavior (hand and head gestures), and speaker’s role (see Section 3 for detail). The paper is structured as follows. Section 2 discusses previous research on the topic. Section 3 is devoted to the data and methods, that is, the RUPEX multichannel corpus (“Russian Pear Chats and Stories”) which has served as the source of the analyzed units of communication, and general principles of manual and cephalic annotation in RUPEX. Section 4 shows the interaction of the three channels (verbal, cephalic and manual) in the context of evidentiality and demonstrates how evidential components are associated with the speaker’s role. The following Section 5 summarizes the results and the final Section 6 concludes the work.

2. Related works

As mentioned above, evidentiality has been primarily regarded in the verbal channel. Most studies are focused on languages where evidentiality is a grammatical category (see [1, 3, 10, 11, 12, 13, 14, 15 *inter alia*]). Furthermore, the past decades witnessed a growing interest in how access to the information is marked in languages where evidentiality is not a grammatical category. In this regard, lexical markers are considered first, as in Russian (see Section 1 and [16; 17, 18] for detail).

However, much less is known about how the direct and / or indirect access to the information is marked on a discourse level, especially in languages of no grammatical evidentiality. Existing studies are first and foremost based on languages where evidentiality is a part of the grammar and has to be expressed, e.g. Bulgarian (East Slavic) [19] or Nganasan (Samoedic) [20]). In this way, prosody and the kinetic behavior of the speaker open a new field of research. Compared with verbal data, evidential components in non-verbal channels of communication have been studied very restrictedly and only for a limited number of languages (see, for example, [21] on hand gestures in Spanish communication). For Russian, no studies on non-verbal evidentiality in manual channel have been done thus far.

As for head movements, they have been much less investigated than hand gestures. In most cases, existing studies are focused on types of gesture and the role of head turns and nods in speaker’s and listener’s kinetic behavior [22, 23, 24, 25, 26, 27]. Here, head gestures are explored for evidential components, thus broadening the issue to the cephalic channel of non-verbal communication.

3. Data and methods

The study is based on the RUPEX multichannel corpus (“Russian Pear Chats and Stories”, see [28] and the website of the project <https://multidiscourse.ru/corpus/?en=1> for more detail). It consists of 40 sessions (communication episodes) in groups of four participants discussing “The Pear film” [29]. Each participant has a fixed role: the Narrator (N), the Commentator (C), the Reteller (R) and the Listener (L). At the preliminary stage, the Narrator and the Commentator watch the film. Then the Narrator describes the content of the film to the Reteller who has not seen it before (“First telling”). No interruption is allowed until the Narrator completes his/her story. This is followed by a conversation stage (“Conversation”), where the Commentator adds details to the Narrator’s monologue and the Reteller asks questions to both interlocutors to better understand the story. Finally (the “Retelling” stage), the Reteller

recounts the film to the Listener, who joins the rest of the participants only at this stage. The general design of the session is presented in Figure 2.

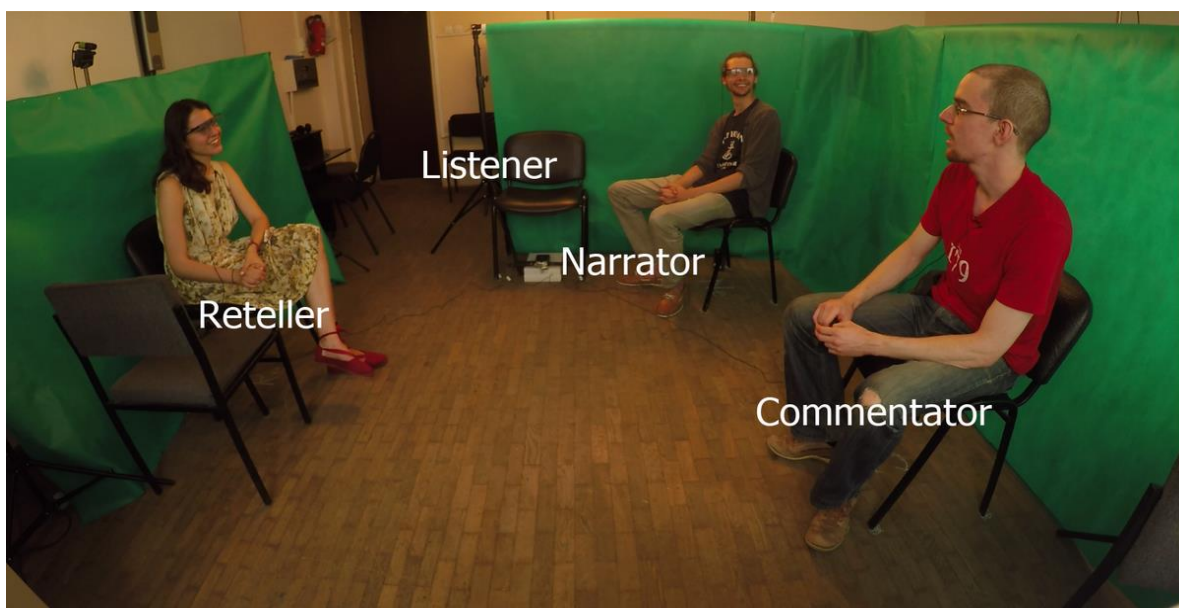


Figure 2: General design of a communication episode

The main mark-up of the RUPEX corpus is carried out with ELAN software (<https://tla.mpi.nl/tools/tla-tools/elan>) and includes vocal, oculomotor, manual and cephalic annotation. This makes it one of the largest resources in terms of annotated communication channels [6]. For the current study of evidential components, we analyzed sessions #04, #22 and #23 (the demo subcorpus), which lasted in total more than an hour.

Our approach was to compare verbal, manual and cephalic patterns of the Narrator and the Reteller. As mentioned above, the Narrator watched the film and thus had direct visual evidence for the events he/she described, while the Reteller did not see the film and based his/her story only on indirect evidence. Given the cross-linguistic prevalence of indirect evidentials over direct markers [1, see also Section 1], we assumed that at least indirect evidentiality might be somehow reflected through the verbal and kinetic behavior of the Reteller. In this study, we discuss hand and head gestures which accompany verbal evidentials, such as *vidimo*, and how gesture type and viewpoint (see Sections 3.1.2 and 3.2.2) are related to the speaker's role.

When analyzing, we relied on RUPEX vocal annotation [30; 31], together with manual and cephalic. The relevant parameters of the latter two taken into consideration are presented below.

1.1. Manual channel

1.1.1. Basic principles of annotation

During the study, we examined depictive and pointing gestures performed from the character's and observer's point of view (OVPT or CVPT, respectively [7]) as a potential evidential component. Manual pragmatic markers were considered as a means to express the speaker's stance, namely (un)certainly. Finally, we compared which gestures accompanied lexical evidentials such as *vidimo* (see Section 4.1), for speakers who have witnessed the event themselves and those who have only heard about it.

The annotation of manual gestures was based on the principles described in [32]. For this study we distinguished between four functional types of gestures regarding how they relate to speech along with their formal features. First, pointing gestures (also known as deictics) have a clearly recognizable form and relate to a referent or a place and associate them with a place in the gesture space around the speaker. Second, depictive gestures can refer to a referent too, but they have a more complex hand form and/or trajectory and convey additional spatial dynamic information such as the position of referents in space, path, speed, trajectory and direction of movement or form and size of an object. Pragmatic gestures correspond to a pre-existing repertoire (see also [33]) and convey not intra-, but meta-discourse meanings

such as discourse structure and the speaker's stance. The most often used pragmatic gestures in the RUPEX narrations are conduit metaphors (palm up open hand gesture, or PUOH), reflecting the transfer of the message from the speaker to the listener, and a swaying gesture used to mark uncertainty. Finally, beats are simple two-fold movements, usually up and down. They correlate to the speech prosody and are considered to highlight the concurrent words [34, 35]. There is some evidence that they perform pragmatic functions too, such as marking boundaries between larger parts of discourse or label erroneous or attracting listeners' attention [7, 36, 37, 38, 39, 40].

1.1.2. Annotating viewpoint in manual gestures

As other channels, manual gestures were annotated for viewpoint with Elan software, see Figure 3.

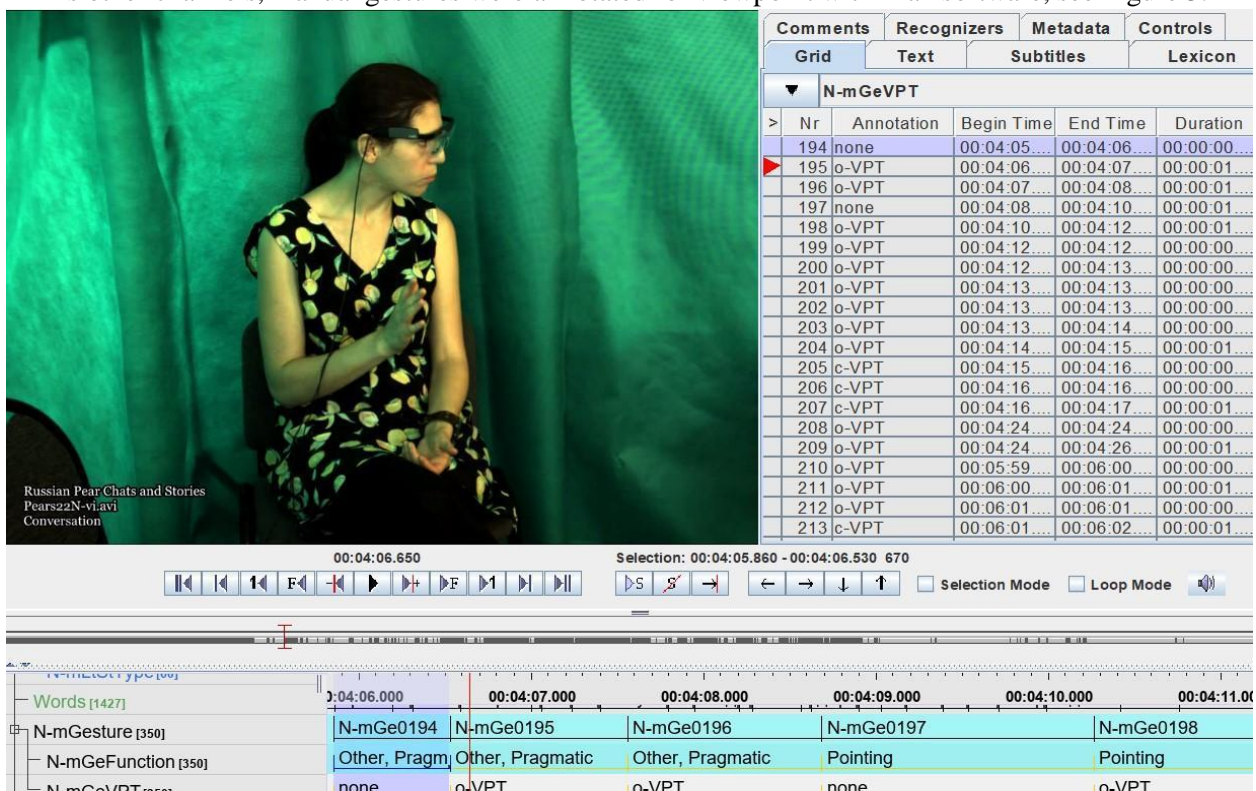


Figure 3: Annotation of gesture types and viewpoint in Elan

To compare the verbal and manual channels, we used a textgrid created by Korotaev, Kibrik and Podlesskaya [41]. Then in Excel the gesture viewpoint was marked (see Figure 4 where orange stands for OVPT, green for CVPT, grey for pragmatic gestures, and chains of sequential coreferential gestures are shown with brackets at the right side of the transcript).

In contrast to pragmatic gestures, depictives and pointings (with rare exceptions) refer to the story itself. As noted above, depictive gestures can be either CVPT or OVPT. Our research shows that the choice of viewpoint in gestures also depends on how the speaker learned about the story they retell. Depictive gestures can show the shape or size of the referent, the position of referents in space, speed, trajectory and direction of movement. When a sequence of depictive gestures presents the same referents maintaining the form and place in gesture space, we can see chains of coreferential gestures.

226.18	227.19				C-vE005	на котором растёт / / дерево.
227.19	227.59				C-vN003	(ч 0.40)
227.59	229.05				C-vE006	к дереву приставлена / / лестница.
229.05	229.12					
229.12	230.36				C-vE007	по этой / лестнице –пазит / фермер.
230.36	230.65					
230.65	232.02				C-vE008	То есть он (с 0.14) (0.08) взлезает /наверх.
232.02	233.34				C-vE009	собирает (0.18) (а 0.63) груши себе в (с 0.36) /фартик.
233.34	235.67				C-vN004	(ч 0.33)
235.67	236.39				C-vE010	но /спускается.
236.39	237.94				C-vE011	и очень /бережно их кладёт в /корзинки.
237.94					C-vE012	У него стоит три /корзины,
238.17	238.80					
238.85	238.85					
238.85	239.26				C-vE013	/больших.)
239.26	239.80				C-vN005	(ч 0.54)
239.80	240.84				C-vE014	две уже-е почти (с 0.11) ==
240.84	241.66				C-vE015	–одна (0.08) /полная,
241.66	241.81					
241.81	242.72				C-vE016	–одна /–почти полная.
242.72	243.87				C-vE017	и в // /третья –пустая.
243.87	244.53				C-vN006	(ч 0.66)
244.53	244.95				C-vE018	–Вот.
244.95	245.08					
245.08	245.55				C-vF001	(еще 0.47) ≈
245.55	245.89					
245.89	246.31	N-vE151	A (0.13) не= ==			
246.31	247.00	N-vE152	/Можно /вопрос?			
247.00	247.80	N-vE153	A –как так получилось.			
247.80	248.90	N-vE154	что когда он /слез с дерева.			
248.90	250.42	N-vE155	(Я ещё /сама /забыла этот вопрос задать.)			
250.42	250.76	N-vN045	(ч 0.34)			
250.76	252.75	N-vE156	если /один-н // (ч 0.26) (с 0.11) /одну корзину –			
252.75	253.64	N-vE157	/полностью /полную.)			
253.64	254.08	N-vE158	– /увезли.			
254.08	255.76	N-vE159	почему /другая полностью /пустая?			
255.76		N-vE160	Она /вообще /пустая совсем.			
256.99						
	257.02				C-vE019	Ну у него было /три корзины /изначально.
	258.80					
258.80	259.10				C-vE020	–Одна-а (0.11) /Тполная.
259.10	260.08					
260.08	260.19				C-vE021	–одна-а (0.10) она (0.09) /дополнил как раз.
260.19	261.74					
261.74	261.82					
261.82	261.88				C-vE022	к= ==
261.88	262.02				C-vE023	ну.
262.02	262.10					
262.10	263.13				C-vE024	/первый раз когда он /слезает.
263.13	263.43				C-vN007	(ч 0.30)
263.43	264.66				C-vE025	он /докладывает туда до /верха.
264.66					C-vE026	и у него остаётся /третья /–пустая /корзина.
264.71	266.17	N-vE161	И после /этого мальчик /привезает?;			
	266.39					
266.39	266.86				C-vN008	(ч 0.47)
266.86	267.00				C-vE027	–Да.
267.00	267.85				C-vE028	после /этого привезает /мальчик.

Figure 4: Mark-up for viewpoint and gesture chains in manual gestures

1.2. Cephalic channel

1.2.1. Basic principles of annotation

The cephalic annotation for this research was realized on Session #22 using the methods discussed in [42]. Functional types of cephalic gestures are similar to those in the manual channel with some additions. We studied monologue parts, when hands do not play an important role in turn-taking. Since hand gesturing is reserved for the speaker, the listener is not supposed to use her/his hands for communicating any information. By contrast, head movements do not stop while listening, thus making head turns a means to show who the gesturer is listening to or watching. The speaker uses head turns to show the actual addressee, to draw someone’s attention or to choose the next speaker. So, one of the added types is regulators – turns and leans at a higher amplitude showing the targeted interlocutor. Apart from pragmatic gestures conveying meta-discourse information, there are head gestures combining pragmatic and pointing functions called “pragmatic-center” and “pragmatic-away” which highlight the synchronous hand gestures [43]. In addition to that, some minor posture changes (called “accommodators”) were found to serve as a cohesive means in discourse, so they were considered to be another gesture type in the cephalic channel.

In the analysis we discuss cephalic behavior which accompanies verbal evidentials, such as *vidimo*, and how cephalic viewpoint is related to the speaker’s role. Since viewpoint has never been investigated on head gestures before, we have developed a special algorithm based on the communication zone [44: 129, see also Section 3.2.2 and Figure 5] and its shift, as well as on type of gesture and its interaction with

other kinetic channels. With all these factors taken into account, each gesture was tagged as CVPT/OVPT.

The main principles of the algorithm are discussed below.

1.2.2. Annotating viewpoint in cephalic gestures

The position of each participant of the session in relation to his/her interlocutors plays a special role in communication. Usually, participants do not look at one another in the same way throughout the conversation, nor do they keep their positions unchanged. The conditions of the situation make them change the communication zone through head and body turns. These turns perform a regulator function and thus can be tagged as OVPT gestures.



Figure 5: Communication zones and their shift during communication: R shares hers with N and C, while N and C share another

In addition to communication zones, we take into consideration the gesture type and visibility range. All depictive gestures used for an illustration of the character being described were marked as CVPT. For other types of gestures and their chains, we examined the position of the head and the visibility range. If the latter was changed, we marked the gesture as OVPT. If not, we consider concomitant gestures of other kinetic channels. If the head gesture in question corresponded in its direction, structure or rhythmical organization with a concomitant gesture in the other kinetic channel [45] of the same meaning, we tagged it with the same viewpoint as for a gesture in the other kinetic channel. Otherwise, we marked it as an OVPT since the observer's viewpoint is more typical for head gestures.

For example, the Narrator's fragment *devočka takaja s dlinnymi černymi kosami* ('such a girl with long black braids', see Figure 6) is accompanied by a manual depictive CVPT gesture. At the beginning of the word *devočka*, 'girl' and before the onset of the manual gesture, we have a cephalic pointing gesture chain directed down towards the hands. Its function is to attract attention to the arms. While the Narrator gesticulates manually, her corpus and head go backward, and her visibility range becomes larger. According to our algorithm, this entire chain of cephalic gestures has to be marked as OVPT.



Figure 6: A chain of OVPT head gestures accompanied with manual gesticulation

4. Results

1.3. Verbal channel

To explore how (in)direct evidence of the speaker's statements was reflected in verbal channel, we looked at the distribution of lexical markers with evidential meaning in the monologue of the Narrator and compared them with the monologue of the Reteller for the same sessions (#04, #22, #23). All the monologues lasted about half an hour in total.

The analysis revealed a strong tendency towards the use of the lexical inferentive *vidimo*: compared to other possible lexical markers of indirect evidence (*vidno*, *kažets'a*, *okazyvaets'a*, *kak budto*, etc.), only *vidimo* was found in the speakers' monologues. As presupposed, the Retellers used it significantly more often than the Narrators, see Table 1 below.

Table 1

Distribution of lexical inferential evidential *vidimo* in the monologues of N and R (binomial test, p-value=0.04)

Session	N	R
#04	2	1
#22	2	12
#23	0	2
Total	4	15

The prevalence of *vidimo* in the Reteller's monologue can be attributed to his/her ongoing processing of the information received from the Narrator. Since the Reteller has not seen the film, it takes him/her more time to understand the sequence of the story and to establish logical connections between the episodes. This results in the overuse of *vidimo* which indicates the inference the Reteller made based on previously mentioned information. Consider the same episode as presented by the Narrator (6) and by the Reteller (7), with *vidimo* marked with bold:

(6) Pears #22N:

N-vE084 ¹	on \↑oboračivaets'a,	'He turns around
N-vE085	(na \neĕ,)	at her
N-vE086	i u /nego padaet \šl'apa.	and his hat falls down
N-vN026	(ŭ 0.47)	
N-vE087	Iz-za togo čto on /oboračivaets'a,	since he turns around
N-vE088	on /-naezžaet peredn'im /kolesom na bolšoj \kamen'.	he get the front wheel onto a big stone.
N-vN027	(ŭ 0.23)	
N-vE089	Dejstvitel'no dovol'no /bol'šoj,	Really quite a big one,

¹ Here and throughout the transcription is given in accordance with principles discussed in [Kibrik, Korotaev, Podlesskaya 2020].

N-vE090	to est' on-n (0.50) {sw 0.22} (0.13) (u 0.06) \objomnyj.	that is, it is huge.
N-vN028	(u 0.25)	
N-vE091	\V _o -ot.	Well.
N-vE092	On padaet s /velosipeda	He falls off the bicycle'

(7) Pears #22R:

R-vE155	— (ə 0.24) šl'apa \sletaet s /mal'čika,	'the hat falls from the boy,
R-vE156	i ostačts'a gde-to na \doroge.	and remains lying on the road.
R-vN040	(u 0.41)	
R-vE157	\V _{ot} .	Well.
R-vE158	I-i vidimo mal'čik kak-to otvl'oksja na etu	And apparently , the boy was distracted
R-vE159	/devočku,	by this girl,
R-vE160	(kotoraja \mimo nego proexala,)	that passed him,
R-vE161	^i-i (? 0.23) naexal na /kamen',	and rode up onto a stone,
R-vE162	(dostatočno \bol'soj vidimo ,)	apparently , a big one,
R-vN041	(u 0.34)	
R-vE162	'i-i ^i (0.26) \upał s /velosipeda	and fell off the bicycle'

Since certain aspects could be learned only through visual evidence, the Reteller apparently has not learned all the information to the same extent as the Narrator, due to his/her indirect access to the events. Subsequently, the Reteller seeks to infer the missing information based on what has been said before. The Narrator does not make any extra inference during the report, as direct visual evidence allows him/her to completely understand the story. The distribution of *vidimo* provides an indication of the speaker's access to the events described and to his role in the session.

4.2. Manual channel

4.2.1. Lexical evidential markers and their accompaniment in the manual channel

Similarly to discourse markers which are not directly related to the story, but reveal its structure and the speaker's stance, pragmatic gestures do not illustrate events from the film, but operate with the narration as an object, pointing at its parts and the connections between them, metaphorically handing the story to the listener, or discarding insignificant details.

Since gestures and speech transfer the same message, among the pragmatic gestures there are those that show the speaker's (lack of) confidence in her words. A typical example of a pragmatic gesture is the "conduit metaphor" [7] or palm up open hand (POUH) gesture [46], see Figure 7.



Figure 7: Emphasized Reteller’s PUOH gesture

With the lexeme *vidimo* the Narrators used either depictive or pragmatic gestures. Depictive gestures transfer additional information and can sometimes complement what is said in words. Our data show that even when verbally declaring uncertainty, speakers can convey gesturally the information they imply. On the contrary, the Retellers who used *vidimo* in abundance used no gesture at all in about half of the cases, the other times they used less informative pointing or pragmatic gestures, sometimes with a delay 250 up to 1150 ms after the word. Often, they accompanied *vidimo* with the pragmatic PUOH gesture (see Table 2). It shows that the Retellers did not have available information to be transferred through any of the channels and sometimes just attenuated the lack of evidence by simply drawing the listener’s attention to the characters without specifying the details of the events.

Table 2

Lexeme *vidimo* with manual gestures

22N, <i>vidimo</i>		22R, <i>vidimo</i>	
PUOH	1	Pointing gesture	1
		No gesture	6
		Delayed PUOH	3
Beat gesture	1	PUOH	2
04N, <i>vidimo</i>		04R, <i>vidimo</i>	
PUOH	2	No gesture	1
Depictive gesture	1		
23N, <i>vidimo</i>		23R, <i>vidimo</i>	
		No gesture	1
N/A		Depictive gesture	1

In addition, eyewitnesses speak of events directly, while people who receive information from others mention it in their words or gestures.

4.2.2. Depictive gestures as direct evidentials in contrast to pragmatics

The Narrators saw the film and have some visual information which can be transferred through hand gestures. Often their gestures did not simply duplicate what was said in words but add some important details. The Retellers have to infer the details from both what they see and hear and yet their mental representation about the story is incomplete. Thus, they do not have much to depict with their gestures, so the Retelles’ gestures are fewer in number and less informative. At the same time, the Retellers use more pragmatic gestures, especially PUOH (see Table 3, Figure 8 and Table 4). It can be explained by the need to keep the listeners’ attention, so the appellative nature of this gesture gives the straightforward explanation. There may also be other reasons for higher number of PUOH gestures in the Retellers, such

as invitation for listeners to imagine other details of the story themselves or a means to cover embarrassment or limited knowledge of the film.

Table 3

The Narrators preferred depictive hand gestures, while the Retellers used more pragmatic hand gestures (χ -square, p -value<0.001)

	Narrator		Reteller	
Depictive	376	50%	168	39%
Pragmatic	216	28%	192	45%
Other	166	22%	69	16%
Total	758	100%	429	100%

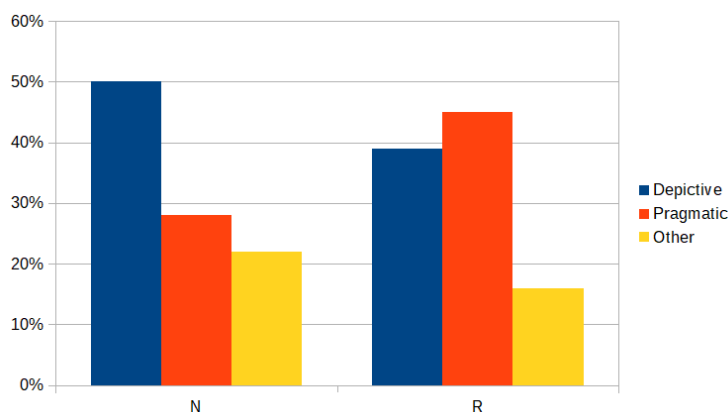


Figure 8: The Narrators preferred depictive hand gestures, while the Retellers used more pragmatic hand gestures

Table 4

PUOH and other pragmatic gestures in the Narrators and the Retellers speech

	Narrator		Reteller	
PUOH	68	31%	92	48%
Other	100	46%	100	52%
Total	216	100%	192	100%

4.2.3. Speaker's viewpoint in manual gestures

Viewpoint was annotated only for pointing and depictive manual gestures. Thus, CVPT is observed when the speaker moves as a character of the story. For example, saying *And he picks up the basket*, the speaker pulls her arms forward and pretends to be holding a basket. For OVPT she gestures as a bystander who is watching the events and is not taking part. Thus, pointing in front of her while saying *The boy was riding along the road* is an OVPT gesture.

The speaker's choice between the two viewpoints is affected by various factors, including syntax (the number of actants in a verb: transitive verbs are more likely to be illustrated by the participant's gestures than intransitive ones), and discourse structure. The study revealed that the speaker's experience of witnessing the events under discussion (direct or indirect) affects his/her gesticulation, even if there is no grammatical evidentiality in the language they speak.

4.2.4. CVPT gestures as direct evidential markers

Typically, depictive gestures illustrate the most dynamic or important events of the plot [8]. "The Pear Story" film includes movements and actions of people with objects, which boosts the use of depictive gestures. The Reteller was told to remember all the details and be ready to describe the story the listener had not seen as best as possible. Obviously, the Narrator and the Commentator who had seen the film used a lot of depictive gestures to illustrate their words. Interestingly, in the Reteller's speech, there were

significantly fewer CVPT gestures out of depictives and pointings (see Table 5 for an illustration of the distribution of CVPT gestures between the Narrator and the Reteller in Session 04, χ -square, p -value = 0.003).

Table 5

CVPT out of all depictive and pointing hand gestures (Session 04)

	Narrator		Reteller	
CVPT	82	50%	74	35%
Others	82	50%	140	65%
Total	164	100%	214	100%

Gestures in Figure 9a and 9b show similar gestures in similar contexts but with an important difference. The speaker in Figure 9a rubs the pear holding it in her hands, while the speaker in 9b talks about the pears and does not pretend to be the person who gathers them

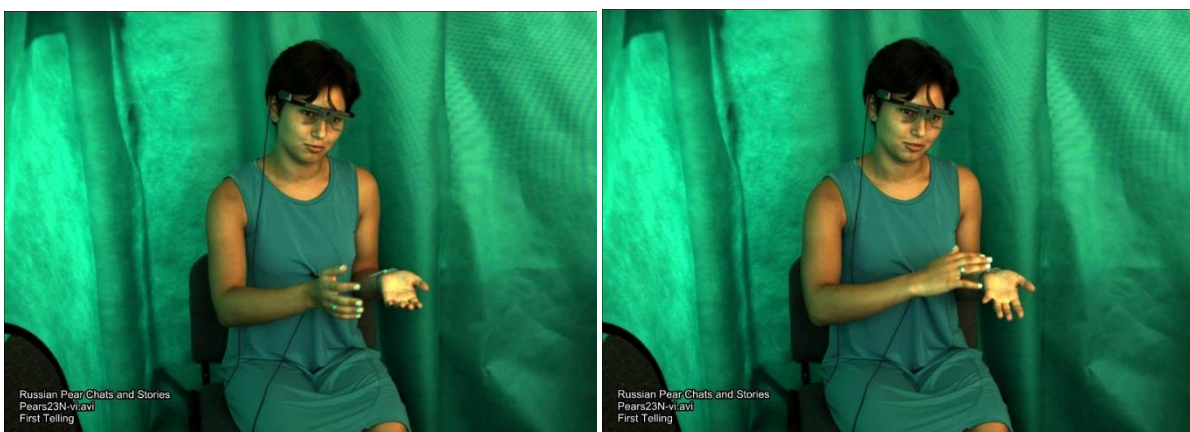


Figure 9a: The Narrator's CVPT gestures on the phrase *Nu pomogayut, sobirayut grushi* 'Well, they help him to gather the pears'

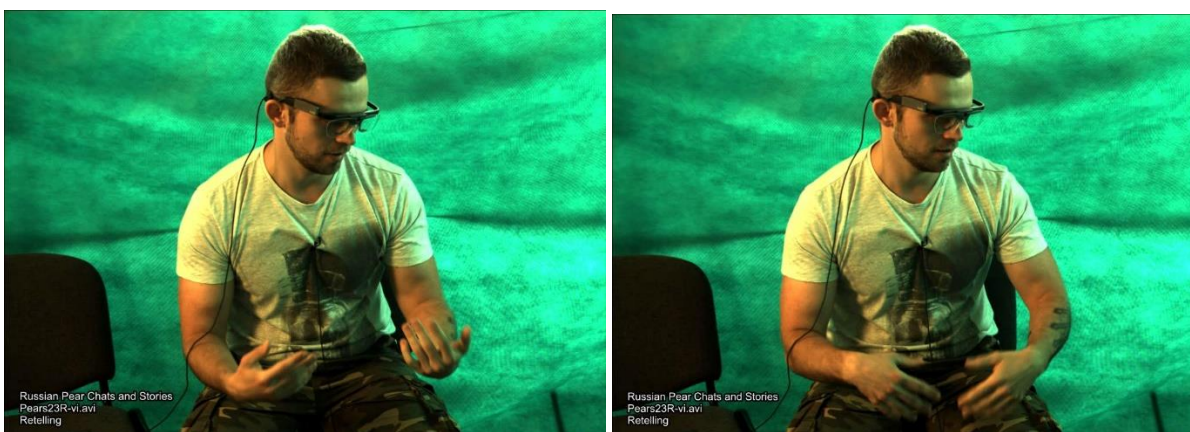


Figure 9b: The Reteller's OVPT gestures on the phrase *Pomogayut emu podniatsa, vot, sobrat' eti grushi* 'They help him to get up, well, to gather these pears'

4.3. Cephalic channel

4.3.1. Distribution of CVPT and OVPT according to the speaker's role

As for the manual channel, we examined how the speaker's cephalic behavior was related to the concomitant lexical evidential *vidimo*.

In Session #22, the Narrator says the word *vidimo* only twice, accompanying it with a depictive gesture with a meaning of doubt (*Rotation*) and a pragmatic gesture (*Shake*), which also expresses doubt.

Cephalic behavior here is similar to hand gestures, since the Narrators accompany their rare verbal evidentials with both depictive and pragmatic gestures.

In contrast, the Reteller uses the lexeme *vidimo* more often than the Narrator (see Section 3) and always accompanies it with pragmatic, pointing and regulator gestures. Absence of depictive head gestures is similar to manual channel and shows their lack of concrete information which could be otherwise shown in gestures. The difference in the Narrator's and the Reteller's behavior could be attributed to their different degree of involvement into the story, which stems from their witnessed and unwitnessed account, respectively.

4.3.2. Distribution of CVPT and OVPT according to the speaker's role

After the annotation, we analyzed how different types of gestures and their viewpoint were distributed between the Narrator and the Reteller.

Head gestures show a great potential for combination of a few functions in one gesture. The preliminary analysis reveals that the Narrator demonstrates a far more varied range of CVPT gestures compared to the Narrator and Reteller's difference in diversity in OVPT gestures, see Table 6 below.

Table 6
Number of sub-types of CVPT and OVPT gestures in the Narrator and the Reteller

	Narrator	Total	Reteller	Total	χ -square, p-value
CVPT	depictive depictive-pragmatic pointing pointing/deictive pointing/pragmatic accomodator accomodator/pragmatic- away pragmatic-away pragmatic-center pragmatic	14	depictive pointing regulator	3	0,027
OVPT	depictive pointing pointing/regulator pointing/deictive pointing/pragmatic pointing/pragmatic- center regulator regulator/pragmatic- center regulator/pragmatic accomodator accomodator/regulator accomodator/pragmatic pragmatic/away pragmatic-center pragmatic- center/pragmatic-away pragmatic	16	depictive depictive/pragmatic pointing pointing/regulator pointing/deictive pointing/pragmatic pointing/pragmatic- center regulator regulator/deictive regulator/pragmatic accomodator pragmatic-away pragmatic-center pragmatic	10	

The distribution of cephalic CVPT-gestures between the Narrator and the Reteller has shown the same tendency as previously found in the manual channel. Thus, the Narrator used CVPT-gestures significantly more often than the Reteller (37 from 234 head gestures in #22N vs. 4 from 179 in #22R, χ -square, p -value<0.001), see Figure 10.

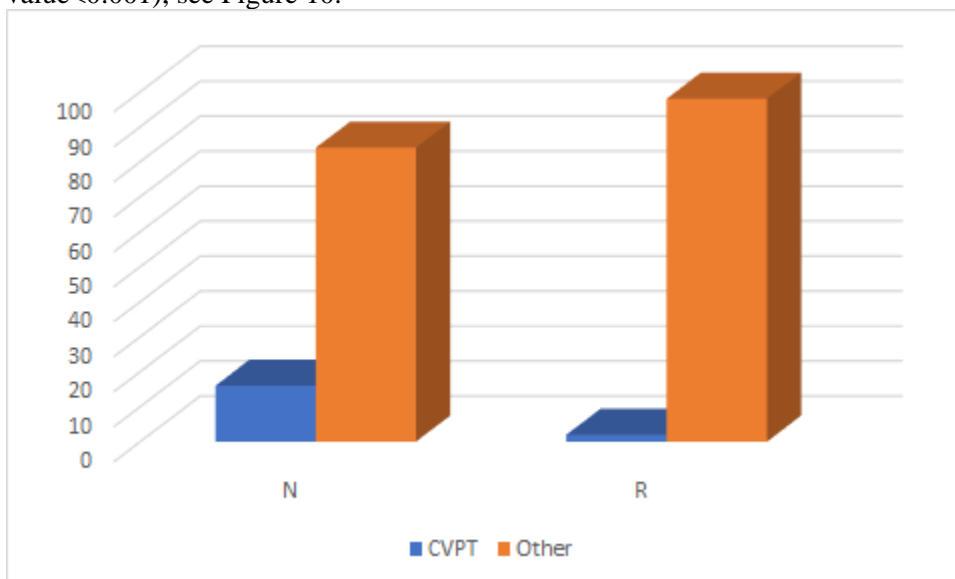


Figure 10: CVPT gestures in the Narrator's and the Reteller's speech, %

Similarly, the Narrator used far more depictive gestures compared to the Reteller (45 from 391 head gestures and gesture movements in #22N vs. 10 from 484 in #22R, χ -square, p -value<0.001), see Figure 11).

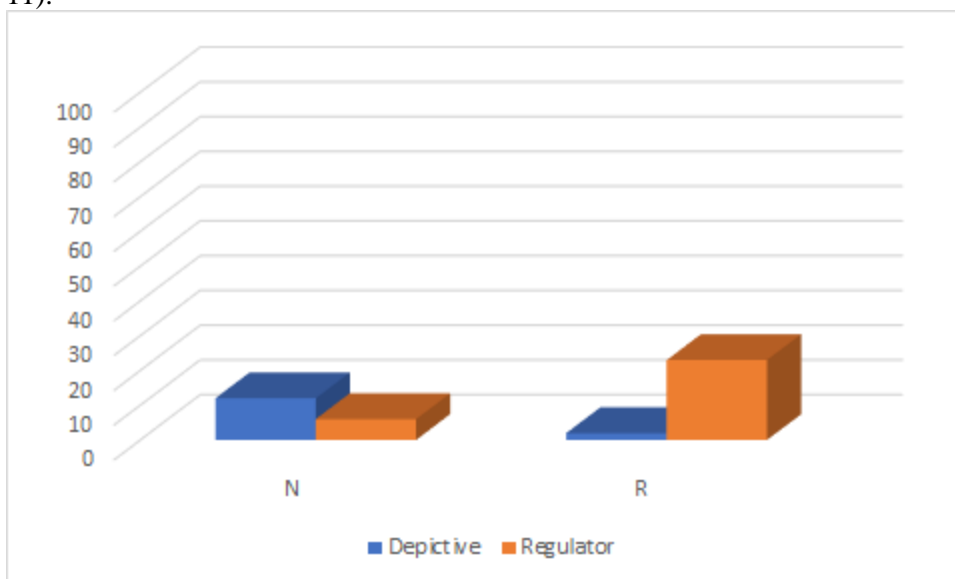


Figure 11: Depictives and regulators in the Narrator's and the Reteller's speech, %

This can be due to the fact that the Narrator has seen the film and got more engaged into the story she was telling. In contrast, the Reteller has not seen the film and thus only put together the non-witnessed facts given by other participants of the conversation.

Finally, the Reteller demonstrates a larger number of regulator gestures (112 from 484 head gestures and gesture movements in #22R, 23 from 391 in the 22N, χ -square, p -value<0.001, see Figure 11). This can be related to the position of the Reteller in relation to other participants and her role in communication. Namely, when the Listener enters, the Reteller changes the position of her corpus previously directed towards the Narrator and the Commentator and subsequently changes the communication zone. However, during her retelling that follows, she still needs to check her story, turning to the Narrator and to the Commentator to look at their non-verbal reaction. The Reteller's cephalic behavior can thus be attributed to her indirect access to the film and considered as a non-verbal marker of indirect evidentiality.

5. Discussion

Though Russian lacks obligatory evidential marking, there are some processes on the discourse level that can point at the speaker's source of information, namely direct or indirect witnessing. As a verbal marker, in our data it turned out to be the lexeme *vidimo*. In manual and cephalic gestures there are some specific elements too which correlate with the speaker's experience. Manual PUOH gestures prevail in the Retellers, who have not seen the story, while the Narrators use more depictive gestures related to the description of spatial and dynamic details of the events in both kinetic channels which can be regarded as a direct evidential marker. Conversely, less number of depictive gestures with increase in pragmatics, especially PUOH hand gestures and cephalic regulators can be interpreted as an indirect evidential. This difference can be explained by lack of the Retellers' visual experience and thus lack of spatial information to be expressed through words or gestures.

The Retellers mask or compensate for this lack with pragmatics and regulators. Interestingly, pointings can both serve as a direct evidential (they prevail in the Narrators' monologues), and as an accompaniment for verbal evidential markers, in this way performing the meta-discourse function of regulating the interlocutors' coordination, highlighting the referents or structuring the discourse. In addition, the speaker's head turns to the source of information about the story that can supposedly be considered an indirect evidential.

Finally, kinetic accompaniment of verbal evidentials differ between the speakers. Depictive gestures in such contexts point to the direct experience and availability of visual information about the events, while pragmatic and pointing gestures, along with their delay or absence in the manual channel, signal absence of such information.

These are pilot results or a study dealing with discourse evidentiality in the verbal, manual and cephalic channels in Russian. Limitations of our findings may be related to the data (monologues and retellings of a short film) and are to be tested on other types of discourse and a larger number of participants.

6. Conclusion

Evidentiality is not obligatory in Russian, but there are oblique means for the speakers to express if they have witnessed certain events personally. These means include lexical markers, manual and cephalic gestures and their combinations. Some of them refer to direct evidentiality, such as more often used CVPT head and hand gestures, some are observed when the speaker relies on other people's testimony, these are repeated lexemes *vidimo* ("apparently"), PUOH gestures and regulating head turns. In addition, combinations of *vidimo* with depictive gestures can be regarded as direct evidentials, while lack of a gesture, delayed or synchronous PUOH gesture with *vidimo* is pertinent as an indirect evidential.

7. References

- [1] F. de Haan, Semantic Distinctions of Evidentiality, in: M. S. Dryer, M. Haspelmath (Eds.), The World Atlas of Language Structures Online, Max Planck Institute for Evolutionary Anthropology, Leipzig, 2013. URL: <http://wals.info/chapter/77>
- [2] J. C. Street, Khalkha Structure, Indiana University Publications, volume 24 of Uralic and Altaic Series, Indiana University Press, Bloomington, 2013
- [3] A. Y. Aikhenvald, Evidentiality, Oxford University Press, Oxford, UK, 2004.
- [4] E. V. Paducheva, Est' li v russkom iazyke grammaticheski vyrazhennaia evidentsial'nost'? [Is there grammatically expressed evidentiality in Russian?], Russkij yazyk v nauchnom osveshchenij 2.26 (2013), 9–29. (In Russian)
- [5] G. Brone., B. Oben. What you see is what you do. On the relation between gaze and gesture in multimodal alignment, Language and Cognition 7.4 (2015): 485–498. doi: 10.1017/langcog.2015.22
- [6] A. A. Kibrik, Russkiy mul'tikanal'nyy diskurs. Chast' I. Postanovka problemy [Russian multichannel discourse. Part I. Setting up the problem], Psikhologicheskij zhurnal, 39.1 (2018a): 70–80. (In Russian)
- [7] D. McNeill, Hand and mind: What gestures reveal about thought, University of Chicago Press, Chicago, 1992.

- [8] Yu. V. Nikolaeva (2004), Funktsionalnye i semanticheskie osobennosti illyustrativnykh zhestov v ustnoy rechi (na materiale russkogo yazyka) [Functional and semantic aspects of illustrative gestures in Russian spoken discourse], *Voprosy Jazykoznanija* 4 (2004): 48–64. (In Russian)
- [9] E. V. Budennaya, Yu. V. Nikolaeva, A. A. Evdokimova, Referential phenomena in speaker's kinetic channels, in: *Computational Linguistics and Intellectual Technologies: Papers from the Annual International Conference "Dialogue"*, RGGU, Moscow, 2020, pp. 118–131.
- [10] W. Chafe, Nichols J. (Eds.), *Evidentiality: The Linguistic Coding of Epistemology*, Ablex (1986). Ablex, Norwood, NJ, 1986.
- [11] Z. Guentchéva (Ed.). *L'Énonciation médiatisée*. Bibliothèque de l'information grammaticale, Éditions Peeters, Louvain, 1996.
- [12] J. van der Auwera, V. Plungjan, Modality's semantic map, *Linguistic typology* 2.1 (1998): 79–124.
- [13] F. de Haan, The Cognitive Basis of Visual Evidentials, in: A. Cienki, B. J. Luka, M. B. Smith (Eds.), *Conceptual and Discourse Factors in Linguistic Structure*, CSLI Publications, Stanford, 2001, pp. 91–106.
- [14] F. de Haan, Coding of Evidentiality, in: M. S. Dryer, M. Haspelmath (Eds.), *The World Atlas of Language Structures Online*, Max Planck Institute for Evolutionary Anthropology, Leipzig, 2013. URL: <http://wals.info/chapter/78>
- [15] A. Y. Aikhenvald. *The Oxford Handbook in Evidentiality*, Oxford Handbooks in Linguistics, Oxford University Press, Oxford, UK, 2018. doi: 10.1093/oxfordhb/9780198759515.013.1.
- [16] T. V. Bulygina, A. D. Shmelev, Gipoteza kak myslitel'nyy i rechevoy akt, in: N. D. Arutiunova, N. K. Riabtseva (Eds.), *Logicheskii analiz yazyka. Mentalnyye deystviya*. Moscow, 1993, pp. 78–82.
- [17] V. S. Khrakovskiy. Evidencialnost, epistemicheskaya modalnost, (ad)mirativnost [Evidentiality, epistemic modality, (ad)mirativity], in: V. S. Khrakovskiy (Ed.). *Evidencialnost v yazykakh Evropy i Azii*. Sbornik statey pamyati N. A. Kozincevoy, Saint-Petersburg, 2007. S. 600–632. (In Russian)
- [18] A. Letučij, Sravnitel'nyye konstrukcii, irrealis i evidencialnost [Comparative constructions, irrealis and evidentiality], in: *Wiener Slawistischer Almanach, Sonderband 72*, Wien, 2008, pp. 215–238. (In Russian)
- [19] M. M. Makartsev, Evidentsial'nost' v prostranstve balkanskogo teksta [Evidentiality in the space of the Balkan text]. Nestor-Istoriya, Moscow, St. Petersburg, 2014. (In Russian)
- [20] A. Yu. Urmanchieva, Skhodstvo narrativnykh strategij nganasanskogo jazyka i srednetazovskogo govora sel'kupskogo (o vozmozhnoj korrel'atsii lingvisticheskikh i etnograficheskikh dannyyh [The similarity of narrative structures in Nganasan and Middle Taz Selkup (some possible correlations of linguistics and ethnography)], *Tomsk Journal of Linguistics and Anthropology* 1.9 (2018): 59–68.
- [21] P. Roseano, M. González, J. Borràs-Comes, P. Prieto, Communicating Epistemic Stance: How Speech and Gesture Patterns Reflect Epistemicity and Evidentiality, *Discourse Processes* 53.3: 135–174. doi:10.1080/0163853X.2014.969137.
- [22] U. Hadar, T. Steiner, E. Grant, F. C. Rose, Kinematics of head movements accompanying speech during conversation, *Human Movement Science* 2.1 (1983), pp. 35–46.
- [23] A. Benoit, A. Caplier, Head nods analysis: interpretation of nonverbal communication gestures, in: *IEEE International Conference on Image Processing, ICIP '2005*, IEEE, Genova, Italy, 2005, pp. III–425. doi: 10.1109/ICIP.2005.1530419.
- [24] J. Allwood, L. Cerrato, K. Jokinen, C. Navarretta, P. Paggio, The MUMIN coding scheme for the annotation of feedback, turn management and sequencing phenomena, *Language Resources and Evaluation*, 41.3 (2007): 273–287. doi: 10.1007/s10579-007-9061-5
- [25] D. Heylen, Listening heads, in: I. Wachsmuth, G. Knoblich (Eds.), *Modeling communication with robots and virtual humans*, Springer, Berlin, 2008. doi: 10.1007/978-3-540-79037-2_13.
- [26] S. Kousidis, Z. Malisz, P. Wagner, D. Schlangen, Exploring annotation of head gesture forms in spontaneous human interaction, in: *Proceedings of the Tilburg Gesture Meeting, TiGeR '2013*, 2013.
- [27] S. Kousidis, J. Hough, D. Schlangen, Exploring the Body and Head Kinematics of Laughter, Filled Pauses and Breaths, in: *The 4th Interdisciplinary Workshop on Laughter and Other Non-verbal Vocalisations in Speech*, Enschede, Netherlands, 2015.
- [28] A. A. Kibrik, O. V. Fedorova. An empirical study of multichannel communication: Russian Pear Chats and Stories, in: *Psychology. Journal of the Higher School of Economics* 15.2 (2018): 191–200.
- [29] Chafe W. (Ed.), *The pear stories: Cognitive, cultural, and linguistic aspects of narrative production*, Ablex, Norwood, NJ, 1980.

- [30] N. A. Korotaev, “Russian Pear Chats and Stories”: Vocal annotation guide, Version 10.01.2019. URL: <http://multidiscourse.ru>
- [31] A. A. Kibrik, N. A. Korotaev, V. I. Podlesskaya, Russian spoken discourse: Local structure and prosody, in: Sh. Izre'el, H. Mello, A. Panunzi, T. Raso (Eds.), *Search of Basic Units of Spoken Language: A corpus-driven approach*, volume 94 of *Studies in Corpus Linguistics*, John Benjamins, Amsterdam, 2020, pp. 35–76. doi: 10.1075/sci.94.01kib
- [32] Litvinenko A. O., Nikolaeva Ju. V., Kibrik A. A. (2017), Annotirovaniye russkikh manual'nykh zhestov: teoreticheskiye i prakticheskiye voprosy [Annotation of Russian manual gestures: Theoretical and practical issues], in: *Computational Linguistics and Intellectual Technologies: Papers from the Annual International Conference “Dialogue 2017”*, RGGU, Moscow, 2017, pp. 255–268.
- [33] J. Bressem, C. Müller. A repertoire of German recurrent gestures with pragmatic functions, in: C. Müller, A. Cienki, E. Fricke, S. H. Ladewig, D. McNeill, J. (Eds.), *Body – Language – Communication: An international Handbook on Multimodality in Human Interaction*, volume 38.2 of *Handbooks of Linguistics and Communication Science*, De Gruyter Mouton, Berlin/ Boston, 2014, pp. 1575–1592
- [34] S. Alexanderson, D. House, J. Beskow, Extracting and analysing co-speech head gestures from motion-capture data, in: R. Eklund (Ed.), *Proceedings of Fonetik 2013, the XXVIth Swedish Phonetics Conference*, volume 21 of *Studies in Language and Culture 21*, Linköping University Electronic Press, Linköping, 2013, pp. 1–4.
- [35] E. Krahmer, M. Swerts, The effects of visual beats on prosodic prominence: Acoustic analyses, auditory perception and visual perception, *Journal of Memory and Language* 57 (2007): 396–414. doi: 10.1016/j.jml.2007.06.005
- [36] D. McNeill, E. T. Levy, S. D. Duncan, Gesture in discourse, in: D. Tannen, H. E. Hamilton, D. Schiffrin (Eds.), *Handbook of discourse analysis*, Blackwell, Oxford, 2015, pp. 262–290.
- [37] M. Theune, C. J. Brandhorst, To beat or not to beat: Beat gestures in direction giving, in: S. Kopp, I. Wachsmuth (Eds.), *Gesture in embodied communication and human-computer interaction*, Springer Verlag, Berlin, Heidelberg, 2010, pp. 195–206.
- [38] E. Biau, S. Soto-Faraco, Beat gestures modulate auditory integration in speech perception, *Brain and Language* 124 (2013): 143–152. doi: 10.1016/j.bandl.2012.10.008
- [39] E. Biau, S. Soto-Faraco, Synchronization by the hand: The sight of gestures modulates low-frequency activity in brain responses to continuous speech, *Frontiers in Human Neuroscience* 9:527 (2015). doi: 10.3389/fnhum.2015.00527
- [40] H. Holle, C. Obermeier, M. Schmidt-Kassow, A. D. Friederici, J. Ward, T. C. Gunter. Gesture facilitates the syntactic analysis of speech, *Frontiers in Psychology* 3:74 (2012). doi: 10.3389/fpsyg.2012.00074
- [41] Korotaev N. A., Kibrik A. A., Podlesskaya V. I., Annotating the vocal modality, in: O. V. Fedorova, A. A. Kibrik (Eds.), *The MCD handbook: A practical guide to annotating multichannel discourse*. To appear.
- [42] N. V. Sukhova, A. A. Evdokimova, Cephalic Annotation Scheme, in: O. V. Fedorova, A. A. Kibrik (Eds.), *The MCD handbook: A practical guide to annotating multichannel discourse*. To appear.
- [43] A. A. Evdokimova. Novye tipy pragmaticheskikh zhestov golovy – Pragmatic center I Pragmatic away [New types of pragmatic gestures, pragmatic-center and pragmatic-away], *Lingvistika I metodika prepodavanija inostrannykh jazykov* 12 (2020): 136–148. doi: 10.37892/2218-1393-2020-12-1-136-148 (in Russian)
- [44] E. A. Grishina, *Russkaya zhestikulyatsiya s lingvisticheskoy tochki zreniya* [Russian gestures from a linguistic perspective], *Jazyki slavyanskoy kul'tury*, Moscow, 2017. (in Russian)
- [45] O. V. Fedorova, A. A. Kibrik (Eds.), *The MCD handbook: A practical guide to annotating multichannel discourse*. To appear.
- [46] C. Müller, Forms and uses of the Palm Up Open Hand. A case of a gesture family?, in C. Müller, R. Posner (Eds.), *Semantics and pragmatics of everyday gestures*, Weidler, Berlin, 2004, pp. 233–256.