CORPUS-BASED INVESTIGATION OF QUOTATION IN RUSSIAN SIGN LANGUAGE[1]

Khristoforova E. A. (evkhristoforova@gmail.com),
Kimmelman V. I. (vadim.kimmelman@gmail.com)

Russian State University for the Humanities, Moscow, Russia

This paper presents corpus-based research of quotation constructions in Russian Sign Language (RSL). Quotation constructions have been observed from different perspective in different signed and spoken languages [Brendel, Meibauer, Steinbach 2011]; [Litvinenko et al. 2009]. Based on the corpus of spontaneous narratives recorded from RSL signers [Burkova 2015], we conducted a quantitative analysis of these constructions. We analyzed constituents of quotation construction, such as the source (author of utterance) indication, the introducing matrix predicate, and the quote. Our investigation of non-manual markers in the corpus revealed that non-manual marking of quotation is optional for RSL quotations. We distinguished direct and indirect quotations in our data based on the reference of indexical elements, the use of subordinating conjunction, and the imperative mood. We found that in RSL non-manuals do not mark the direct/indirect type of quotation. Our data show that RSL signers tend to use direct quotation much more frequently than indirect quotation. In addition, we compared our findings with the data on quotation constructions in some other sign languages and with the studies of quotation in natural discourse of spoken languages. This comparison showed that RSL quotations share core properties with quotations in spoken and signed languages [Litvinenko et al. 2009].

Key words: quotation, sign languages, RSL, corpus-based research, non-manual markers

КОРПУСНОЕ ИССЛЕДОВАНИЕ ЦИТАЦИОННЫХ КОНСТРУКЦИЙ В РУССКОМ ЖЕСТОВОМ ЯЗЫКЕ

Христофорова Е. А. (evkhristoforova@gmail.com),
Киммельман В. И. (vadim.kimmelman@gmail.com)

Российский государственный гуманитарный университет, Москва, Россия

В данной статье представлено корпусное исследование цитационных конструкций в русском жестовом языке (РЖЯ). Цитационные конструкции были изучены с разных точек зрения на материале как жестовых, так и звуковых языков [Brendel, Meibauer, Steinbach 2011; Litvenenko et al. 2009]. Для настоящего исследования мы использовали корпус спонтанных нарративов, записанных от носителей РЖЯ [Burkova 2015]. Анализ корпуса позволил количественно описать такие составляющие цитационных конструкций, как указание на автора высказывания, вводящую предикацию и собственно саму цитацию. В процессе анализа немануальных маркеров, представленных в корпусе, было обнаружено, что немануальное маркирование цитаций не является обязательным в РЖЯ. В нашем корпусе мы разделяли прямую и косвенную цитацию, основываясь на следующих критериях: сдвиг референции индексикалов, наличие подчинительного союза и показателей императива. Мы обнаружили, что различие между прямым и косвенным типом цитации не маркируется немануально. Мы отметили, что носители РЖЯ используют прямую цитацию значительно чаще, чем косвенную. Сравнив наши результаты с данными исследований цитационных конструкций в других жестовых языках и в естественном дискурсе звуковых языков, мы пришли к выводу, что цитационные конструкции РЖЯ имеют много общего с цитацией в звучащих и жестовых языках [Litvinenko et al.2009].

Ключевые слова: цитация, жестовые языки, РЖЯ, корпусное исследование, немануальные маркеры

1. Introduction

Quotation constructions, that is, the means of conveying other's words and thoughts have been the topic of numerous studies [Brendel, Meibauer, Steinbach 2011]. Typically, direct and indirect quotation are distinguished. Direct quotation is almost verbatim representation of an utterance, while indirect quotation is a report from narrator's perspective. In writing, one may observe the difference in punctuation between these two types. In natural discourse, different intonation can be used for direct or indirect quotation. In addition, these two types of quotation differ in the reference of indexicals—elements whose reference depends on the context, such as personal pronouns and time adverbials. In direct speech, indexicals must be interpreted within the context of the quoted situation, while in indirect quotation they are interpreted within the overall context of the narration. There are also structural differences between direct and indirect quotation. Indirect quotation in most languages tends to be expressed by an embedded clause. Direct quotation, on the contrary, is syntactically independent. Notwithstanding all the differences between the types of quotation, it is not always easy to distinguish direct and indirect quotation in natural discourse [Litvinenko et al. 2009].

Sign languages also have quotation constructions. Investigating quotation in sign languages, researchers observed a phenomenon called "role shift" [Herrmann & Steinbach 2012; Quer 2011; Schlenker 2017]. Signers tend to shift into the role of a character using the range of non-manual markers. Among these non-manual markers are leans or turns of the body and head, change of eye gaze direction, and

different emotional facial expressions. Performing role shift signers not only sign from a character's point of view conveying his words, emotions, or thought, but they can also act from character's perspective presenting his actions. This is different from the quotation in spoken languages [Liddell & Metzger 1998].

Role shift has been explored from different perspectives. Some researchers concluded that the use of role shift does not always clearly indicate direct quotation, but shows some properties of indirect speech. Much research has been done to identify non-manual markers, which accompany quotation in sign languages [Herrmann & Steinbach 2012 i.a.].

This paper describes quotation in Russian Sign Language (RSL). We identified quotation constructions in a corpus of spontaneous narratives. We described the constituents of quotation constructions, the most frequent non-manual markers, and the differences between direct and indirect quotation. We also compared our findings with quotation in other sign languages and in natural discourse of spoken languages.

2. Methodology

We have chosen to investigate quotation in RSL based on corpus data. Corpus-based research methods have only recently been applied to sign languages (see [Lucas, Bayley and Valli 2001] for one of the first studies). Corpus research gives a possibility for quantitative studies of natural language, and sign languages are not the exception. Modern computer-based devices, such as ELAN or SignStream, provide sign language linguists with tools for multilayered annotation [Safar & Glauert 2012]. Considering multi-modal nature of sign languages, these tools are essential for full description and profound investigation of sign languages. There are also structured search engines that allow users to study overlap between distinct layers of description. A common obstacle that sign language linguists are bound to deal with is the lack of automatic annotation devices. Tagging, parsing and transcription are all completely manual. That makes annotation process complicated, time-consuming and highly annotator specific. In addition, manual factor leads to the low productivity of the annotation process, which may excuse relatively small sizes of most corpora [Safar & Glauert 2012].

On-line RSL corpus was created by a team of sign language researchers lead by [Burkova 2015]. Currently, this corpus is the only RSL corpus available for public use. It contains around 200 video recordings of different text types (picture-based storytelling, interviews, spontaneous narrations, elicitations, etc.). Most recordings are annotated in ELAN software[2] using four layers: glosses for the right hand signs, glosses for the left hand signs, overall translation of the clause and comments (**Figure 1**).

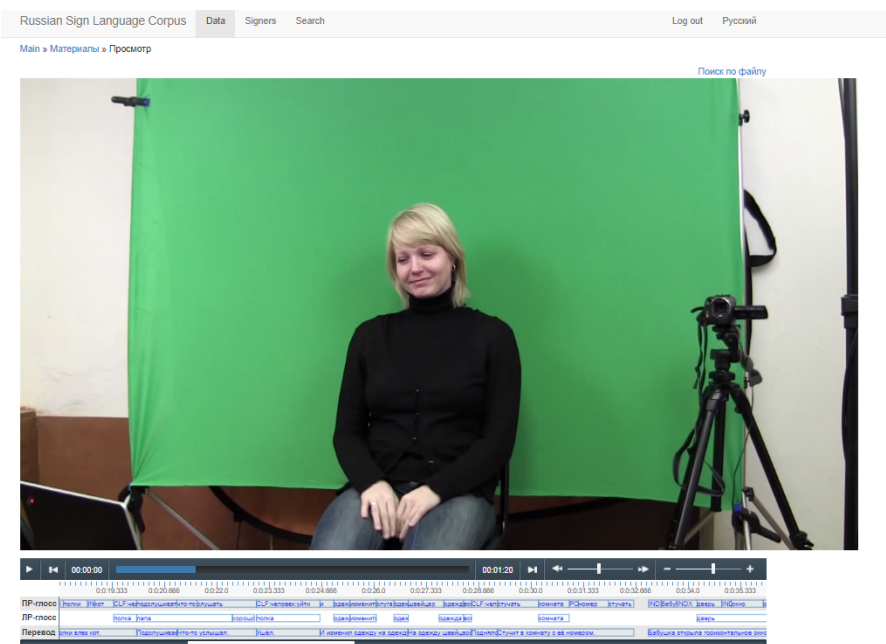[2]   https://tla.mpi.nl/tools/tla-tools/elan/

Fig. 1. Screenshot of the website

RSL is spread across the large territory of Russian Federation, which necessarily leads to some dialectal variation (Schembri & Johnston 2012). On-line RSL corpus mostly represents Moscow and Novosibirsk variants of RSL. These two sections of data must be carefully analyzed in order to sort out the influence of dialectic variation, but this goes beyond the purposes of thisstudy. For this study, we have chosen spontaneous narrations recorded in Moscow. This part of the corpus contains videos recorded from eleven RSL signers.

The chosen section of corpus contains almost 8,000 signs, forming nearly 1,200 clauses. Although this amount of data is relatively small, it contains 341 quotations. We would like to highlight that we consider not only reporting the speech of others but also reporting thoughts and attitudes. Conveying these meanings, speakers and signers also use direct or indirect quotation. Using ELAN, we annotated all the videos using the following tiers:

- Type of quotation (speech/thought/attitude);
- Matrix verb, introducing quotation: the verb of speech (or thought/attitude) if used;
- Source of quotation: whether the author of quotation is lexically introduced or not. If there is no lexical sign for the source but it matches the subject of the preceding action enacted by the signer, the source may be easily reconstructed from the context and its lexical representation would be excessive. We indicated such cases as "constructed action";

- Author of quotation: whether the signer himself is the author of quotation (s/he uttered it in the past) or some other person uttered it;
- Non-manual markers: eye gaze direction (distinct from the one towards the interlocutor), head turns, and body leans and turns;
- Indexicals: indexical elements (deictic elements, time indications, agreeing verbs personal pronouns, possessive markers) in the quotation and their reference (shifted or not);
- Subordination features: elements moved from the quotation to the matrix clause;
- Markers of direct or indirect speech: imperative markers or conjunctions.

The non-manual markers we annotated (eye gaze, head turns, body leans and turns) are the ones most commonly associated with quotation in sign languages [Herrmann & Steinbach 2012]. Another common marker is emotional facial expressions attributable to the author of the quote. However, various researchers (ibid. a.o.) have demonstrated that facial expressions cannot be analyzed as markers of quotation as they are a part of the quoted utterance; they also clearly occur without quotation. We thus left them out of our analysis.

3. Properties of quotation in RSL

For RSL quotation, we have revealed the same basic constituents as previously identified for quotation constructions in other languages: the indication of the source (author) of quotation; introducing matrix predicate; the quotation itself (1).

(1) I $_{source}$ SAY $_{matrix\ predicate}$ [I YES-YES VIA MOSCOW GO_BY_TRAIN]$_{quotation}$
I say: "Yes, I go via Moscow by train".

Source indication and introducing matrix predicates are optional. Source is indicated in 160 (47%) cases, while in 118 (35%) cases it is not mentioned at all. Authors of other 52 (15%) quotations matched subjects of previously described actions, so that the source of these quotations was not expressed lexically, but was easily retrieved from the context. According to [Mathis & Yule 1994], an optional source indication is also common for spoken languages.

Introducing matrix predicates share the same property of optionality. 218 (64%) quotations are not introduced by any predicate. We have also found that in 27 (8%) cases the quotation is proceeded by a sign called "palms up" instead of matrix predicate (2). We assumed that this sign might also be analyzed as a quotation marker. In RSL "palms up" is considered to be a multifunctional sign without any specific lexical meaning. Although we suggest that it is used to introduce quotations in our data, we still cannot exclude that this sign may bear other functions. Interestingly, according to our data, "palms up" tend to introduce quotations whose author is the signer himself but in past. Optional indication of matrix predicates of quotation is also among the properties of quotation in spoken languages [Litvinenko et al. 2009].

(2)   I PALMS_UP YES YES OKAY IMPERATIVE
      I say: "Yes, okay, let's go!"

Figure 2. "Palms up" sign

As for 123 (36%) cases of matrix predicates of quotation, the most frequent are THINK, TELL, SAY, ASK and CALL. In addition, we have identified a class of verbs that introduce quotation quiet frequently but do not belong to verbs of speech or thoughts such as CALL: syntactically, it introduces quotation, but in fact, it describes the action preceding quotation (3).

(3)   DAVYDENKO(proper noun) CALL-1 NUMBER BOX WHAT BOX
      Davydenko calls me: "What is the number of box?"

Non-manuals are another typical marker of quotation. As was already mentioned in section 2, in this paper we analyze eye gaze, head turns, and body movements and do not take into account facial expressions (4). Non-manual marking in RSL also turned out to be non-obligatory. 195 (57%) quotations are marked by the change of eye gaze direction, 175 (51%) are marked by body movements, and 287 (84%) by head turns. Head turns thus seem to be the most reliable marker, but see below. We have found 95 (26%) quotations without gaze or body turns, and 16 (5%) that simply have no non-manual marking.

(4)   <u>eye gaze, head turn, body lean</u>
      CALL-3 INTERESTING THIS
      I call her: "This is interesting!"

Figure 3. Non-manual markers of quotation

In our data, non-manuals in RSL, if they are present, do not always accompany only quotation itself as typically observed for other sign languages. As shown in Table 1, only in 47 (14%) cases for eye gaze, in 58 (17%) cases for body movements, and in 62 (18%) for head turns these non-manuals mark the whole quotation and nothing else. Our data shows that non-manuals in RSL can also accompany other constituents of quotation constrictions. In fact, these results challenge our assumption that eye gaze and body movement mark quotations in all the cases.

Table 1. Non-manual marking of different constituents of quotation constructions

	Eye gaze	Body movement	Head turns
Part of quotation	52	63	62
Whole quotation	47	58	62
Whole quotation + matrix predicate	38	22	46
Whole quotation + matrix predicate + source indication	58	32	117

Note that body leans are relatively more often used to mark the whole quotation than eye gaze (χ^2=12.9, p = 0.005) which might indicate that it is a better marker overall (although also highly optional). Head turns are also different significantly from both gaze and body movement because they are very often used even on the source. This shows that they are not really good markers of the quote itself even though they are very frequent.

Considering all our findings discussed above, we tried to come up with a hypothesis explaining why these markers are present in some cases and absent in others. For example, we checked whether non-manual marking is connected with the report of thoughts instead of the report of speech (Table 2). We found that body movement is used significantly more often for reported thought than speech (χ^2 = 9.6, p = 0.02). Importantly, body movements are used with both types. The differences for eye gaze and head turns were not significant. We also proposed that non-manual markers are less frequent when the signer himself is the author of quotation. However, this hypothesis was not supported by our data either: no differences are statistically significant (Table 3).

Table 2. Non-manual marking of different types of quotation

	Eye gaze	Body movement	Head turns	Total
Reported speech	155 (56%)	131 (47%)	245 (88%)	277
Reported thoughts/ attitudes	40 (62%)	44 (68%)	52 (81%)	64

Table 3. Non-manual markers and source of quotation

	Eye gaze	Body movement	Head turns	Total
Author of quotations is the signer	107 (61%)	94 (55%)	141 (82%)	171
Author of quotation is not the signer	99 (58%)	86 (51%)	147 (86%)	170

We further hypothesized that non-manual markers would be used more frequently when not a single utterance but a whole dialogue is quoted. We found nineteen cases of reported dialogues in the corpus. It turned out that that non-manual marking of quotations within the dialogues is similar to the overall pattern, so this hypothesis was not confirmed.

Finally, we hypothesized that non-manual marking correlates with the type of quotation (direct or indirect). Before we test this prediction, it is necessary to present our findings concerning direct and indirect speech in RSL.

4.   Direct and indirect quotation

As was already mentioned in section 1, we can identify direct or indirect quotation basing on the following basic criteria:
- Reference of indexicals (shifted/non-shifted)
- Syntactic status of quotation as a sentence (embedded clause/independent clause).
- Possibility of imperative quotes (possible/impossible)

First, we analyze indexical elements in our data. Most frequent of them are personal pronouns (I, YOU, HE/SHE); possessive markers; agreeing verbs; time and place adverbials like HERE, and NOW; tense markers for past and future.

It turned out that 196 (57%) quotations do not have indexical elements. Another obstacle is that, in the case of the signer quoting his own speech in the past, the first-person pronoun can refer to the signer in the context of signed quotation (as the author of quotation) and to the signer in the context of general narration within which the quotation is reported. In such cases, it is impossible to define the reference of the pronoun as shifted or not.

Considering indexical element whose reference can be identified without complications, we found that 95% (86/91) of them have shifted reference. It implies that

for the quotations with the use of indexical elements direct quotation type is much more frequent in RSL.

As far as we have at least 86 examples of direct quotations and 8 examples of indirect quotations, we can proceed to the investigation of the connection between non-manual marking and the type of quotation. We hypothesized that non-manuals might only mark direct quotations, as direct speech is supposed to be more emotional (as it is in spoken languages). However, our data shows that the use of non-manuals is not defined by the type of quotation. Among 86 direct quotations, we have found 31 quotations without eye gaze, 30 quotations without body movements, 8 without head turns, and 3 without any non-manual marking. Also recall that in many examples, including those with shifted indexicals, the non-manuals are not aligned with the quote. Thus, non-manual markers are not obligatory for direct quotations.

Apart from the reference of indexicals, we also analysed features of subordination. The most obvious marker of quotation represented by an embedded clause is a subordinating conjunction THAT. This conjunction is used after matrix predicate in order to introduce embedded quotation. We assume that if THAT is used to introduce particular quotation, this quotation should be considered as indirect.

Constructions with the use of subordinating conjunction THAT are quite rare in our data: only seven cases were found (5). Most likely such constructions are borrowed from Russian. It is also important to note that among seven quotations with subordinating conjunction some are accompanied by non-manual markers. This is another proof that non-manuals do not only mark direct quotations.

(5) EXPLAIN THAT DEFINITELY THERE PST GO PST
     We explain that we definitely came.

Our last criterion to identify the type of quotation is the use of imperatives quotes. According to [Kuno 1988], imperatives within quotation are a clear attribute of direct speech. Although not all researches agree with this statement, we decided to examine all the cases of imperatives within quotations in our data in order to investigate whether the use of imperatives correlate with the type of quotation. We have found 15 cases of the use of imperative manual marker within quotations. None of these quotations are introduced by subordinating conjunction and none of them contain indexical elements with non-shifted reference. Consequently, we consider them direct quotations, which implies that imperatives can be considered as a direct speech indication in RSL.

Summing up this section, we can give the following description of direct and indirect quotation constructions in RSL: direct quotation is the most frequent pattern to convey somebody's words/thoughts/attitude in RSL. It is explicitly indicated by the shifted reference of indexical elements, by the lack of subordinating conjunction THAT, or by imperatives within quotation. Our assumption about non-manual marking of direct or indirect quotations has not been confirmed: according to our data non-manual marking is optional and not fully aligned with quoteseven in direct speech.

5. Summary

Our aim was to describe quotation constructions in RSL based on corpus materials. Within on-line RSL corpus we have chosen spontaneous narrations, recorded in Moscow. In this part of data, we found 341 quotations. Using ELAN, we annotated each quotation by tiers listed in section 2, which allowed us to investigate the properties of quotation constructions in RSL from different perspectives.

Quotation constructions in RSL consist of the same elements as quotations described in other natural languages. These elements are source (author) indication, introducing matrix predicate, and the quotation itself. Source indication and introducing matrix predicate are optional. Although at least one of these elements accompanies the majority of quotations in our data, it is possible to find quotation construction consisting only of quotation: 12% (40) of quotation constructions in our data do not have either source indication or introducing matrix predicate.

We also investigated non-manual marking of quotations in RSL. Although many sign language researchers state that non-manual marking of quotations is obligatory or at least highly frequent [Herrmann & Steinbach 2012], our data shows that for RSL quotations non-manual markers are optional. The most common marker is the head turns but it usually marks the source and (a part of the) quote, not just the quote.

Following generally accepted criteria of direct/indirect speech (reference of indexical elements, syntactic structure, and possibility of imperatives), we found that for more than a half of quotations in our data we cannot define the type of quotation. Among those quotations that can be identified as direct or indirect, the majority of quotations (95%) are direct. RSL shares this property with spoken languages, in which direct quotations are also more frequent [Litvinenko et al. 2009]. This fact may be also explained by the relatively young age of RSL. It is possible that syntactic structure of indirect quotation in RSL may be in its early stage of grammaticalization [Pfau et al. 2016], but such processes are still understudied.

We also hypothesized that non-manual markers may indicate the type of quotation (that is, that only direct speech will have them). However, our data contradicted this hypothesis as described in detail in section 4. Difficulties with identifying the type of quotation are not unique for RSL: natural discourse of spoken languages has the same property [Litvinenko et al. 2009].

References

1. *Brendel E., Meibauer J., Steinbach M.* (2011), Exploring the meaning of quotation, Understanding Quotation, Berlin: De Gruyter Mouton, pp. 1–34.
2. *Burkova, S.* (2015), Russian Sign Language Corpus: http://rsl.nstu.ru/.
3. *Herrmann A., Steinbach M.* (2012), Quotation in sign languages: A visible context shift, Quotatives. Cross-linguistic and Cross-disciplinary Perspectives, Amsterdam: John Benjamins, pp. 203–230.
4. *Litvinenko, A., Korotaev N., Kibrik A., Podlesskaya V.* (2009), Constructions with quotation or "reported speech" [Konstruktsii s tsitatsiey, ili "chuzhoi rechyu"], Night Dream Stories: A Corpus Study of Spoken Russian Discourse [Rasskazi

o snovedeniah: Korpusnoye issledovanie ustnogo russkogo diskursa], Moscow: Language of Slavic Cultures, pp. 288–307

5. *Kuno S.* (1988), Blended quasi-direct discourse in Japanese, Papers from the Second International Workshop on Japanese Syntax, Stanford: CSLI, pp. 75–102.

6. *Liddell S. K., Metzger M.* (1998), Gesture in sign language discourse, Journal of Pragmatics, Vol. 30, № 6, pp. 657–697.

7. *Lucas C., Bayley R., Valli C.* (2001), Sociolinguistic Variation in American Sign Language. Washington, DC: Gallaudet University Press.

8. *Mathis T., Yule G.* (1994), Zero quotatives, Discourse Processes, Vol. 18, № 1, pp. 63–76.

9. *Pfau R., Steinbach M., Herrmann A.* (2016), A matter of complexity: subordination in sign languages, Boston: De Gruyter Mouton.

10. *Safar E., Glauert J.* (2012), Computer modelling, Sign Language: An International Handbook, pp. 1075–1101.

11. *Schlenker P.* (2017) Super monsters I: Attitude and Action Role Shift in sign language, Semantics and Pragmatics, Vol. 10, № 9.

12. *Schembri A., Johnson T.* (2012) Sociolinguistic aspects of variation and change, Sign Language: An International Handbook, pp. 788–816.

13. *Quer J.* (2011), Reporting and quoting in signed discourse, Quotation, Berlin: De Gruyter Mouton, pp. 277–302.