

**ИССЛЕДОВАНИЕ ВАРИАТИВНОГО ОФОРМЛЕНИЯ
ПРЯМОГО ДОПОЛНЕНИЯ В УДМУРТСКОМ ЯЗЫКЕ
В ЗЕРКАЛЕ ДВУХ МЕТОДОЛОГИЧЕСКИХ ПОДХОДОВ:
ЭЛИЦИТАЦИЯ И КОРПУСНЫЙ АНАЛИЗ***

Н. В. Сердобольская

Институт языкознания РАН,

Государственный институт русского языка им. А. С. Пушкина

Данные элицитации выделяют референциальный статус в качестве базового фактора распределения маркирования прямого дополнения в бесермянском удмуртском, в то время как для литературного языка более важным фактором принято считать одушевленность. В настоящей работе анализируются данные корпуса бесермянского удмуртского (10 539 предложений, 2187 именных групп в позиции прямого дополнения). Корпусной анализ позволяет определить конкретный вес различных значений двух факторов. В контексте одушевленных прямых дополнений аккузатив расширяет свое употребление на неопределенные и нереперентные группы (за исключением некоторых лексических классов), а в контексте неодушевленных начинает употребляться в контексте определенных или «тяжелых» групп.

Ключевые слова: вариативное оформление прямого дополнения, одушевленность, определенность, аккузатив, посессив, удмуртский язык, бесермянский диалект, бесермянское наречие

* Работа поддержана грантом РФФ №18-18-00462. Автор приносит благодарность С. Ю. Толдовой, Д. В. Горшкову и О. И. Беляеву за помощь в обработке данных.

**DIFFERENTIAL OBJECT MARKING IN UDMURT
APPROACHED BY TWO METHODOLOGIES:
ELICITATION VS. CORPUS ANALYSIS***

*Natalia Serdobolskaya
Institute of Linguistics RAS,
Pushkin State Russian Language Institute*

Elicitation sessions in Beserman Udmurt show the relevance of referential properties of the DO for the choice of its marking; however, for Standard Udmurt animacy has been claimed to play a more significant role. This study involves a corpus analysis of Beserman Udmurt (10 539 sentences, 2187 DOs), which permits to establish the exact rank of each value of both parameters. For animates, the use of the accusative is expanded onto indefinite/non-specific DOs, excluding some lexical classes; for inanimates, its use is shifted from definite DOs to definite and “heavy” DOs.

Keywords: differential object marking, animacy, definiteness, accusative, possessive, Udmurt, Beserman

* The work is supported by the RSF grant No 18-18-00462. I would like to thank Svetlana Toldova, Dmitriy Gorshkov and Oleg Belyaev for their help with data processing.

1. Introduction

Linguists are not always unanimous in regards to the choice of the methodology while working with minority languages. Specifically, many debates are raised around the advantages and drawbacks of two methodologies, elicitation and corpus-based analysis. This study contributes to this debate, showing what information can be found by application of those methods to a specific phenomenon in one and the same idiom.

The study is focused on differential object marking in Beserman Udmurt. The phenomenon of differential object marking (DOM) has gained wide attention of researchers since [Bossong 1985] on Persian. It involves constructions with transitive verbs that mark their objects in two or more different ways. For example, in Spanish animate DOs can take a preposition *a*, while inanimate DOs occur unmarked [de Swart 2007: 129]:

- (1) a. *Mari vió a la mujer.*
Mari saw a the woman
'Mari saw the woman.'
- b. *Mari vió (*a) la mesa.*
Mari saw a the table
'Mari saw the table.'

DOM has been analyzed both as a separate phenomenon [de Swart 2007] and as part of a large number of phenomena under the label of differential argument marking [Witzlack-Makarevich, Seržant 2017], together with differential subject marking. DOM is widespread among language families and areals, comprising the Uralic languages.

The list of factors influencing the choice of the marker in cases similar to (1)–(2) includes the following [Moravcsik 1978]: tense, aspect, modality, polarity, information structure of the sentence, referential properties of the DO, DO's animacy.

This paper treats this phenomenon in one of the Permic languages, Udmurt (Beserman dialect), comparing the results obtained in elicitation sessions and the results of the corpus analysis. The data come from my field data collected in 2003–2005 and 2009–2012 and the corpus of Beserman Udmurt (http://beserman.ru/corpus/search/?interface_language=en; approx. 75 000 tokens).

2. DOM in Permic languages and in Beserman Udmurt

[Witzlack-Makarevich, Seržant 2017] classify the systems of DOM based on morphological markedness of the DO, among other parameters; the systems with non-marked DOs are termed as asymmetrical, as they present the opposition of presence vs. absence of morphological encoding. Permic languages offer the threefold variant of DOM, where the non-marked variant (2) is opposed to two DO markers, one of which is the accusative case marker (3) and the one belongs to the paradigm of possessive markers inflected for person and number (4). See some examples from Beserman Udmurt:

- (2) *uj-âñ nâl-âz gur est-i-z.*
 night-LOC girl-P.3(SG) oven stoke-PST-3(SG)
 ‘At night the girl has stoked the oven.’ [Corpus]

- (3) *ž’ič’â až’-i-z kâš’pu-ez.*
 fox see-PST-3(SG) birch-ACC
 ‘(The wolf carried a birch to make a new shaft for the cart.) The fox saw the birch and scolded the wolf.’ [Corpus]

- (4) *nu nâl-de ta-t-âš’ č’ašša-je gu-e.*
 carry(IMP.SG) daughter-ACC.P.2(SG) this-OBL-EL forest-ILL pit-ILL
 ‘(The step-mother said her husband) Carry your daughter away to the forest hut.’ [Corpus]

Udmurt has developed a special marker *-tâ* for plural DOs (non-possessive):

- (5) *so vel’t-e č’ašja-je č’âž-jos-tâ âb-âl-ânâ.*
 that go-PRS.3SG forest-ILL duck-PL-ACC.PL shoot-ITER-INF
 ‘He often goes to the forest to shoot ducks.’

This marker is in complimentary distribution with the possessive, cf. (5) and (6).

- (6) *Vaš’a pâd-jos-se kott-i-z.*
 Vasya leg-PL-ACC.P.3(SG) wet-PST-3(SG)
 ‘Vasya has wetted his legs.’

DOs with plural semantics can also occur without the marker of plurality; the distribution of the plurality marker is beyond the scope of this paper.

The possessive markers of DO are part of the large paradigm that differentiates markers based on the possessor’s person, number, syntactic position of the

head and (in)alienability. As shown in Table 1, the DO set is differentiated from other markers both in form (different vowel and vowel/consonant order, special marking of plural possessors) and function (it does not distinguish between alienable and inalienable possession).

Table 1. The paradigm of possessive suffixes in Beserman Udmurt¹

	DO set	non-DO set: inalienable	non-DO set: alienable
P.1SG	<i>-me</i>	<i>-(j)â / -m²</i>	<i>-(j)e</i>
P.2SG	<i>-de / -te</i>	<i>-(j)âd / -d</i>	<i>-(j)ed</i>
P.3SG	<i>-ze / -se</i>	<i>-(j)âz / -z</i>	<i>-(j)ez</i>
P.1PL	<i>-mes</i>	<i>-(â)mâ</i>	
P.2PL	<i>-des / -tes</i>	<i>-(â)dâ / -tâ</i>	
P.3PL	<i>-zes / -ses</i>	<i>-(â)zâ / -sâ</i>	

The possessive markers can denote the possessive relation between the DO and some participant in the discourse, as in (6). However, they are widely used as referential devices [Suihkonen 2005, Winkler 2011], as in (7), where the DO ‘dust’ is aforementioned.

- (7) *val = no kopot'-se š'ij-e = ke kâz-e.*
 horse = ADD dust-ACC.P.3(SG) eat-PRS.3SG = if cough-PRS.3SG
 ‘(Our hay is with dust.) And each time the horse eats (some of) the dust, it coughs.’ [Corpus]

In similar cases the possessive relation can hardly be observed, and the possessive suffixes are obviously employed as referential markers. This leads some researchers to consider the hypothesis of the article-like status of possessive markers in Udmurt [Fraurud 2001]; however, this hypothesis is rejected based on the non-obligatoriness of possessives in contexts of definiteness/specificity (7) and the large spectrum of meanings they develop (see [Fraurud 2001] for details). The following range of meanings is observed both in Komi varieties and in Udmurt: definiteness, endearment, vocative function, anaphoric function, ethical function (“associative relation” to the hearer/the protagonist), syntactic function: agreement with the modifier in nominal and cardinal phrases

¹ The pronominal set of markers in *-im / -id / -iz* is not included since the present work is only focused on nouns.

² The *-m / -d / -z* variant is used after case markers ending in a vowel.

and in non-finite clauses, see [Alatyrev 1983, Edygarova 2010, Kel'makov 1996, Kuznecova 2012, Suihkonen 2005, Winkler 2011] for Udmurt and [Schlachter 1960, Klumpp 2008] for Komi.

In addition, [Serdobolskaya 2017] and [Serdobolskaya, Usacheva, Arkhangelskiy 2019] identify the following functions: partitive indefinite (indefinite part of a definite set or mass), contrastive topic, semi-active DOs (re-activation of the previous topic in the discourse), introduction of a new topic of the discourse. [Serdobolskaya, Usacheva, Arkhangelskiy 2019] argue for the analysis of possessive markers in terms of pragmaticization: the possessive markers are used in pragmatic functions and constructions associated with these meanings.

In sum, the following markers of DO are available in Beserman Udmurt: no marking, accusative *-ez* for singulars and *-tâ* for plurals, and the possessive markers (Table 1). The present study is mostly focused on the distribution of the non-marked variant and the accusative. The detailed analysis of possessives in DO is given in [Serdobolskaya 2017].

The major works on DOM in Standard Udmurt claim that the following factors govern the distribution of non-marked vs. accusative DOs: definiteness [Perevoshchikov et al. 1962: 93; Csucs 1990: 34; Winkler 2001: 20; Kondrat'eva 2002], quantification and partitivity of the DO [Perevoshchikov et al. 1962; Kondrat'eva 2002; Winkler 2011: 46], animacy of the DO [Kondrat'eva 2010]. For Beserman, the quantification factor plays a minor role in a limited number of contexts. The following discussion is mostly focused on animacy and definiteness of DOs.

3. Animacy vs. definiteness in DOM: results of elicitation sessions in Beserman Udmurt

[Kondrat'eva 2010] claims that animate DOs are mostly marked with the accusative, while the definiteness factor is also of importance. However, the elicitation sessions on Beserman offer the following results: all animacy-based classes (human animate, non-human animate and inanimate) can occur without the accusative. As to the referential properties, native speakers show a strong tendency to mark definite (and attributive-used) DOs with the accusative (or with the possessive), while indefinite specific and non-specific DOs are not marked. Generic DOs are marked with the accusative if they constitute the topic of the sentence (as in 'Potatoes, we dig them in autumn').

Thus, the elicitation sessions show the significance of referential properties of DOs and non-significance of the animacy factor.

This conforms to the well-known generalizations about the use of the accusative in the earlier state of Permic languages. It has been claimed that the accusative goes back to the possessive of 3rd person, which in turn was used a definiteness marker. There is even a point of view that the definiteness function did not arise from the possessive function, but was their original function [Majinskaya 1979; Raun 1988].

The elicitation sessions took place in years 2003–2005 and continued in 2009–2012. During this period of time, the corpus-based methodology could not have been applied until the morphologically-tagged corpus appeared in 2011–2012. The first version of the corpus was developed in the Fieldworks SIL software by Olga Biryuk and contained 33 000 tokens. It was then enlarged by Timofey Arkhangel'skiy and transferred to the online search platform that he developed (see http://beserman.ru/corpus/search/?interface_language=en). The present work is based on the version of the corpus dating December 2017, when it contained 75 000 tokens.

Thus, it is important that the elicitation results appeared earlier and were not influenced by the results obtained by corpus analysis. It should also be emphasized that the corpus materials roughly belong to the same time period as the data collected by elicitation (a number of texts date in 2013–2016, when I did not elicit, but I assume that the language did not change significantly between 2013 and 2016).

4. Animacy vs. definiteness in DOM: results of the corpus study of Beserman Udmurt

4.1. Basic distribution

In order to conduct the corpus study, all the corpus texts have been automatically extracted into Microsoft Excel file and divided into separate sentences (I thank Svetlana Toldova and Dmitriy Gorshkov for completing this work). The resulting file contained 10 539 sentences. Each sentence was then manually annotated based on the DO marking and animacy (if the sentence contained a DO). I differentiated between five DO marking types: 2nd person singular possessive, possessive³, (singular) accusative, plural accusative and no marking.

³ The 2nd person singular possessive is considered separately for the reason that this marker developed a number of specific pragmatic functions, presumably different from the other markers.

In case a sentence included more than one DO, separate lines were created. The sentences were then analyzed and some subtypes were tagged according to specific parameters (see below).

The distribution of animacy classes and DO marking is shown in Table 2.

Table 2. Distribution of DO marking among animacy-based groups of nouns in the Beserman corpus (10 539 sentences, 2187 DOs)

$$\chi^2 = 367.51, df = 8, p < .0001$$

marker	human animate	non-human animate	inanimate	total
2 nd person possessive	4 (2%)	3 (1%)	42 (2.5%)	49
accusative <i>-ez</i>	72 (33%)	83 (26%)	126 (8%)	281
possessive	69 (32%)	87 (27%)	500 (30%)	656
accusative plural <i>-tâ</i>	26 (12%)	53 (17%)	33 (2%)	112
no marking	45 (21%)	95 (30%)	949 (57.5%)	1089
total	216	321	1650	2187

First, it can be observed that the basic rule formulated in [Kondrat'eva 2010] for Standard Udmurt is in part confirmed for Beserman (contrary to the results of elicitation sessions): accusative (singular and plural) is much more frequent for animates than for inanimates. Vice versa, the frequency of non-marked DOs raises with inanimates. Note that this is not a strict grammatical rule (as shown by elicitation sessions), rather a tendency.

The frequency distribution is significant according to the chi-square test. To understand each cell's departure from independence I used the standardized residuals post-hoc test. The results are represented in the mosaic plot in Figure 1.

The Figure 1 shows that both possessive markers do not demonstrate any significant difference with respect to the animacy parameter. By contrast, accusative (singular and plural) DOs show significantly bigger frequencies for animate DOs (human and non-human) and significantly smaller frequencies for inanimates. Conversely, non-marked DOs are significantly more frequent for inanimates and significantly less frequent for animates (human and non-human).

Thus, there are four cells that show significantly low frequencies: accusative singular and plural with inanimates, non-marked DO for animates of both types. Those cells are marked with bold font in the Table 2. In what follows I am going to consider the material that gave the results for these cells, and try to find an explanation for the disbalanced frequencies.

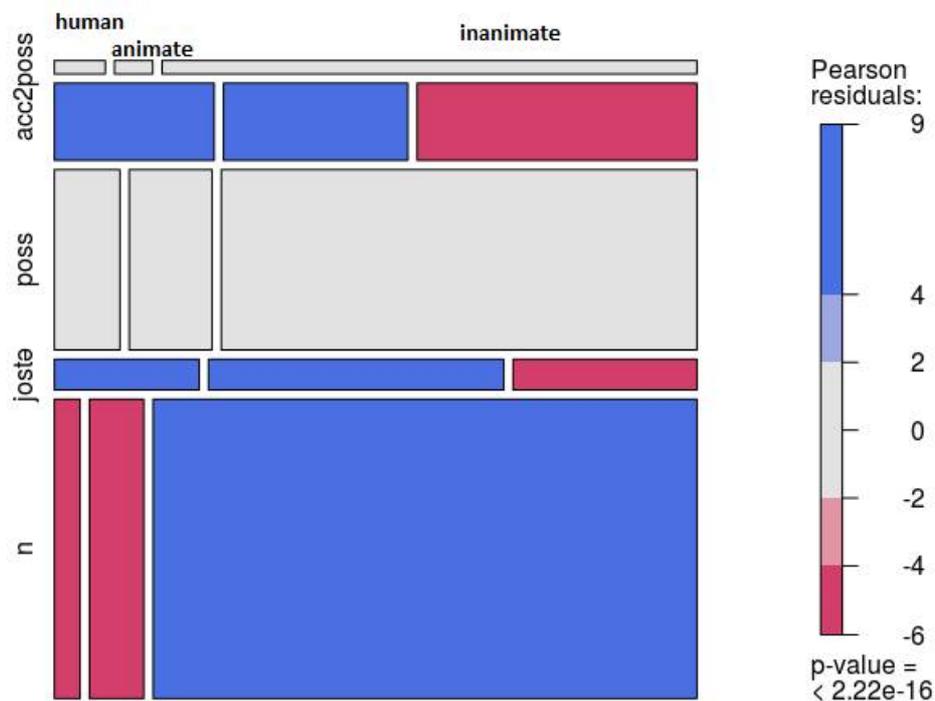


Figure 1. The mosaic plot showing departure of each cell from independence based on Table 2. (Grey is used for non-significant departure, blue for significantly higher frequencies, red for significantly lower frequencies; ACC – accusative, POSS – possessive, N – non-marked).

4.2. Non-marked human animates

The examples with non-marked human animates seem to form a heterogeneous class. At least, they include DOs with different referential properties: there are specific indefinites (8), definites (9), non-specific indefinites (10), generic DOs (11).

(8) *nāl so vaj-i-z, kāk-t-et'i-ze.*
 girl that give-PST-3SG two-OBL-ORD-ACC.P.3(SG)
 'As a second child she bore a girl.' [Corpus]

(9) *prez'ident bārj-em ber-e...*
 president choose-NMLZ behind-ILL
 'After the president has been elected (everything changed).' [Corpus]

(10) *a mužik ton šed'-t-â.*
 and husband you be.found-CAUS-IMP.SG
 '(I'm not married.) And you should find a husband.' [Corpus]

(11) *kāno et'-iš'ko-m, mād-mād-a-mâ š'i-iš'ko-m – ju-iš'ko-m.*
 guest call-PRS-1PL RECP-RECP-LOC/ILL-P.1PL eat-PRS-1PL drink-PRS-1PL
 'We invite guests, eat and drink at each other's houses.' [Corpus]

However, all these examples share one common property: the DO and the verb describe a situation of creating a new object, either physically ('give birth to a child', as in (8)) or socially ('find a husband', 'elect a president', 'invite a guest'). In all those cases the object does not exist as such until the situation described by the verb takes place (the president is not a president until s/he has been elected, as well as the guest is not yet a guest until s/he has been invited).

Thus, all of these situations can be described as "creation of a new object"⁴. Judging from the corpus data, I can conclude that such situations require for DOs to occur in the non-marked form. It could be argued that this rule does not depend on the referential properties of the DO, given that definite DOs are also non-marked (9); however, it is a debatable issue whether such DOs can be analyzed as instances of referential or attributive use [Donnellan 1966; Kripke 1977]. Even if at the given point of time the interlocutors could have the particular president in mind, the noun "president" may be used to point at the specific time period (this clause could be replaced by the phrase "after the elections", without mentioning the president), irregardless of the issue who was president at the current time (see ex. 21 in [Abbott 2011: 62]). Taking this point of view, we need to conclude that the lexical pairs of creation of a new object are likely not to have referential definite DOs at all. However, the verification of this hypothesis needs checking based on elicitation of thoroughly elaborated contextual minimal pairs, which are absent from the corpus.

There are other examples where DOs of this class occur without any marking; these include two particular lexemes 'people' and 'baby' (the non-human-like morphological behavior of these two is quite expected knowing that these lexemes are treated similarly in other languages, e.g. 'baby' in English) and the subclass of non-human mythological characters, as wood-goblin etc.:

- (12) *ken'a-ke* *až'-i* *č'aššja-jân* *č'aššja kuž'o* *kad'* *mar = a.*
 how.many-INDEF see-PST-1SG forest-LOC forest goblin similar what-Q
 'Once, when I was in a forest, I saw a wood-goblin, or someone similar.'
 [Corpus]

⁴ The relevance of such lexical pairs has been first demonstrated for Komi-Zyrian in [Serdobolskaya, Toldova 2013].

With all the other lexical classes, the accusative is present irregardless of the referential properties of the DO, see example (13) with an indefinite non-specific DO.

- (13) *kak = pe č'eber nāl murt-ez, kâšno murt-ez až'-e – fš'o.*
 as = CIT beautiful girl man-ACC woman man-ACC see-PRS.3SG everything
 'Each time he sees a beautiful girl, a beautiful woman – he comes on to her.' (lit. it's everything) [Corpus]

Hence, I can formulate the following rule: human-denoting DOs are mostly marked with the accusative (or the possessive), with the exception of some lexemes (baby, people, mythological characters) and a specific class of lexical pairs of DO + verb (creation of a new object).

This rule works irregardless of the referential properties of the DO; however, the creation of a new object as such presupposes indefiniteness or non-specificity of the referent (see the discussion after the example (9) above). Thus, it can be claimed that the function of the absence of DO marking has narrowed from indefiniteness/non-specificity to situations of creation of a new object.

4.3. Non-marked non-human animates

For non-human animates, the context of creation of a new object also requires the absence of marking:

- (14) <...> *tin' ož' tin', podruža šed'-t-âsa.*
 here so here girlfriend be.found-CAUS-CVB
 '<...> in this way [the rooster] found himself a girlfriend.' [Corpus]

However, there is a large class of uses that falls out of this rule. These examples either include indefinite/non-specific DOs (as predicted by the referential properties rule in the beginning of the section 3) or are characterized by the semantics of a typical situation (see [Kretov 1992] for the relevance of a similar parameter for Russian; its relevance for DOM is shown in [Serdobolskaya, Toldova 2013] based on the data on Komi-Zyrian). This includes pairs of nouns and verbs denoting specific farm activities, as milking cows, feeding the cattle, shepherding cattle etc. In these cases the non-marking variant is often used even if the DO is definite:

- (15) *parš'-jos-tâ š'ud-iš'ko-m, sre skal kâsk-iš'ko-m.*
 pig-PL-ACC.PL feed-PRS-1PL, then cow milk-PRS-1PL
 '(We usually get up in the morning...) feed the pigs, then milk the cows.'
 [Corpus]

The speaker tells about her everyday activities, thus mentioning the cattle that actually belongs to her. It is therefore not aforementioned, but its existence and uniqueness is established due to presupposition accommodation (as the context implies that it belongs to the speaker). Note that the speaker uses the accusative (plural) for the first DO, and no marking for the DO in the second clause, even if they have the same referential properties.

The relevance of the semantic pairs denoting typical situation is confirmed both by chi-square and Fisher's criteria, see Table 3.

Table 3. Distribution of DO marking with typical situations (for non-human DOs)
 $\chi^2 = 3.551$, $df = 1$, $p = .06$; $F = 0.04594$, $p < .05$

typical situations	yes	no
accusative	4 (14%)	6 (46%)
non-marked DOs	25 (86%)	7 (54%)
total	29	13

The Table 3 shows that most DO occurring in typical situation pairs are not marked, while for other cases the distribution is almost fifty-fifty.

Therefore, for non-human animate DOs the following rules are relevant: 1. There is a tendency for definite DOs to take the accusative, and for all other referential types not to take it; 2. The DO + verb pairs that denote creation of a new object have non-marked DOs; 3. The DO + verb pairs that denote typical situations show a strong tendency towards the absence of the accusative.

4.4. Accusative inanimates

As shown in Figure 1, the situation for inanimates is reversed, which in part follows the generalizations for Standard Udmurt. The whole number of accusative inanimates is 126, which makes 8% of all the inanimate DOs in the corpus. 62 of these examples include definite DOs (16) and 3 of them include attributive DOs.

(16) *kâk-na-ze-s* *kut-i* *mon,* *odig-ze* *ki-t'i-z*
 two-COLL-ACC.P.3-PL catch-PST(1SG) I one-ACC.P.3(SG) pyka-PROL-P.3(SG)

kut-i *ðagaũme mešok-ez* *pi* *nu-e* *kâ-t-âš'* *baš'-t-i-d-â!*
 catch-PST-1SG come.on bag-ACC AUTOCIT carry-IMP.PL where-OBL-EL take-PST-2-PL

(The speaker is telling the story of two thieves who tried to take a big bag out of the warehouse.) 'I caught them both, I caught one of them by the hand and said: "Come on, carry the bag back to the place you've taken it from!"'

However, definite DOs can also occur non-marked:

- (17) *mašina* [...] ⁵ *mar kar-e, baš't-o-d,*
 car what do-PRS.3SG take-FUT-2(SG)

uža-l-o-d vāldā, baš't-i molokovoz.
 work-EXP-FUT-2(SG) after.all братъ-PST(1SG) milk.tanker

‘(Then I was given a car, a milk tanker.) What shall you do with a car, if you take you’re going to work on it, I took the milk tanker.’ [Corpus]

Topical generic DOs (59 tokens) may take the accusative:

- (18) *jetân kiž'-āl-i-z-â, jetân-ez âšk-iš'ko-m tin' taž'.*
 flax seed-ITER-PST-3-PL flax-ACC pull-PRS-1PL here so

(The speaker is telling about flax breeding and processing.) ‘We seeded the flax, then we pulled the flax this way.’ [Corpus]

However, they may also be non-marked, as the first DO in the same sentence.

The rest of inanimate accusative DO (outside of the class of definite, attributive and topical generic DOs) makes 32 examples. The question arises, what triggers the presence of the accusative in these cases. It turns out that most of these examples (25, that is 78%) include DOs that contain some material other than the head noun (adjectives, juxtaposed nouns, pronominal modifiers etc.). By contrast, DO that only contain one lexical unit, i.e. the head noun, are mostly not marked. This is shown in Table 4, where I compare the distribution of marking with one-word DOs and DO containing more than just the head.

Table 4. Distribution of DO marking depending on the NP structure⁶
 $\chi^2 = 68.088$, $df = 1$, $p < .001$

inanimate DOs	no marking	accusative
head-only (one-word) DOs	764 (80.5%)	47 (45%)
DO containing modifiers, juxtaposed nouns etc.	185 (19.5%)	58 (55%)
total	949	105

⁵ Unintelligible fragment.

⁶ This table includes all the results on inanimate DOs, irregardless of their referential status, as the annotation based on referential properties was only made for the narrow class of accusative inanimates. We excluded proper nouns from Table 4, since the choice of DO marking with them is subject to lexical restrictions.

The distribution is relevant for the choice of DO marking, as shown by the chi-square test; the post-hoc Pearson residuals test shows that the most significant (departing from the independence hypothesis) are the results in the second column, that is, by accusative DOs. Thus, it may be concluded that the “heaviness” factor is relevant. DOs that contain more than one lexical unit are more likely to be marked with the accusative than one-word DOs. It must be, however, specified that this generalization is very rough: there are components that require the absence of marking, for example cardinals, numeric groups, negative polarity items and indefinite pronouns:

- (19) *a d'en'is so mar-ke kâl so vala = wa?*
 and Denis that what-INDEF language that understand-PRS.3SG = Q
 ‘What about Denis, does he speak any other language?’ [Corpus]

In case of indefinite pronouns, the accusative is often banned by native speakers. As far as numeric groups and cardinals are concerned, I do not have such data, and the examples of definite DOs of these structural types are absent from the corpus. Thus, it is unclear, which factor “wins” in these cases, the definiteness of DO (requiring the accusative) or the syntactic structure factor (requiring no marking).

By contrast, demonstrative pronouns require the accusative:

- (20) *mon so-je, so šâd-ez ž'už'-i, ž'už'-i gine*
 I that-ACC that soup-ACC take.a.gulp-PST(1SG) take.a.gulp-PST(1SG) only

âm dur-â bâdes kwal'ek-ja-nâ kuč'k-i-z...
 mouth edge-P.1SG full tremble-MULT-INF begin-PST-3(SG)
 (The speaker was given a plate of goose soup.) ‘I have just taken a gulp of this soup, and my lips started trembling (it was too hot).’ [Corpus]

Thus, it can be concluded that the referential properties of the DO do play a significant role in the choice of the marking of inanimate DOs. Indefinite and non-specific inanimate DO are always non-marked. By contrast, for definite and attributive DOs the accusative is not obligatory. It is required if the definiteness semantics is “reinforced” by the presence of lexical expressions signaling definiteness, such as demonstrative pronouns.

In other words, the functions of the accusative are narrowed in case of inanimate DO: it is not obligatory with definites (as in case of animate DOs); however, it is obligatory with DOs including demonstrative pronouns (and several lexical classes of proper nouns).

5. Conclusions

The elicitation-based studies show the relevance of referential properties of DOs for the choice of DO marking in Beserman Udmurt. For Standard Udmurt, the animacy factor has been reported to play a more significant role [Kondrat'eva 2010]; in elicited examples from Beserman this factor seems to be much less important. The corpus analysis enables us to make a more precise picture of the interplay of the two factors, as different values of those factors make different impact into the choice of DO encoding. Namely, for human animate DOs the non-marked form is only used for the situations of creation of a new object and for some specific lexemes. Non-human animates are non-marked if they make part of the situation of creation of a new object or typical situations. Otherwise, definite animates always take the accusative, and indefinite, generic and non-specific DOs show fifty-fifty distribution of the marking. As for inanimates, indefinite and non-specific DOs are always non-marked; definites can take the accusative. It is obligatory with NPs including demonstrative pronouns and preferred with “heavier” NPs, that is, NPs including more than the nominal head. However, some modifiers favour the absence of the accusative; these include numeric groups, cardinals, negative polarity items and indefinite pronouns.

Therefore, the factor of referential properties plays an unequal role for each animacy-based class of nouns: for the animate classes, the use of the accusative is expanded onto indefinite/non-specific DOs (excluding some lexical classes); for inanimates, its use is narrowed from definite DOs to the “heavier” DOs. This distribution has been discovered by the means of the corpus analysis, as it enables to make frequency-based judgments. However, it does not offer the possibility to test less frequent types. For instance, the corpus does not have any examples of cardinals or numeric groups in DOs with definite semantics, as well as examples of the referential use inside the groups denoting creation of a new object. These gaps can only be covered by elicitation of thoroughly elaborated contextual minimal pairs. Until such pairs are found and tested, we cannot conclude, which factor “wins” in case of interaction of factors.

Abbreviations

DO – direct object, ACC – accusative, ADD – additive particle, ART.DEF – definite article, AUTOCIT – autocitative, CAUS – causative, CIT – citation marker, COLL – collective numeral, CVB – converb, EL – elative, EXP – expanded stem, FUT – future, ILL – illative, IMP – imperative, INDEF – marker of indefinite pronouns, INF – infinitive, ITER – iterative, LOC – locative, MULT – multiplicative, NMLZ – nominalization, OBL – oblique nominal stem, ORD – ordinal numeral, P.1/2/3SG/PL – possessive markers, PL – plural, PREP – preposition, PROL – prolativ, PRS – present, PST – past, Q – question marker, RECP – reciprocal, SG – singular.

References

- Abbott 2011 — Abbott B. Reference: Foundational issues. Maienborn C., von Stechow P., Portner P. (eds.). *Semantics. An international handbook of natural language meaning*. Vol. 1. Mouton de Gruyter, 2011. P. 49–74.
- Alatyrev 1983 — Alatyrev V.I. *Vydelitel'no-ukazatel'naya kategoriya v udmurtskom yazyke* [Category of emphatic demonstrative in Udmurt]. Izhevsk: Udm. NII istorii, ehkonomiki, literatury i yazyka, 1970.
- Bossong 1985 — Bossong G. *Differentielle Objektmarkierung in den neuirischen Sprachen*. Tübingen: Gunter Narr, 1985.
- Csucs 1990 — Csucz S. *Chrestomathia Votiacica*. Budapest: Tankönyvkiadó, 1990.
- de Swart 2007 — de Swart P. *Cross-linguistic variation in object marking*. Ph.D. thesis. Radboud Universiteit Nijmegen, 2007.
- Donnellan 1966 — Donnellan K.S. 1966. Reference and definite descriptions. *Philosophical Review* 77. P. 281–304.
- Edygarova 2010 — Edygarova S. *Kategoriya posessivnosti v udmurtskom yazyke*. [Category of possession in Udmurt] Tartu: Tartu Ülikooli Kirjastus, 2010.
- Fraurud 2001 — Fraurud K. Possessives with extensive use: A source of definite articles? Baron I., Herslund M., Sørensen F. (eds.). *Dimensions of possession. Typological studies in language*. Amsterdam; Philadelphia: John Benjamins, 2001. Pp. 243–267.
- Kel'makov 1996 — Kel'makov V.K. 1996. *Formy sub'ektivnoj ocenki imen sushchestvitel'nyh v udmurtskom yazyke*. [Substantive forms of subjective evaluation in Udmurt]. *Congressus Octavus Internationalis Fenno-Ugristarum III*. Jyväskylä, 1996. P. 131–134.
- Klumpp 2008 — Klumpp G. *Differentielle Objektmarkierung und Informationsstruktur in Dialekten des Komi*. Hab. diss., München, 2008. Ms.
- Kondrat'eva 2002 — Kondrat'eva N.V. *Vyrazhenie pryamogo ob'ekta v udmurtskom yazyke (v istoriko-sopostavitel'nom plane)* [The expression of direct object in Udmurt (in historic-comparative aspect)]. *Avtoref. dis. ... kand. filol. nauk*. Izhevsk: Udmurtskij gos. un-t, 2002.
- Kondrat'eva 2010 — Kondrat'eva N.V. *Mezhkategorial'nye svyazi v grammatike udmurtskogo yazyka (na materiale padezha pryamogo ob'ekta)* [Intercategorical relations in the grammar of Udmurt: the direct object case]. Izhevsk: Udmurtskij universitet, 2010.
- Kretov 1992 — Kretov A.A. «S'edobnoe-nes'edobnoe», ili kriptoklassy russkikh sushchestvitel'nyh [“Edible/inedible”, or criptoclasses of Russian nouns]. *Lingvistica Silesiana*. 1992. Vol. 14. P. 103–114.
- Kripke 1977 — Kripke S. Speaker's reference and semantic reference. French P.A., Uehling T.E., Wettstein H. (eds.). *Midwest Studies in Philosophy*, vol. II: *Studies in the Philosophy of Language*. Morris, MN: University of Minnesota, 1977. P. 255–276.
- Kuznecova 2012 — Kuznecova A.I. *Kumulyaciya grammaticheskikh znachenij v agglutinativnyh pokazatelyah: dejkticheskie funkcii posessiva v ural'skikh yazykah*. [Cumulation of grammatical meanings in agglutinative markers: deictic functions of possessive forms in Uralic]. Kuznecova A.I. (ed.). *Finno-ugorskie yazyki: Fragmenty grammaticheskogo opisaniya. Formal'nyj i funkcional'nyj podhody*. Moscow: Yazyki slavyanskih kul'tur, 2012. P. 250–252.
- Majtinskaya 1979 — Majtinskaya K.E. *Istoriko-sopostavitel'naya morfologiya finno-ugorskikh yazykov*. [Historic comparative morphology of the Finno-Ugrian languages] Moscow: Nauka, 1979.
- Moravcsik 1978 — Moravcsik E. On the case marking of objects. Greenberg J.H. (ed.). *Universals of human language*. V. 4: *Syntax*. Stanford: Stanford University Press, 1978. P. 249–289.
- Perevoshchikov et al. 1962 — Perevoshchikov P.I., Vahrushev V.M., Alatyrev V.I. (ed.). *Grammatika sovremennogo udmurtskogo yazyka. Fonetika i morfologiya* [A grammar of Modern Udmurt: Phonetics and Morphology]. Izhevsk: Izhevskoe knizhnoe izd-vo, 1962.

- Raun 1988 — Raun A. Proto-Uralic comparative historical morphosyntax. Sinor D. (ed.). *The Uralic languages. Description, history, and foreign influences*. Leiden, New York, København, Köln: Brill, 1988. P. 555–571.
- Schlachter 1960 — Schlachter W. *Studien zum Possessivsuffix des Syrjänschen*. Berlin: Akademie Verlag, 1960.
- Serdobolskaya, Usacheva, Arkhangel'skiy 2019 — Serdobolskaya N., Usacheva M., Arkhangel'skiy T. Grammaticalization of possessive markers in the Beserman dialect of Udmurt. Johanson L., Nevskaya I., Mazzitelli L.F. (eds.). *Linguistic possession. New insights from the languages of Europe and North and Central Asia*. (Studies in Language Companion Series). John Benjamins, 2019. In press.
- Serdobolskaya 2017 — Serdobolskaya N.V. Invariant opredelennosti i differencirovannoe markirovanie pryamogo dopolneniya v besermyanskom udmurtskom [Uniqueness and familiarity in Beserman Udmurt direct object marking]. *Acta Linguistica Petropolitana*. 2017. Vol. 13, No. 3. Saint Petersburg: Nauka. P. 76–122.
- Serdobolskaya, Toldova 2013 — Serdobolskaya N.V., Toldova S.Ju. Leksicheskie svojstva glagola i oformlenie pryamogo dopolneniya v komi-zyryanskom yazyke (pechorskij dialekt). [Lexical properties of verb and direct object marking in Komi-Zyrian (the dialect of Pechora)]. *Lingvisticheskij bespredel-2. Sbornik nauchnyh trudov k 80-letiyu A.I. Kuznecovoj*. Moscow: Izdatel'stvo MGU, 2013. P. 164–175.
- Suihkonen 2005 — Suihkonen P. On the categories and functions developed from the possessive and deictic suffixes in Udmurt. Hasselblatt C., Koponen E., Widmer A. (eds). *Lihkkun lehkos! Beiträge zur Finnougristik aus Anlaß des sechzigsten Geburtstages von Hans-Hermann Bartens*. Wiesbaden: Harrassowitz Verlag, 2005. P. 401–432.
- Winkler 2011 — Winkler E. *Udmurtische Grammatik*. Wiesbaden: Harrassowitz Verlag, 2011.
- Witzlack-Makarevich 2017 — Witzlack-Makarevich A., Seržant I.A. Differential argument marking: Patterns of variation. Seržant I.A., Witzlack-Makarevich A. (eds.). *The diachronic typology of differential argument marking*. [Studies in diversity linguistics]. Berlin: Language Science Press, 2017.

Статья поступила в редакцию 21.11.2018

The article was received on 21.11.2018

Наталья Вадимовна Сердобольская

кандидат филологических наук; старший научный сотрудник, Институт языкознания РАН; старший научный сотрудник, Государственный институт русского языка им. А. С. Пушкина

Natalia V. Serdobolskaya

Ph.D.; senior researcher, Institute of Linguistics of the Russian Academy of Sciences; senior researcher, Pushkin State Russian Language Institute

serdobolskaya@gmail.com