NATALIA SERDOBOLSKAYA (Moscow)

A CORPUS ANALYSIS OF DIFFERENTIAL OBJECT MARKING IN BESERMAN UDMURT*

Abstract. This paper is focused on differential object marking in Beserman Udmurt. Elicitation sessions show the relevance of referential properties of the DO for the choice between accusative and zero marking; however, for Standard Udmurt animacy has been claimed to play a more significant role. To establish the exact ranking of each parameter, we conducted an analysis of corpus data involving 10 539 sentences with 2187 DOs. With human animates, the unmarked DO is only possible with some lexical classes; with non-human animates, referential properties play the central role, as predicted by elicitation; with inanimates, the frequency of the accusative is higher among DOs with certain types of modifiers.

Keywords: Udmurt, differential object marking, animacy, corpus analysis, referential properties.

1. Introduction

It is well-known that Uralic languages exhibit the phenomenon of differential object marking (DOM) described in (de Swart 2007; Iemmolo 2011; Witzlack-Makarevich, Seržant 2017). This term is used to describe constructions with transitive verbs that encode their objects in two or more different ways, see (1). DOM in Uralic languages and varieties has been first described in Wickman 1955, who gives a detailed overview of DOM in 13 Uralic varieties and summarizes the information on semantics of DOM found in different sources and texts available at the time.

Standard Udmurt:

(1) *mon uram-jš' l'ek* ***puni-jez*** *~ čeber* ***korka*** *adž'-i*
I street-EL savage dog-ACC pretty house see-PST(1SG)
'I saw a savage dog ~ a pretty house in the street' (Winkler 2011 : 46)[1]

Modern grammars of Uralic languages typically describe the distribution of DO markers, providing a list of factors influencing the choice of marking

* This project 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, and the participants of the conferences "Language, History and Culture of the Beserman: state-of-the-art and perspectives", "Typology of morphosyntactic parameters 2018" for the fruitful discussion.

[1] I retain the author's transcription, however, changing the glosses according to the rules used in the present paper for Beserman.

together with the relevant lexical restrictions. However, the interaction between different factors is often not examined: for example, the Udmurt grammar (Грамматика современного удмуртского языка 1962 : 94; Грамматика современного удмуртского языка 1970 : 168—169) claims that the most important factors are definiteness and membership of specific lexical classes (physiological and psychological verbs). Winkler (2001) states that definite DOs are most often marked with the accusative, in contrast to indefinite DOs; then, he mentions resultativity, animacy and focus as important factors (animate DOs tend to occur with the accusative, as well as DOs in resultative contexts and focused DOs). Such rules adequately describe cases where both factors work in the same direction. For example, if the DO is both definite and animate or if it is definite and the verb is psychological, the DO is expected to be marked. However, if the two rules make different predictions (e.g. the DO is animate and indefinite; the verb is psychological and the DO is indefinite), the resulting marking is unclear: Does one factor outrank the other, or is there free variation? In the latter case, which outcome is more frequent?

This study is aimed at answering these questions for Beserman Udmurt. Based on the data of the Beserman corpus (http://beserman.ru/corpus/search/?interface_language=en; ca. 75 000 tokens), I analyze the statistical distribution of DO markers for various lexical classes of nouns and identify the exact weight of the factors of animacy and referential properties of the DO in the choice of DO encoding.

This paper is structured as follows. In section 2 I provide the background on the studies of DOM in Beserman Udmurt. Section 3 is focused on the opposition of the two factors, animacy and referential properties of the DO. In section 4 I consider the results of the quantitative study of corpus data.

2. Background on DOM in Beserman Udmurt

Cross-linguistically, differential object marking (DOM) has been viewed both as a separate phenomenon (de Swart 2007) and as part of phenomena labelled as differential argument marking (Witzlack-Makarevich, Seržant 2017), together with differential subject marking. Typological studies show that in a large number of languages DOM is determined by prominence scales, such as animacy scale (human > non-human animate > inanimate) and definiteness scale (pronoun > proper name > definite > indefinite specific > non-specific). Aissen (2003) proposes to model DOM in a figure representing interaction of animacy and definiteness. Each combination of the parameters' values is regarded separately (e.g. human indefinite specific and non-human indefinite specific); however, for each combination of the two parameters the result can be predicted based on other combinations and on above-mentioned scales (e.g. if human indefinite specific DOs take the case marker, then human definite DOs do so, as well). I will provide a detailed overview of DOM in Beserman showing the relevance of lexical classes for a full account of DOM. The Beserman DOM is determined by a number of lexical parameters, which interfere with prominence scales in a nontrivial way, namely, they are only relevant for specific combinations of animacy and definiteness.

Witzlack-Makarevich and Seržant (2017) classify the systems of DOM based, among other parameters, on morphological markedness of the DO. The systems with unmarked DOs are called asymmetrical, as they involve

the privative opposition of presence vs. absence of morphological encoding. Permian languages possess a threefold variant of DOM, see Table 1 (with the exception of some Komi dialects offering four variants of marking, which include dative case, see Klumpp 2014), where the unmarked variant (2) is opposed to two DO markers: the simple accusative case marker (3) and the possessive accusative, inflected for person and number (4). In (2)—(4), this distinction is illustrated by examples from Beserman Udmurt.

(2) *Uj-ə̑n nə̑l-ə̑z* ***gur*** *est-i-z*
night-LOC girl-P.3(SG) stove heat-PST-3(SG)
'At night the girl heated t h e s t o v e' (Corpus)

(3) *ʒ̑ić́ə̑ aʒ̑-i-z kə̑š́ pu-**ez***[2]
fox see-PST-3(SG) birch-**ACC**
'(The wolf carried a birch tree to make a new shaft for the cart) The fox saw t h e b i r c h t r e e and scolded the wolf' (Corpus)

(4) *Nu nə̑l-**de** ta-t-ə̑š́ ć́aššа-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 y o u r d a u g h t e r away to the forest hut' (Corpus)

Table 1

The paradigm of the possessive markers in Beserman Udmurt

Person, number	DO set	Non-DO set: inalienable	Non-DO set: alienable
P.1SG	*-me*	*-(j)ə̑/-m*	*(j)e*
P.2SG	*-de/-te*	*-(j)ə̑d/-d*	*(j)ed*
P.3SG	*-ze/-se*	*-(j)ə̑z/-z*	*(j)ez*
P.1PL	*-mes*	*-(ə̑)mə̑*	
P.2PL	*-des/-tes*	*-(ə̑)də̑/-tə̑*	
P.3PL	*-zes/-ses*	*-(ə̑)zə̑/-sə̑*	

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

The possessive accusative markers are part of a large paradigm of possessive markers that differentiate the possessor's person (e.g. 2nd person *nəl-əd* 'your daughter' vs. 3rd *nəl-əz* 'his/her daughter'), number, syntactic position of the head (DO vs. other positions) and (in)alienability. As shown in Table 1, the DO set is differentiated from other markers both in form and function (unlike the non-DO set, the DO set does not distinguish between alienable and inalienable possession).

Plural DOs have a special accusative marker *-tə* (non-possessive), see the first DO in (5), or take possessive accusative markers. The zero marking is not available for DOs with the plural marker; note, however, that Udmurt exhibits optional plurality marking (in terms of Corbett 2000 : 70) and singular unmarked DOs can refer to multiple entities (the third noun in (5)).

(5) *So vel'-e ć́ašja-je ć́əž-**jos-tə̑** / *ć́əž-**jos** / **ć́əž** ə̑b-ə̑l-ə̑nə̑*
that go-PRS.3SG forest-ILL duck-PL-**ACC.PL** duck-**PL** / duck shoot-ITER-INF
'He often goes to the forest to shoot ducks' (Elicited)

[2] In Beserman, *pu* means 'tree', while *kəš́* is not attested as a separate lexeme.

The marker *-tə* is in complimentary distribution with the possessive, cf. (5) and (6).

(6) *Vaśa pə̑d-jos-**se** kott-i-z*
Vasya leg-PL-**ACC.P.3(SG)** wet-PST-3(SG)
'Vasya has drenched his legs' (Elicited)

As shown in (5), DOs referring to multiple entities can occur without any overt plural morphology and without the accusative, while plural DOs must take the accusative *-tə* or the possessive markers. Hence, unlike singular DOs, plural DOs have a binary opposition (non-possessive vs. possessive marking), which must be analyzed together with the distribution of the nominal plurality marker. Therefore, plural DOs will not be considered in detail; however, I am providing the statistical results for the sake of the future research agenda.

Discussion of DO markers requires a few remarks on the use of the possessive markers. Besides from marking the possessive relation, as in (4) and (6), the possessive markers are widely used as referential devices (Suihkonen 2005; Winkler 2011), as in (7), where the DO 'dust' is aforementioned in the discourse; the possessive here acts as a kind of definiteness marker.

(7) *Val=no kopoť-**se** šə̑j-e=ke kə̑z-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) t h e d u s t, it coughs' (Corpus)

In such cases a possessive relation can hardly be observed, and the possessive suffixes are obviously employed as referential markers. This has led some researchers to consider the hypothesis that possessive markers in Udmurt are article-like (Fraurud 2001; É. Kiss, Tánczos 2018). However, this hypothesis is challenged by 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 in Beserman 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 and in non-finite clauses (see Алатырев 1970; Едыгарова 2010; Кельмаков 1996, Кузнецова 2012; Suihkonen 2005; Winkler 2011).

In addition, in Сердобольская 2017 and in Serdobolskaya, Usacheva, Arkhangelskiy 2019 are identified the following functions: definiteness by bridging, 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 into the discourse. Serdobolskaya, Usacheva and Arkhangelskiy (2019) argue for the analysis of possessive markers in terms of pragmaticization: the possessive markers are used in pragmatic functions.

To sum up, the following markers of DO are available in Beserman Udmurt:

- no marking,
- accusative *-ez* for singulars and *-tə* for plurals,
- possessive markers.

The present study is mostly focused on the distribution of the unmarked variant and the accusative. The detailed analysis of possessives with DO is given in Сердобольская 2017.

3. Animacy vs. definiteness in Udmurt DOM: previous studies

In this section, I provide an overview of previous studies of the distribution of the accusative vs. unmarked DO in Standard Udmurt and in Beserman.

For Standard Udmurt, reference grammars claim that definiteness is the most important factor determining the choice between the unmarked DO and the accusative (Грамматика современного удмуртского языка 1962 : 93; Csúcs 1990 : 34; Winkler 2011 : 20; É. Kiss, Tánczos 2018; see also the special study by Kondraťjeva (Кондратьева 2002)). Winkler lists three main factors that influence the choice of the DO marker: definiteness, quantification and animacy: "Der Nominativ (auch im Plural) ist Objektkasus, und zwar indefiniter direkter Objekte (AKK im Falle definiter direkter Objekte [–––] Der Akkusativ ist auch der Kasus des Totalobjekts [–––] Auch Belebtheit spielt hier eine Rolle" (Winkler 2011 : 44, 46). For the closely related Komi languages animacy is reported to be among the most important factors (see Klumpp 2008; 2014 : 421—427; also see the discussion of other important factors in Komi and the survey of the relevant literature in Klumpp 2008).

Kondraťjeva points out that there are a number of lexical and grammatical factors that can be more important than definiteness, including animacy, totality/partiality of the DO, resultative/non-resultative interpretation of the situation and a nuber of lexical factors (see Кондратьева 2002).

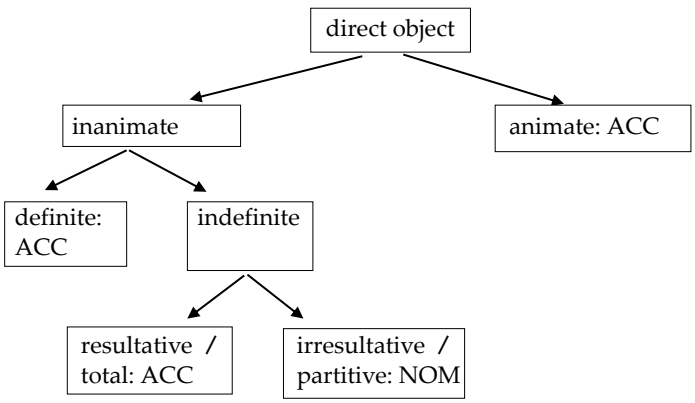

Figure 1. **The distribution of accusative and nominative with DOs in Standard Udmurt (Кондратьева 2010 : 135)**

It means that the DO marker can be unambiguously predicted for each combination of factors' values. However, Kondraťjeva (Кондратьева 2011 : 42) also describes an apparent exception based on information structure: a topical DO is likely to be marked with the accusative even if Figure 1 predicts zero marking to be chosen.

Winkler (2011 : 47) and Kondraťeva (Кондратьева 2000; 2011 : 55) also describe some purely lexical restrictions. For example, nouns denoting intervals and the first object of the verb *kožanə* 'consider (somebody to be something)' usually take the accusative, as well as the DOs of mental verbs and evaluative predicates; abstract nouns and DOs in combinations with verbs denoting profession are preferably not marked.[3]

[3] I did not test these lexical restrictions for Beserman Udmurt, since combinations of such DO + verb are very infrequent in the corpus.

The first study to identify the factors influencing the DO marking in Beserman was Сердобольская, Толдова 2012, based on fieldwork conducted in 2003—2005 in the village of Šamardan. This paper describes that the referential properties of DOs and the information structure of the sentence as the most significant factors, while animacy and quantification factors are shown to play a minor role in a limited number of contexts.

Specifically, elicitation shows that all animacy-based classes (human animate, non-human animate and inanimate) can occur non-marked:

(8) *So bɔ̂gat-e ďeš́ **piń́al** biń-ɔ̂nɔ̂*
that can-PRS.3SG good baby swaddle-INF[4]
'She knows very well how to swaddle a baby' (Elicited)

(9) *So **skal** baš́t-ɔ̂nɔ̂ med-e*
that cow buy-INF want-PRS.3SG
'He plans to buy a cow' (Elicited)

(10) *So **ľɔ̂pet** tupa-ť-ťa*
that roof fit-CAUS-MULT(PRS.3SG)
'He is repairing the roof [of his house]' (Elicited)

As for referential properties, native speakers show a strong tendency to mark definite, universal and attributive 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 only if they constitute the topic of the sentence (as in 'Potatoes, we dig them in autumn'). The relevance of referential properties of the DO for the choice of the encoding is illustrated in (11)—(17):

• Definite: referential use in terms of Donnellan 1966, cf. (12):

(11) *peľ-o **kuč́ɔ̂ran-ez** korka-š́ uj-i-z-ɔ̂*
ear-ATTR hawk-ACC house-EL turn.out-PST-3-PL
'(They approached the house and saw an owl with its young) They turned t h e o w l out of the house' (Corpus)

• Definite: attributive use (in terms of Donnellan 1966: in the case of the referential use the speaker is able to identify the referent of the description, while in the case of the attributive use the speaker is describing the referent on the basis of his/her properties or situation s/he is part of, e.g. *the best student in the class*, *the one who is going to leave the class last*, *Smith's murderer*):

(12) *Mon kos-o **ber** **pot-iš́** **pi-jez** tɔ̂l kɔ̂s-ɔ̂nɔ̂*
I order-FUT(1SG) behind go.out-PTCP.ACT boy-ACC light turn.out-INF
'I want (lit. I am ordering) the last one to leave the classroom to turn out the lights' (Elicited)

[4] In (8)—(9) and several other examples the DO belongs to a non-finite clause. Non-finiteness is an additional parameter that influences the choice of the DO marking in Finno-Ugric languages, see Сердобольская, Толдова 2012 for Mari. However, according to my data, this parameter is not relevant for Beserman Udmurt. To test the importance of this factor, I have tagged 110 arbitrary chosen clauses from the corpus (my search was restricted to non-animate DOs without any modifiers), 81 of which are finite and 29 non-finite and analyzed the DO marking in these sentences with the chi-square. The obtained results show that (non-)finiteness is irrelevant at $p = 0.6149$ ($\chi^2 = 1.8$, $df = 3$).

- Indefinite specific:

(13) *odig starik mən-e bazar-e vuza-nə, vuza-šk-ənə* [---]
one old.man go-PRS.3SG fair-ILL sell-INF sell-DETR-INF
kureg *baš-t-e, **ə̑ž** baš̀t-e*
hen take-PRS.3SG sheep take-PRS.3SG
'(The beginning of a tale.) Once an old man went to the fair to trade, he took a h e n, a s h e e p' (Corpus)

- Non-specific:

(14) *so **skal** baš̀t-ənə med-e*
that cow buy-INF want-PRS.3SG
'He plans to buy a cow.' (Elicited)

- Universal:

(15) *ton l'uka vič̀ak kwaka-**os-tə̑**, a mon l'uka-l-o*
you assemble(IMP.SG) all bird-PL-ACC.PL and I assemble-EXP-FUT-1SG
*vič̀ak tak život-**jos-tə̑***
all so(RUS) cattle-PL-ACC.PL
'You should assemble all the birds, and I will assemble all the animals' (Corpus)

- Generic topical:

(16) *mi val'ˊ-**os-ez** baš̀t-ənsa końušńa-je końux-lə̑ zdatˊ*
we horse-PL-ACC take-CVB stable-ILL stableman-DAT return(RUS)
*kar-əl-i-m val'ˊ-**os-tə̑***
do-ITER-PST-1PL horse-PL-ACC.PL
'(The fragment is about farm horses and the way they were looked after in the speaker's youth. Context: We waited for the horses to be returned to the stables) We took the horses to the stables to return them to the stablemen' (Corpus)

- Generic non-topical:

(17) *nu tiń šeďˊ-i-z kolxoz, **skal** vožˊma val*
well here find.oneself-PST-3(SG) kolkhoz cow pasture(PRS.3SG) be.PST
'[My husband] entered the kolkhoz, he pastured the cows' (Corpus)

However, Serdobolskaya and Toldova (see Сердобольская, Толдова 2012) are also aware of the exceptions to these rules: as stated in Кондратьева 2011, the information structure and the discourse properties of the DO sometimes override them, making a DO marked in cases it is expected to be non-marked, and vice versa (see Сердобольская, Толдова 2017 for details; see also Klumpp 2008 for information structure and discourse factors in Komi DO marking).

The prevalence of the definiteness factor conforms to the well-known reconstructions of the use of the accusative in earlier stages of the Permian languages. It has been claimed that the accusative goes back to the 3rd person possessive, which in turn was used a definiteness marker (Майтинская 1979 : 102; Rédei 1988 : 382—383). Some scholars even maintain that definiteness was the original function of the so-called possessive markers and is not secondary to possession (Майтинская 1979; Raun 1988). See also É. Kiss, Tánczos 2018 on the topicality-based hypothesis on the earlier use of possessive markers.

The elicitation sessions took place in 2003—2005, while the present study is based on the Beserman corpus containing oral texts collected from 2003 to 2017 (http://beserman.ru/corpus/search/?interface_language=en). The difference in our results is, thus, more likely to be explained by the change of methodology, rather than by language evolution (however, the influence of the latter factor cannot be completely rejected).

The size of the corpus and the possibility to use the morphologically-based search enables us to easily calculate the relevance of the above-mentioned factors for Beserman and test the results obtained for Standard Udmurt and for Beserman using statistical criteria.

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

4.1. Basic distribution

This study is based on the version of the Beserman corpus dating December 2017, when it contained 75 000 tokens. Before conducting the study, all corpus texts were automatically extracted into a spreadsheet file and split into separate sentences (I am grateful to Dmitriy Gorshkov for performing this conversion). 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). The annotation distinguishes five DO marking types:

- 2nd person singular possessive,
- possessive,[5]
- singular accusative in *-ez*,
- plural accusative in -(*jos-*)*tə*,[6]
- no marking.

Three classes of animacy are distinguished:

- human animates,
- non-human animates,
- inanimates.

In case a sentence included more than one DO, separate lines were created manually. After the first calculations the analysis was refined for some lexical classes using other semantic parameters discussed in this and the following sections.

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

It can be observed that the basic rule formulated in Кондратьева 2010 for Standard Udmurt is partly confirmed for Beserman (contrary to the results of elicitation): accusative is much more frequent for animates (about 30% for

[5] 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. Namely, in Udmurt, as well as in other Uralic languages, this marker is used for ethical reference to the addressee (Künnap 2006; Кузнецова 2012). For Komi languages, it has been argued that the markers of 2nd person singular carry an additional component of meaning termed as subjective or intimate relation of the referent, see the interpretation of Komi-Zyrian suffixes in Schlachter 1960; Baker 1986.

[6] There are three examples in *-jos-ez*.

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
2nd person possessive	4 (2%)	3 (1%)	42 (2.5%)	49
possessive	69 (32%)	87 (27%)	500 (30%)	656
accusative singular *-ez*	72 (33%)	83 (26%)	**126** (8%)	281
accusative plural *-tə*	26 (12%)	53 (17%)	**33** (2%)	112
no marking	**45** (21%)	**95** (30%)	949 (57.5%)	1089
Total	216	321	1650	2187

accusative singular) than for inanimates (only 8%).[7] Conversely, the frequency of unmarked DOs raises with inanimates: it is 57.5%, while for animates it is 21—30%. Note that in both cases it is not a strict grammatical rule (as shown by elicitation sessions), but rather a tendency showing that there are discourse factors that may override the basic rules (see Сердобольская, Толдова 2017).

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 2.

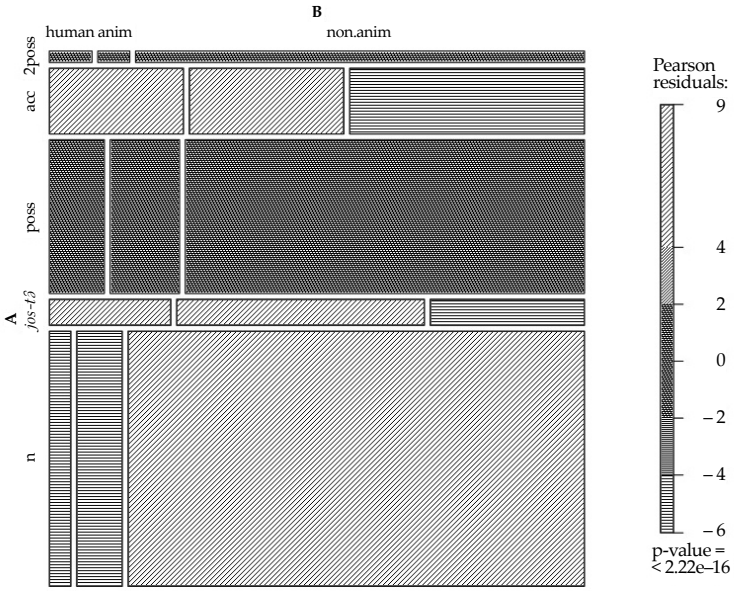

Figure 2.[8] **The mosaic plot showing departure of each cell from independence based on Table 2.**

[7] The accusative plural is much more frequently used with animates (12 and 17%) than with inanimates (only 2%), the major tendency being for inanimate multiple entities to occur unmarked (both for case and number). This conforms to the basic rules of optional plurality marking in Beserman formulated in Ишматова, Черниговская 2012.

[8] Intersecting shading is used for non-significant departure, diagonal shading for significantly higher frequencies, horizontal shading for significantly lower frequencies; n — non-marked, *jos-tə* — accusative plural, poss — possessive, acc — accusative, 2poss — 2nd person possessive. The width of each cell is in direct proportion to the percentage within the row of Table 2; the height of each cell is proportional to the percentage within the column.

Figure 2 shows that neither type of possessive markers demonstrates any significant difference in distribution with respect to the animacy parameter (both lines are shaded with intersecting lines). By contrast, accusative (singular and plural) DOs show significantly higher frequencies for animate DOs (the cells for humans and non-humans are diagonal-shaded) and significantly smaller frequencies for non-animates (horizontal shading). Conversely, unmarked DOs are significantly more frequent for non-animates (diagonal shading) and significantly less frequent for animates (horizontal shading for humans and non-humans).

Therefore, animacy seems to play a major role in the choice between accusative vs. absence of marking. Within the animate class, the distinction between human and non-human referents could also conceivably play a role. To test this, the chi-square test was applied to the first two columns of Table 2. The result is relevant ($\chi^2 = 9.93$, df = 4, p = 0.0416); however, the chi-square value is only a little higher than the distribution value (9.49). The standardized residuals test shows that the departure from the expected values is higher for the last row, i.e. for the absence of marking (–1.51 for humans and +1.24 for non-humans), while the first row (accusative) does not show significant departure. That means that zero marking is differently distributed for human and non-human animate DOs; the use of the accusative, obviously, follows the same pattern. This is confirmed by the study of concrete examples with unmarked human and non-human DOs, see below.

Thus, the following major tendencies are confirmed: animate DOs tend to take the accusative (singular or plural), while non-animate DOs tend to be unmarked. These results are consistent with the rule of animacy given in Кондратьева 2010 (and seem to conflict with the results of elicitation sessions given in Сердобольская, Толдова 2012); however, the rule is not observed strictly. I assume that the rarer patterns (unmarked animates and accusative inanimates) should uncover other semantic (or lexical) factors underlying the mechanism of DOM in Beserman.

Let us now focus on the rarer patterns. 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 boldface in Table 2. In what follows I am going to consider the data that gave the results for these cells, and try to find an explanation for the difference in frequencies. Different animacy types are going to be considered in separate sections.

4.2. Human animate DOs

At first sight, the examples with non-marked human animates (45 instances) seem to form a heterogenous class. At least, they include DOs with different referential properties: there are specific indefinites (18), definites (19), non-specific indefinites (20), generic DOs (21).

(18) ***nə̑l*** *so vaj-i-z,* *kə̑k-t-eti̯-ze*
girl that give-PST-3(SG) two-OBL-ORD-ACC.P.3(SG)
'As a second child she gave birth to a girl.' (Corpus)

(19) ***preźident*** *bə̑rj-em* *ber-e*
president choose-NMLZ behind-ILL
'After the president has been elected (everything changed)' (Corpus)

(20) *a* ***mužik*** *ton šeď-t-ə̂*
and husband you be.found-CAUS-IMP.SG
'(I'm not married) And you should find a husband' (Corpus)

(21) ***kə̂no*** *eť-iśko-m, mə̂d-mə̂d-a-mə̂ śi-iśko-m — ju-iśko-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, many of 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 (18)) or socially ('find a husband', 'elect a president', 'invite a guest'). In the latter set of cases the object mentioned (president, husband, guest) does exist, but it does not fit the description until the situation described by the verb takes place: the president is not a president until s/he has been elected,[9] as well as the guest is not yet a guest until s/he has been invited, and the husband is not yet a husband until the marriage takes place.

Thus, all of these situations can be described as "creation of a new object".[10] While this rule is semantic, it is strongly bound to individual DO—verb pairs. It does not depend on DOs only, because the nouns 'child', 'husband' 'guest', 'president', 'deputee' etc. can largely be used in situations when the discussed object already exists (e.g. *I asked the guests to sit down*). It cannot be attributed to verbs either — probably the only good candidate for such an interpretation is 'to give birth to', but it can also be used with an aforementioned object: *When did she give birth to her son? — In 1996*. What is relevant for such cases is the exact pair of two lexemes, the verb and the DO, and the wider context.

Judging from the corpus data, it can be concluded that DOs in such context are required to occur in the unmarked form. Among the 45 unmarked human animate DOs, there are 29 examples that I classify in this way, while no such DO—verb pairs with accusative are attested.

It could be argued that this rule does not depend on the referential properties of the DO, given that definite (19) and generic (21) DOs are also unmarked, as well as indefinites. However, there is only one example of a definite DO (19), and it is a debatable issue whether it is used to refer to a particular individual or to the event of president's election. In the latter case it can be analyzed as an example of semantic incorporation (as in typical activity pairs discussed in 4.3.2), see Stvan 2009; de Swart, Zwarts 2009 and, hence, ruled out of the set of definite DOs. Due to the lack of data, these two hypotheses cannot be verified on corpus data alone; elicitation of thoroughly elaborated contextual minimal pairs is required. The rest of the 29 unmarked human animates include indefinite DOs, specific and non-specific.

Hence, I conclude that there is another semantic parameter interfering with the choice of DO marking in Beserman Udmurt: in contexts involving creation of a new object (physical or social) even human animate DOs occur unmarked.

There are 16 other examples of non-marked human animate DOs, all of which can be clearly described in terms of idiosyncratic lexical restrictions. First, these include two lexemes *kalə̂k* 'people' and *piń́al* 'baby':

[9] This sentence refers to the first elections of the president of the Udmurt republic.
[10] The relevance of similar lexical pairs has been demonstrated for Komi-Zyrian in Сердобольская, Толдова 2013 and for Hungarian in Piñón 2006.

(22) ***kalə̑k*** *vaj-ə̑nə̑=pe ot-əś́ lə̑kt-o-z mašina*
people give-INF-CIT that.OBL-EL come-FUT-3(SG) car
'(We worked all day long gathering the hay.) It was planned that a car would come to take people away from there.' (lit. a car will come, saying, to take people from there) (Corpus)

(23) *nu* ***pińal****=to ug*[11] *kel'-t-ə̑=ńi mon*
well baby=PTCL NEG.SG remain-CAUS-SG=already I
'(Now you're saying the same thing, what shall I do with the baby?) But I won't leave the baby, will I?' (Corpus)

These two lexemes can occur without the accusative (not necessarily, though) even in case they are definite, as in (22) and (23).[12] The non-human-like morphological behavior of these two nominals can be expected knowing that similar lexemes can be treated as inanimate in other languages, e.g. 'baby' in English (which can be referred by the pronoun 'it', unlike other nouns denoting humans).

Second, there is a class of human-like mythological characters, such as 'wood-goblin', etc.:

(24) *keńa-ke aǯ́-i čaššja-jə̑n* ***čaššja kuǯ́o*** *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 something similar' (Corpus)

(25) *a soldat so-je* [---] *so-len kə̑nar-ez vań=no*
and soldier that-ACC that-GEN1 strength-P.3(SG) COP-ADD
babajaga *zək kut-i-z, paldurt-i-z*
Baba.Yaga big grab-PST-3(SG) push.aside-PST-3(SG)
'(The soldier and Peter came into Baba-Yaga's house, but she does not want to give them dinner) And the soldier her [---] the soldier is strong, he grabbed Baba-Yaga[13] violently and pushed her aside' (Corpus)

This group includes the lexeme *vožo* denoting people who go costumed from house to house on Chrismas Eve (Russian *ряженый*).

As well as with the lexemes 'people' and 'child', the unmarked variant is even found in the context of definite DOs, as in (25). However, those lexemes are not necessarily unmarked (as well as 'people' and 'child'), see (26) with the accusative (the same context as in (25)).

(26) *soldat* ***babajag-ez*** *kut-i-z, ǯ́ut-i-z, muket aǯ́-e*
soldier Baba.Yaga-ACC grab-PST-3(SG) lift-PST-3(SG) other place-ILL
puk-t-i-z
sit-CAUS-PST-3(SG)
'The soldier grabbed Baba-Yaga, lifted her up and put her into another place' (Corpus)

Thus, human-like mythological characters and the lexemes 'people' and 'child' show free variation in DO marking.

[11] The form *ug* in Udmurt is used both for the 1st and the 3rd person singular in present and future. Therefore, I do not specify the person and tense in the glossing.
[12] The same variation is observed in Komi with *kaga* 'baby, child', see Klumpp 2014.
[13] Baba-Yaga is a Russian folklore mean character.

With all the other lexical classes, the accusative is present regardless of the referential properties of the DO, see example (27) with an indefinite non-specific DO and (28) with an indefinite specific DO.

(27) *kak=pe čeber nəl murt-**ez**, kəšno murt-**ez** aǯ́-e —*
as=CIT beautiful girl human-ACC woman human-ACC see-PRS.3SG
fšo
everything
'Each time he sees a beautiful girl, a beautiful woman — he comes on to her (lit. it's everything)' (Corpus)

(28) *jara-t-i l'egit' pi-***ez** *ta-t-əś́, ta šamardan-əś́=ik*
please-CAUS-PST(1SG) young son-ACC this-OBL-EL this Šamardan-EL=EMPH
'(I was born in Šamardan, I lived here and married in Šamardan) I fell in love with a young man from here, from the same Šamardan' (Corpus)

Hence, the generalization for unmarked human animate DOs is the following: a DO is unmarked either if it occurs in a lexical pair meaning 'creation of a new object', or if it is a lexeme *kalək* 'people', *pińal* 'child', or if it belongs to a class of human-like mythological characters (irregardless of the referential properties of these lexemes). This can be generalized in the following rule:

(a) human-denoting DOs must be marked with the accusative (or the possessive), with the exception of specific lexical DO—verb pairs (denoting creation of a new object).
NB. Some lexemes are classified by native speakers differently, along with humans or non-humans, including 'baby', 'people', human-like mythological characters.

This rule works irregardless of the referential properties of the DO; however, it is noteworthy that the creation of a new object as such most often presupposes indefiniteness, non-specificity or generic status of the referent (see the discussion after (19) above[14]). If we assume that in Permian the accusative was consistently used to mark definiteness (since it emerged from the possessive marker that could encode definiteness, along with possessive function, see Майтинская 1979; Raun 1988, Décsy 1990), it can be claimed that in case of human-denoting DOs the function of the absence of DO marking has narrowed from indefiniteness/non-specificity to situations of creation of a new object. This is illustrated in figure 3. It is noteworthy that in closely related Komi languages human animates also can occur non-marked; however, the conditions of the omission of the marking are not the same as in Beserman, see Klumpp 2008.

If further investigations reveal unmarked human definites in Beserman, it will mean that the function of the unmarked constructions has then been reanalyzed as a property of lexical classes, whereafter it expanded onto contexts of definiteness, as in (19). At this point, we can only argue for the process of narrowing of function to have taken place, see figure 3.

Indeed, the situation of creation of a new object does not presuppose indefiniteness or non-specificity. A lot of counterexamples can be imagined with an aforementioned or situationally given DO, e.g. (talking to a married couple) "*Where did you meet your husband*?". However, indefinite specific/non-specific uses, apparently, cover the absolute majority of contexts where such

[14] Definite DOs belonging to such lexical pairs are not attested in the corpus.

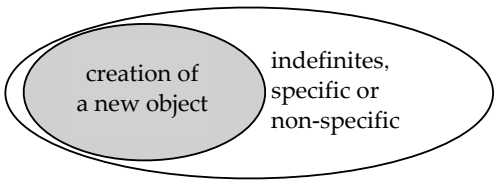

Figure 3. **The change in meaning of non-marked human-denoting DOs.**

pairs are attested (there is only one instance that could be classified as a definite DO, see (19)). Hence, the transition of meaning from pure indefiniteness (non-specificity) to these lexical pairs seems to be quite expectable.

4.3. The marking of non-human animates

4.3.1. The referential properties rule

As demonstrated in 4.1, the statistical criteria show that the unmarked variant is distributed differently with human and non-human animate DOs. The semantic factors favouring the choice of this variant differ indeed, as shown below. There are 95 examples of unmarked non-human animate, which comprise a much higher percentage (30%) than the one observed with humans (21%). One of these examples is obviously explained in terms of creation of a new object:

(29) *tiń ož́ tiń*, ***podruga*** *šeď-t-ə̂sa*
here so here girlfriend be.found-CAUS-CVB
'In this way [the rooster] found himself a girlfriend' (Corpus)

Obviously, this context requires the absence of DO marking with non-human animates, as well as with humans. However, there is only one example that fits this explanation, the other 94 fall out of this rule. The first hypothesis that I am considering is the referential properties rule (a) suggested in Сердобольская, Толдова 2012. To test this hypothesis, I annotated all the examples with non-human animates according to their referential properties. The results are as follows.

A large subset of unmarked DOs includes indefinite, non-specific, and generic non-topical DOs (the number of examples is 47), exactly as it is predicted by the referential properties rule (a):

- Indefinite specific DO:

(30) *čukna keľa-l-o tə̂n-ə̂-d* ***kə̂k val*** *i ə̂robo azveš́*
morning send-EXP-FUT(1SG) you-DAT-P.2(SG) two horse and cart silver
'(Well done, my girl, I will give you a reward.) In the morning I will send you two horses (not mentioned before) and a cart full of silver' (Corpus)

- Indefinite non-specific DO:

(31) ***keš́*** *vož́-em-e pot-e*, ***kureg*** *vož́-em-e*
goat hold-NMLZ-P.1SG go.out[15]-PRS.3SG chicken hold-NMLZ-P.1SG
pot-e
go.out-PRS.3SG
'I want to keep goats, I want to keep chicken' (Corpus)

[15] The construction "nominalization + possessive + go out in PRS.3SG" is used to denote will.

• Generic non-topical DO:

(32) *zək prazńik il'jin deń,* ***ťaka*** *vand-iško-m*
big feast Ilya'(RUS) day(RUS) ram slaughter-PRS-1PL
'Ilya's day (the day of the prophet Elijah) is a big feast, we usually slaughter a ram for it' (Corpus)

The DO 'two horses' in (30) is indefinite specific (it is mentioned for the first time in the current discourse); the DOs 'goat' and 'chicken' are non-specific, since the speaker only plans to keep them, and their referent cannot yet be identified; the DO 'ram' in (32) is generic, since each year different rams are chosen for the feast.

Hence, it seems that the referential properties rule does play a significant role in the choice of the DO marking. The results are given in Table 3.

Table 3

The marking of non-human animate DOs and their referential properties[16]
χ^2 = 8.7, df = 2, p = 0.0129

DO marking	Accusative (singular or plural)	No marking	Total
Definite	**71 (68%)**	**34 (32%)**	**105**
Indefinite specific	**26 (44%)**	**33 (56%)**	**59**
Indefinite non-specific	5 (56%)	4 (44%)	9
Generic topical	9	0	9
Generic non-topical	0	9	9
Universal	18	0	18
Total	129	80	209

The standardized residuals test attributes the highest significance to the first two rows (+1.12/–1.35 to the first row and –1.43/+1.72 to the second row), while the third row does not show a significant departure from the expected values. Hence, it can be concluded that definite DOs (animate non-human) take the accusative, and indefinite specific DOs are unmarked, with more than chance frequency. For non-specific DOs the difference seems to be statistically insignificant (however, I will return to this in section 4.3.4). The distribution of marking among generic topical, generic non-topical and universal DOs conforms to the referential properties rule (a).

Thus, there is a rather large percentage of definite DOs that occur without marking and a large percentage of accusative indefinites. Hence, there must be another factor interfering with the referential properties of the DO.

4.3.2. Definite unmarked DOs

Let us consider some examples with definite unmarked DOs:

(33) *mar-ke ʑ́ə̑t-až́-e* [---] *ber ʑ́ə̑t val=ńi, olo* ***skal***
what-INDEF evening-front-ILL late evening be.PST=already, whether c o w

[16] The DOs marked with possessives are not taken into account here. When calculating the chi-square, I did not take into account generic and universal DOs, since the corresponding rows contain zero values.

o-mar kə̑sk-em-ə̑šʹ pə̑r-iśko val
INDEF-what pull-NMLZ-EL enter-PRS-1SG быть.PST
'(Have you seen Lada's daughter? — Come on, I haven't seen Lada yet! They don't go out) Once in the evening [−−−] it was late, I came in after having milked m y / t h e c o w or something' (Corpus)

(34) *parśʹ-jos-tə̑ šud-išʹko-m, sre* ***skal*** *skə̑sk-iśko-m*
pig-PL-ACC.PL feed-PRS-1PL then c o w milk-PRS-1PL
'(We usually get up in the morning) feed the pigs, then milk t h e c o w s' (Corpus)

In (33) the speaker is talking about her cow and in (34) about her family's cows, 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 (see Karttunen 1974; Stalnaker 1974), as the context implies that it belongs to the speaker. Note that in (34) the speaker uses the accusative (plural[17]) for the first DO, and no marking for the DO in the second clause, even if they have the same referential properties.

The DOs in nearly all such cases denote cattle and farm animals. It is possible that lexical factors are relevant here (see Кретов 1992 on the relevance of lexical classes for grammatical phenomena in Russian and von Heusinger 2008 for the relevance of lexical classes of verbs in Spanish DOM). I consider two hypotheses in connection with these data: 1) the lexical class of farm animals allows no marking (contrary to wild animals); 2) the lexical DO—verb pairs denoting typical farm activities (and, supposedly, other typical activities) allow no marking. I annotated all the examples with non-human animates (excluding possessive-marked DOs) and tested both hypotheses with the chi-square test. The result for the first hypothesis is shown in Table 4.

Table 4

Distribution of DO marking among the class of non-human animates in the Beserman corpus (10 539 sentences, 2187 DOs; $\chi^2 = 92.59$, df = 2, p < 0.0001

DO marking	Farm animals	Other lexical classes	Standardized residuals
ACC	**13**	**71**	**–4.46/+4.44**
ACC.PL	21	32	–1.05/+1.04
No marking	**80**	**12**	**+5.05/–5.03**
Total	114	115	

Thus, the statistical criterion confirms the hypothesis that the lexical class of farm animals significantly differs from other nominals with respect to the frequency of unmarked DOs. The standardized residuals test shows the significance of the first and the third rows, while the accusative plural row is hardly significant at all (the values are –1.05/+1.04). The accusative plural DOs are going to be addressed below in section 4.3.4.

Let us now consider the second hypothesis. I did not only test the pairs denoting farm activities, taking a larger class of lexical pairs that denote all kinds of typical activities. A typical activity can be defined as an activity

[17] The occurrence of the accusative on the noun 'pigs' may be caused by the presence of the plural marker (see Klumpp 2008 : 151—152). Such examples, at first sight, seem to break the referential properties rule (i); see section 4.4.4 for the explanation.

predominantly associated with a particular kind of object, e.g. milking a cow, writing a letter, reading a book etc.[18] In English, such pairs do not take articles, e.g. *leave town, attend school* and have been claimed to exhibit properties of semantic incorporation (Stvan 2009; de Swart, Zwarts 2009). The relevance of such lexical pairs for DOM is shown in Сердобольская, Толдова 2013 based on the data of Komi-Zyrian. The authors demonstrate that the pairs of nouns and verbs denoting specific farm activities, such as milking cows, feeding cattle, fishing (lit. catching fish), shooting hares etc., demonstrate special behavior with respect to DO marking. In Beserman the DO in such pairs can remain non-marked even if it is definite, as in (33) and (34). The relevance of this semantic parameter is confirmed by the chi-square criterium, see Table 5.

Table 5

Distribution of DO marking with typical activities (for non-human DOs)
$\chi^2 = 63.65$, df = 2, $p < 0.0001$

Typical activities	yes	no	standardized residuals
ACC.SG	**4**	**80**	**–4.47/+3.11**
ACC.PL	15	39	–0.62/+0.43
non-marked DOs	**56 (74.7%)**	**36 (23.2%)**	**+4.75/–3.3**
Total	75	155	

Therefore, DOs occurring in typical activity pairs tend to be unmarked or take the accusative plural, while in other contexts accusative marking is by far more frequent. It is thus unclear which of the two parameters (class of nouns or lexical pairs) is more significant. Given that nearly all the typical activity pairs in the corpus include nouns denoting farm animals, and vice versa, this question cannot be clarified at the moment.

Note that, just as with the hypothesis in Table 4, the standardized residuals test shows the significance of the first and the third rows, while the accusative plural row is not significant at all (the values are below 1).

I can offer the following explanation for these facts. If it is the parameter of the lexical class of farm animals that is significant, it could be explained by individualization. Farm animals (dogs/cats living on the farm are not included) are often dealt with in herds and flocks, where the individualization of particular animals is not required, while wild animals are seen one by one (during hunting or in case they come into the village or attack the herd).

With typical activities the distribution is also quite understandable. Let us return to the examples (33) and (34). The speaker tells about her everyday activities, and her intention is not to refer to particular referents, but rather to describe the situation as a whole (similarly to electing the president in 4.2). Stvan (2009) argues for the analysis of English bare singular nouns in examples like *attend school*, *leave town* in terms of semantic incorporation, based in their inability to take modifiers, allow plurality marking

[18] The following lexical pairs have been attested in the corpus of Beserman Udmurt: 'milk a cow', 'feed/water/keep/raise/shepherd/slaughter/bring in/assemble/turn out to grass cows/pigs/geese etc.', 'harness a horse/ox', 'dig out worms', 'fish (lit. catch fish)'. In Komi-Zyrian, such pairs also include hunting terms; it is, however, absent in the Beserman corpus (in general, hunting is not very much widespread in the Beserman village I worked in).

and anaphoric reference. She generalizes that bare singulars show "effects as if they are part of some type of larger fused predicate that does not introduce a discourse transparent referent and does not imply a singular referent" (Stvan 2009 : 331). In line with this, the examples (33) and (34) are explained by assuming that even if at a given point of time the speaker could have a particular cow/cattle in mind, the noun 'cow/cattle' is used to point to a larger situation of cow-milking (in particular, in (33), the speaker's intention is to locate another event at a specific time interval, namely, after the milking took place) and not to introduce the particular cow as a discourse referent (cf. Abbott 2011 : 62, example 21 and its interpretation). Given that the DOs in (33) and (34) do not introduce a discourse referent, they occur unmarked (similar cases in other Finno-Ugric languages are considered in detail in Толдова, Сердобольская 2002). Most examples with typical activity pairs are similar: the narrators tell something about their own cattle, which is definite and possessed (and, hence, likely to take the possessive accusative); however, the corresponding DOs occur unmarked.

Thus, both hypotheses are confirmed by the chi-square test and have clear semantic explanations. Both of them can account for the majority of unmarked definite DOs belonging to the non-human class. The number of unmarked definites is 34, and about 85% of them belong to farm animals/ typical activity class. Another 15% belong to the experimental texts with cards representing animals:

(35) *odig-ez-len* *jə̑r-ez* *və̑lla* *pal-a-z*
one-P.3(SG)-GEN1 head-P.3(SG) upper side-LOC/ILL-P.3(SG)
sə̑l-tə̑-t *eto* [–––] ***kə̑k kureg pi***
stand-SMLF-CAUS(IMP.SG) this(RUS) two chicken son
'(Then on top of the (card with) the goose are two (cards with) yellow chicken. You have two geese, right?) Put these, two (cards with) chicken on top of one of the geese's head' (Corpus)

The experiment was carried out by Olga Biryuk and Maria Usacheva in 2011; it was designed according to the referential communication method: one subject was told to put the cards into the picture, and the second subject had to explain the first subject how they should be placed (see Бирюк, Усачева 2012 for details). All the cards were given to the subjects in advance, which means that they are definite. These cards with pictures could have been interpreted as inanimate objects and, hence, unmarked even if definite, see 4.4. However, they represented animals and, hence, could have been interpreted as animate. In the corpus the distribution of accusative/no marking with experimental cards is approximately equal; this obviously means that it is a point of variation among the speakers.[19]

Thus, the following rule can be proposed for the definite non-human animates:

(b) definite non-human animate DOs take the accusative unless they belong to a lexical class of farm animals / to the typical activity pairs;
NB. Illustrations, cards and pictures are obviously treated differently by native speakers and can thus follow both patterns, animate or inanimate.

[19] It is noteworthy that the DOs referring to cards representing humans are all treated in the same way as human DOs, i.e. they are all marked with the accusative.

4.3.3. Indefinite DOs with the accusative

As shown in Table 3, indefinite DOs are often unmarked (56%). However, they occur with the accusative in 44% of cases. Serdobolskaya and Toldova show the relevance of discourse factors for the marking of indefinite specific DOs (Сердобольская, Толдова 2017). In a specific experiment they tested the difference in marking of protagonists (characters mentioned in the discourse no less than 10 times) and incidental characters. According to this experiment, the protagonists occurring in introductory contexts, and, thus, having indefinite specific status, take the accusative in all 100% of cases, while other indefinite specific DOs are unmarked. Compare the following examples:

(36) *uj-ə̂n nə̂l gur-ze est-i-z*, [–––] *ʒ́uk*
night-LOC girl stove-ACC.P.3(SG) stoke-PST-3(SG) porridge
*pəž́-t-i-z i aʒ́-e šə̂r-**ez***
be.cooking-CAUS-PST-3(SG) and see-PRS.3SG mouse-ACC
'At night the girl stoked the stove, cooked porridge and saw a mouse' (Corpus)

(37) *i vaj-e so so-lə̂ **val**, bə̂des ə̂robo azveš́*
and give-PRS.3SG that that-DAT horse full cart silver
'And he gives her a horse and whole cart of silver' (Corpus)

In (36) the mouse is one of the main characters in the text, since it helps the girl escape from the bear (the mouse is mentioned 10 times in the text). Although it occurs in the introductory context (as a new participant of the discourse), it takes the accusative in all 4 texts (other native speakers used another construction, without the DO). By contrast, the DO *val* 'horse' in (37), which is mentioned only once in the text, is unmarked in all 6 texts.

Given that this factor is highly relevant for the choice of DO marking, I annotated the indefinite specific DOs according to the parameter of protagonism. The results are given below.

Table 6

The marking of indefinite non-human animates in the corpus
χ^2 = 27.84, df = 2, p < .0001

DO marking	Accusative (singular or plural)	Acc.sg	Acc.pl	No marking	Total
Indefinite specific	26	25	1	33	59
— protagonist in introductory context	19	19	0	3	22
— other	7	6	1	30	37
Indefinite non-specific	5	0	5	4	9

It is evident from the Table 6 that the protagonist explanation covers the vast majority (73%) of cases where the indefinite specific DO takes the accusative (although the protagonism rule is not strict: there are 3 instances of unmarked protagonists). Other 7 instances comprise one example of accusative plural (see 4.3.4 below) and 6 examples of the following origin: they all belong to one and the same sentence in the same experimental text and describe an incidental character occurring only once in the whole text.

Among 7 experimental texts this character (a wild duck) occurs with the accusative in 6 texts, and it is unmarked in one text only (while all the other incidental characters in experimental and spontaneous texts are unmarked):

(38) *kut-i-z-ɔ̂* *so-os* ***lud*** ***ćɔ̂ž-ez*** *i* *so-je* *śi-i-z-ɔ̂*
grab-PST-3-PL that-PL wild duck-ACC and that-ACC eat-PST-3-PL
'They caught a wild duck and ate it' (Corpus)

(39) *so* *šer-k-ɔ̂n* *gondɔ̂r-en* *kijon-en* *šij-o* ***lud*** ***ćəž*** *kut-ɔ̂sa*
that behind-OBL-LOC bear-INS wolf-INS eat-PRS.3PL wild duck grab-CVB
'Then the bear and the wolf caught a wild duck and ate it (lit. ate a wild duck after having caught)' (Corpus)

It could be that the presence of the lexeme 'wild' plays a significant role (see 4.4.2 for the relevance of the presence of modifiers in inanimate DOs). However, this cannot be tested for non-human animate DOs, since there are no more such DOs with modifiers in the corpus.

As for non-specific DOs, note that all the 6 instances that are not explained by the referential properties rule (a) take the accusative plural. The specific distribution of the accusative plural (in comparison with the accusative singular) is dealt with in the next section.

Thus, I can formulate a rather strict rule for the indefinite non-human animate DOs:

(c) indefinite non-human animate DOs occur unmarked, unless they introduce a new protagonist in the discourse (in that case they take the accusative).

The exceptions to this rule concern DOs with modifiers (this will be confirmed by the results in 4.4.2).

4.3.4. Accusative plural DOs

Let us examine the specific case of accusative plural DOs in -(*j*)*ostə*. Note that they show peculiar distribution compared to the accusative singular DOs (see previous sections and Кондратьева 2000 on Standard Udmurt). First, they are largely used in typical activity pairs or with nouns denoting farm animals. Second, they mark indefinite specific (non-protagonists) and non-specific DOs (Table 6), while the accusative singular usually does not.

Note that in the DO position the plural suffix requires the presence of the accusative (see example (5); the explanation is offered in 4.4.4). Hence, the combination of both must be dealt with separately from accusative singular. The peculiarities of the use of -(*j*)*ostə* need to be analyzed together with the distribution of the plural suffix -(*j*)*os* in Beserman (as well as in Standard Udmurt). The examined variety has optional plurality marking (in terms of Corbett 2000 : 70), which means that the nominal plurality suffix is not used in all contexts even if the nominal refers to a non-singular entity, see example (34) with the singular nominal *skal* 'cow' denoting a quantity of cows. Rather, its distribution is regulated by several factors, such as animacy, referentiality and information structure (see Шматова, Черниговская 2012 for details). However, the exact rules of distribution are yet to be formulated.

Hence, the distribution of the accusative plural combination in Beserman is a matter of future studies; in any case, it differs from the distribution of accusative singular.

4.3.5. Non-human animate DOs: generalizations

It can be seen from sections 4.3.1—4.3.4 that it makes no sense to study isolated parameters influencing the choice of DO marking. In all cases (indefinite vs. generic vs. definite DOs; typical activities, farm animals etc.) a great number of exceptions is found, making it impossible to formulate any strict rules. However, if we consider the combinations of the parameter values (definite DOs in typical activity pairs; generic DOs in the topic), we can formulate rather strict rules and the exceptions make isolated instances. The rules for non-human animates are the following (this repeats the rules (b) and (c) given above):

(d) The distribution of accusative singular and no marking with non-human animate DOs is regulated by the following rules:
- definite non-human animate DOs take the accusative unless they belong to the lexical class of farm animals / to typical activity pairs;
- indefinite (specific and non-specific) non-human animate DOs occur unmarked, unless they introduce a new protagonist in the discourse;
- generic DOs are only marked if they are topical;
- universal and attributive DOs take the accusative.

There are three reservations:

1) Illustrations, cards and pictures can follow the animate or the inanimate pattern.
2) The distribution of accusative plural is yet to be studied.
3) The marking of lexical pairs denoting creation of a new object is yet to be studied.

These rules slightly differ from the ones for human DOs. The latter do not remain unmarked even if indefinite/generic non-topical. As for the parameter of typical activities, similar pairs for human DOs are unattested in the corpus, and, hence, this parameter cannot be tested.

4.4. The marking of inanimate DOs

4.4.1. Referential properties and the marking of inanimate DOs

The data in Table 2 and Figure 2 show that inanimates occur unmarked more frequently (57.5%) than animate DOs. This in part follows Kondraťeva's generalizations for Standard Udmurt (Кондратьева 2010), see scheme 1. The whole number of accusative (singular) inanimates is 126, which makes 8% of all the inanimate DOs in the corpus (for the accusative plural the number is 33, which makes 2%; those are going to be dealt with in section 4.4.4). The question is, then, what triggers the accusative marking with these 126 DOs. The first hypothesis to test was the referential properties rule (a). It predicts that definite, universal, topical generic and attributive DOs take the accusative. I annotated all the accusative singular examples for the referential properties of the DO and found out that 52 of these examples include definite DOs (40), one includes a universal DO, 11 include topical generic DOs as the second DO in (43), two include attributive DOs. Five DOs are designations (e.g. "Do you know what sickle is?" lit. 'do you know sickle'); according to the corpus data, they also require the accusative.

(40) *kɔ̂k-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) hand-PROL-P.3(SG)
*kut-i davayte mešok-**ez** pi nu-e kɔ̂-t-ɔ̂ś́*
catch-PST(1SG) come.on(RUS) bag-**ACC** AUTOCIT carry-IMP.PL where-OBL-EL
baś́t-i-d-ɔ̂!
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 t h e b a g back to the place you've taken it from!"' (Corpus)

Note that this rule is not observed strictly: definite inanimate DOs may occur unmarked:

(41) *mašina* [---][20] *mar kar-e, baś́t-o-d, uža-l-o-d*
car what do-PRS.3SG take-FUT-2(SG) work-EXP-FUT-2(SG)
*vɔ̂ldɔ̂, baś́t-i **molokovoz***
after.all take-PST(1SG) milk.truck
'(Then I was given a car, a milk truck) What shall you do with a car, if you take it, you're going to work on it, (so) I took t h e m i l k t r u c k' (Corpus)

(42) *kɔ̂-t-ɔ̂ś́ so aʒ́-i-z pič́i=ges **korka***
where-OBL-EL that see-PST-3(SG) small=CMPR house
'(The soldier saw a small house from the top of the tree. [---] They went to the place) where he saw t h e s m a l l h o u s e' (Corpus)

In these examples the nouns *molokovoz* 'milk truck' and *korka* 'house' are aforementioned in the previous sentence and, hence, definite. However, they are unmarked, contrary to the referential properties rule (a) and scheme 1.[21]

Topical generic DOs most often take the accusative, in line with the rule in (a):

(43) ***jetɔ̂n** kiž́-ɔ̂l-i-z-ɔ̂, jetɔ̂n-**ez** ɔ̂šk-išḱo-m tiń taž́*
flax seed-ITER-PST-3-PL flax-**ACC** pull-PRS-1PL here so
'(The speaker is telling about flax breeding and processing in the years of her youth) We seeded the flax, then we pulled the flax in this way' (Corpus)

However, they may also be unmarked, as the first DO in the same sentence[22] (cf. Klumpp 2008 : 144 on unmarked topical DOs in Komi).

Thus, the referential properties rule accounts for the accusative marking of 66 cases (52 definite DOs + 1 universal DO + 11 generic DOs + 2 attributive DOs, see Table 7 below).

There is another class of generic DOs that most often occur with the accusative. Consider two examples:

(44) *baś́t-išḱo-m val kruška-je šukaś́* [---] *i obezaťel'no č́ag*
брать-PRS-1PL be.PST mug-ILL kvass and necessarily kindling

[20] Unintelligible fragment.
[21] This is in line with Klumpp's (2014) observation on unmarked discourse topical DOs in Komi.
[22] Note that this DO is pronounced without a pause with the verb, and cannot be interpreted as a clause-external topic.

čag-ez *ʒ́ua-t-išḱo-m* *korka-šen=ik* *i* *ʒ́ut-ə̂sa*
kindling-ACC burn-CAUS-PRS-1PL house-EGR-EMPH and lift-CVB
'(The speaker is telling about the custom of seeing off masked people after Christmas) We pour kvass [–––] And a kindling, necessarily. We light a kindling when going out of the house and carry it lifted high in the air' (Corpus)

(45) *so-kə* *vaj-iš'ko-m* ***viľ*** ***veńə̂k***
that-when give-PRS-1PL new broom
'(The beginning of the text: The sixth of July is St. John the Baptist's day) At the time we usually bring new brooms' (Corpus)

Both DOs in bold are generic, since these texts describe the activities that are undertaken in the village every year during the feasts. However, in (44) the generic DO is aforementioned, while in (45) it is not. Accordingly, in (44) the DO takes the accusative, while in (45) it is unmarked. It is quite expectable that anaphoric DOs take the accusative: they are aforementioned and, hence, in each situation a particular NP refers to an already established referent (even the situation is generic). Hence, it is already given in the preceding context and treated as definite, that is, marked with the accusative. Anaphoric DOs of this class comprise 28 instances among the accusative DOs, while generic non-anaphoric DOs are only 3.

The distribution of inanimate DOs marked with the accusative singular is summed up in Table 7.

Table 7

Distribution of referential types among accusative DOs

Referential properties	**Number of accusative singular DOs**
Definite	52
Designations (descriptions)	5
Universal	1
Generic topical	11
Generic anaphoric	28
Attributive	2
Not defined[23]	11
Other generic DOs	3
Indefinite (specific and non-specific)	13
Total	126

The results observed in Table 7 are mostly predicted by the referential properties rule (a): definite, universal, generic topical and attributive DOs take the accusative. Two types are added to this rule, generic anaphoric expressions and designations. However, there are 16 DOs that violate this rule (13%). These exceptions require an explanation, which is provided in the following section.

Another question concerns the obligatoriness of the referential properties rule. We have seen that it is not obligatory: there are exceptions, like (41)

[23] This includes instances that allow several interpretations from the point of view of referential properties.

and (42) with a definite DO and (43) with a topical generic DO (the first DO in the example). Hence, there must be another factor influencing the presence/absence of the accusative with inanimates.

4.4.2. NP structure and the marking of the DO

Let us examine the 16 generic and indefinite DOs in detail. It turns out that most of these NPs (13, that is 81% of 16) contain some material other than the head noun (adjectives, pronominal modifiers, juxtaposed dependent nouns, e.g. *ǯaǯeg siľ* 'goose meat' etc.).

(46) *gondə̂r č́ašja -je mə̂n-i-z* [−−−] *i šeď-t-i-z*
bear forest-ILL go-PST-3(SG) and be.found-CAUS-PST-3(SG)
kyz č́ašja-jez *i vaj-i-z ǯič́ə̂-lə̂*
thick tree-ACC and bring-PST-3(SG) fox-DAT
'The bear went to the forest [−−−] and found a thick tree and brought it to the fox' (Corpus)

This sentence is part of the experimental texts where the speakers retold the same story. First, the bear went to the forest to find a good tree to make a shaft, then the wolf went to find a tree. It is important that the tree should not be too thick or too thin. Consequently, these sentences provide a good opportunity to compare the marking of indefinite specific DOs (non-protagonists, since both trees are not taken for the shaft and not discussed afterwards).

In these texts the DO 'thick tree' is marked in 3 out of 7 texts (as in (46)), while the DO 'birch' is never marked in any of the 7 texts:

(47) *kijon mə̂n-i-z, va-i-z* ***kəǯpu***
wolf go-PST-3(SG) bring-PST-3(SG) birch
'The wolf went [to the forest] and brought a birch' (Corpus)

It is notable that both DOs 'thick tree' and 'birch' occur in similar contexts, in texts authored by the same speakers. The only significant difference between (46) and (47) is the presence of an adjective. Consider also (38) and (39) with non-human animate DOs. Hence, it can be supposed that the presence of modifiers plays a role in the choice of the DO marking. I annotated all the examples with inanimate DOs (1054 instances) according to their internal structure. The results are shown in Table 8 and analyzed below.

Table 8

Distribution of DO marking depending on the NP structure[24]
$\chi^2 = 66.088$, df = 1, $p < 0.001$

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**

[24] 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. Proper nouns are excluded, since the choice of DO marking with them is subject to lexical restrictions (see section 4.4.3).

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, with accusative DOs (they are given in boldface). Thus, it might be concluded that the factor of morphosyntactic heaviness is relevant for the choice of the accusative. 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 NP modifiers that require the accusative, and there are ones that require (or favour) the absence of marking (indefinite pronouns).

Let us examine various groups of modifiers with inanimate DOs. First, there are modifiers that nearly always occur with accusative nominals. These are demonstrative pronouns (48), which require the accusative on the DO in 22 out of 25 cases, accusative plural in one case, and no marking in two cases (these two instances, however, include the accusative on the pronoun, e.g. *so-je mad'*, that-ACC song, which is a very marginal construction for Udmurt).

(48) *mon so-je,* ***so*** *šɜ̂d-**ez*** *ǯuž̌-i,* *ǯuž̌-i*
I that-ACC t h a t soup-**ACC** take.a.gulp-PST(1SG) take.a.gulp-PST(1SG)
gine ɜ̂m dur-ɜ̂ bɜ̂des kwal'ek-ja-nɜ̂ kučk-i-z
only 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)

This result is predicted by the referential properties rule: demonstrative pronouns mostly introduce definite or generic anaphoric NPs, and therefore they take the accusative.

Second, there are modifiers that, expectedly, require or strongly prefer the absence of marking (or the accusative plural): indefinite pronouns (all 4 instances), pronouns meaning 'such' (all 4 instances), the interrogative 'what' (all 6 instances), numerative groups with mass nouns (all 33 instances):

(49) *a ďeńis, so* ***mar-ke*** ***kɜ̂l*** *so svala=wa*?[25]
and Denis that what-INDEF language that understand(PRS.3SG)=Q
'What about Denis, does he speak any other language?' (Corpus)

(50) *mi pun-išk̆o-m* ***šiž̌ɜ̂m-ťamɜ̂s kilogram pesok***
we put-PRS-1PL seven-eight kilogram granulated.sugar
'We (usually) put seven or eight kilogram of granulated sugar' (Corpus)

Most of these facts are predicted by the referential properties rule: indefinite pronouns and the wh-question word 'what' introduce indefinite NPs and, hence, occur unmarked. The pronouns meaning 'such' introduce NPs with complicated referential properties that require a separate study. As for numerative groups, so far I cannot offer any explanation for the observed fact.

In case of indefinite pronouns, the elicitation sessions conducted by Serdobolskaya and Toldova (Сердобольская, Толдова 2012) show that the accusative is most often disallowed by native speakers.

The third group comprises modifiers that allow both the accusative, accusative plural and no marking. The number of occurrences is given in Table 9.

[25] The demonstrative in this example refers to Denis and does not make part of the DO.

Table 9

The marking of inanimate DOs with various types of modifiers

Modifier type	ACC.SG	ACC.PL	No marking
adjectives	12	2	56
juxtaposed nouns	10	1	31
cardinals	2	0	17
'many'	4	1	4
quantifiers	2	0	4

I omit the data on comitative nouns (e.g. 'porridge with sugar'), nominalizations ('talk'), participles, dependent infinitives and relative clauses, since these groups are represented by a very small number of examples. The data on 'many' and quantifiers in DOs also do not allow to draw any valid conclusions, since there are less than 10 examples in each of these groups. Hence, I focus on the first three rows of Table 9, where the figures are high enough.

Table 10

The marking of DOs with adjectives (F = 0.0011, p < 0.01)

Modifier type	ACC.SG	no marking	Total
Adjectives	12 (18%)	56	68
Head-only (one-word) DOs	47 (6%)	764	811

Table 11

The marking of DOs with juxtaposed nouns (F = 0.0002, p < 0.01)

Modifier type	ACC.SG	no marking	Total
Juxtaposed nouns	10 (15%)	31	41
Head-only (one-word) DOs	47 (6%)	764	811

Table 12

The marking of DOs with cardinal numerals
(F = 0.3104. The result is not significant at p < 0.1)

Modifier type	ACC.SG	no marking	total
Cardinals	2 (11%)	17	19
Head-only (one-word) DOs	47 (6%)	764	811

Tables 10 and 11 demonstrate that inanimate DOs with adjectives and juxtaposed nouns show a significantly higher percentage of accusative than DOs that only include the nominal head (confirmed by the Fisher Exact Test). By contrast, DOs with cardinals do not seem to show a significant deviation from the expected values.

Thus, the factor of NP structure seems to influence the choice of DO encoding, independently of the referential properties factor: DOs with adjectives/juxtaposed nouns are more likely to take the accusative than single-word DOs. However, it is not a strict rule and even not a tendency, but a matter of a slightly higher frequency. This observation is supported both by the data on inanimates from experimental texts and by the results on non-human animate DOs, see section 4.2.3.

The obtained result may be explained by the "heaviness" of the NP that influences the choice of DO marking. "Heavier" DOs are more likely to be marked than light (one-word) DOs. However, this generalization does not work in case of DOs with cardinal numerals. It might be that the system of DO encoding is changing, and this factor is slowly being integrated into it.

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 DOs are preferably unmarked. By contrast, with definite, attributive and some other DOs the accusative is frequent, even if not obligatory. The accusative is strictly required when the definiteness semantics is "reinforced" by the presence of lexical expressions signaling definiteness, such as demonstrative pronouns. Another intervening factor is the heaviness of the DO: NPs with adjectives and juxtaposed nouns are more likely to be marked than one-word NPs.

4.4.3. Proper names of inanimates

Proper names are usually definite, and, hence, they are expected to always take the accusative. This is most often true for proper names of animates (both human and non-human[26]), while inanimates can be divided into two classes with respect to the choice of the DO marking. Toponyms (as countries, cities, villages etc.) and specific institutions take the accusative, while names of feasts, journals and broadcasts occur without marking:

(51) *a armij-e=ke e-j mən-əsal ku mon so-je aʒ́-isal*
and army-ILL=if NEG.PST-1SG go-COND when I that-ACC see-COND
germańi-ez*=no*?
Germany-ACC=ADD
'If I hadn't been to the army, would I ever have seen Germany? (lit. when would I)' (Corpus)

(52) ***akaška****=no kar-ə̑l-i-z-ə̑, korka-ś́ korka=no pə̑ra-l'l'a-z-ə̑*
Akashka=ADD do-ITER-PST-3-PL house-EL house(ILL)=ADD enter-ITER-3-PL
'They celebrated Akashka, they went from house to house' (Corpus)

This is, obviously, a lexically-based distribution.

4.4.4. Accusative plural marking of inanimates

Inanimates in the accusative plural comprise 33 instances. This combination is observed with both types of referential classes — those that most often take the accusative and those that mostly occur without it. The distribution of the two types of classes is 45 vs. 48%, as demonstrated in the Table 13.

Hence, no valid generalizations can be made about the referential properties of DOs in accusative plural. The data on modifiers (adjectives, juxtaposed nouns etc.) does not allow to make any conclusions either. Apparently, plural inanimate DOs are more frequently marked with the accusative than singular inanimate DOs. The same tendency has been observed for Standard Udmurt in Кондратьева 2000 : 100 and for Komi in Rounds 1990. Klumpp (2008) explains

[26] Other referential types are also possible, e.g. generic, as *All Marys are gifted* (the girls with the name Mary). However, such contexts did not occur in our corpus.

Table 13

The distribution of the accusative plural *-jos-tə* among referential classes

Referential status	Number of instances
Referential classes that take the accusative	**15 (45%)**
Definite	11
Generic anaphoric	2
Generic topical	2
Referential classes that do not take the accusative	**16 (48%)**
Generic	7
Indefinite non-specific	2
Indefinite specific	7
Not defined	2
Total	33

this on the basis of the diachronic origin of the plural markers in Permian languages: in both Komi and Udmurt these markers go back to lexemes denoting humans ('people' in Komi and 'member of a group' in Udmurt). Therefore, they may preserve the preference for the accusative marking in spite of the change in meaning and use. Thus, the distribution of the accusative plural presents an interesting topic for further investigation.

4.4.5. Generalizations

The quantitative analysis of inanimate DOs in the corpus allows formulating some tendencies (not strict rules as in case of animate DOs) in the choice of marking of these DOs.

The distribution of accusative singular and zero marking with inanimate DOs is regulated by the following rules:

- inanimate DOs are marked if they occur with demonstrative pronouns;
- they are unmarked if they occur with indefinite pronouns, the interrogative 'what', the pronoun 'such' and in numerative phrases;
- otherwise, they are more likely to take the accusative if they are definite, generic topical, generic anaphoric, universal, attributive, or if they are designations;
- indefinite and generic DOs are more likely to take the accusative if they include adjectives or juxtaposed nouns;
- proper names of inanimates: toponyms (such as countries, cities, villages etc.) and specific institutions take the accusative, while names of feasts, journals and broadcasts occur without marking.

The distribution of accusative plural is yet to be studied.

Thus, the functions of the accusative (singular) are narrowed in case of inanimate DO: it is not obligatory with definites (as in case of animate DOs); it is only obligatory with DOs that include demonstrative pronouns (and several lexical classes of proper names). In part, the distribution of the accusative is widened to indefinite/generic DOs with adjectives and juxtaposed nouns; however, this is a matter of frequency and not a strict rule.

5. Discussion

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 (Кондратьева 2010); in elicited examples from Beserman this factor seems to be much less important. The corpus analysis enables us to draw 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 unmarked form is restricted to the situations of creation of a new object and to some specific lexemes (*ɫalə̑k* 'people', *pińal* 'child' and mythological characters).

Non-human animates are mostly unmarked if they belong to the class of farm animals or take part in typical activities (supposedly, also in the context of creation of a new object). Otherwise, definite, generic topical and universal DOs always take the accusative. Indefinite and generic (non-topical) DOs are unmarked unless they introduce a new protagonist into the discourse.

The marking of inanimates does not depend as much on the referential properties of the DO. The most important factor is its heaviness and the type of modifiers included. Demonstrative pronouns mostly require the accusative, while indefinite pronouns, the interrogative 'what', pronouns meaning 'such' and numerative phrases mostly take zero marking. Bare nouns may take the accusative if they are definite, generic topical, generic anaphoric, universal, attributive, or if they are designations (descriptions). Indefinite and generic (non-topical) DOs are mostly unmarked; they might take the accusative, though, if they include adjectives or juxtaposed nouns. Proper names of inanimates are treated as follows: toponyms (such as countries, cities, villages etc.) and specific institutions take the accusative, while names of feasts, journals and broadcasts occur without marking.

The relevance of lexical classes of DOM is in line with recent studies such as Piñón 2006; Сердобольская, Толдова 2013; von Heusinger 2008.

Therefore, the factor of referential properties plays an unequal role for each animacy-based class of nouns: for human animates, the use of the accusative is expanded onto indefinite/non-specific DOs (excluding some lexical classes); for inanimates, its obligatory use is narrowed from all definite DOs to DOs with demonstrative pronouns and some classes of proper names. The influence of the heaviness of the DO is observed for inanimates only.

The data that do not conform with the tendencies listed above can be accounted for on the basis of information structure factors: according to my observations, *ceteris paribus*, DOs tend to be unmarked if they are part of the wide focus, despite the fact that they are definite (41); it might be that the sentence-initial position influences the absence of marking in (43) and similar examples (M. Usacheva and T. Arkhangelskiy, p.c.). However, the corpus data does not offer a possibility to give enough evidence in support of these hypotheses.

This distribution has been discovered by the means of corpus analysis, since 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 cardinal numerals or numerative phrases in DOs with definite semantics. These gaps can only be covered by elicitation of thoroughly

elaborated contextual minimal pairs. Until such pairs are found and tested, it cannot be concluded which factor prevails in case of the competition of factors.

6. Conclusions

The present study contributes to the studies of DOM in languages of the world in the following way. Many existing works study various factors in isolation (information structure, referential properties, animacy, lexical classes), giving evidence for or against a particular factor. However, I show the importance of studying combinations of factors' values: for some lexical classes, the referential properties are less relevant, since the corresponding DOs are always unmarked (e.g. typical activities); some combinations of the factors' values allow to predict the choice of marking on more firm grounds (generic anaphoric DOs) than others (indefinite inanimate DOs with modifiers). Each combination of the factors' values has to be studied independently from other combinations. To arrive at strict rules, minimal pairs have to be elicited so that specific sets of factors can be isolated.

DOM in Beserman does not yield neither to one-parameter models, nor to two-parameter models as the one proposed in Aissen 2003. Each combination of the parameters' values shows different lexical and syntactic restrictions on DOM. Moreover, the rules for each combination are not equally strict. The following lexical classes are relevant for Beserman: the DO-verb pairs denoting typical activities and creation of a new object (similar pairs are relevant for DOM in Komi-Zyrian, see Сердобольская, Толдова 2013). The importance of these pairs for Hungarian DOM and for English bare singulars is shown in Piñón 2006 and in Stvan 2009, respectively, and the theoretical explanation is offered based on referential properties of DOs. Given that one of the functions of the DO marker in Beserman is indication of definiteness, the relevance of these lexical pairs is expectable.

Another interesting issue is the "heaviness" of the DO, namely, DOs with modifiers tend to take the accusative more frequently than single-word DOs. This factor could be analyzed in terms of semantic incorporation; however, this analysis needs further elaboration.

It is noteworthy that many rules are not strict, especially in the domain of inanimate DOs. I hypothesize that it is due to the fact that DOM in Udmurt is currently in its transitional state from the definiteness-based system towards the animacy-based one. Many researchers claim that the accusative marker in Udmurt goes back to the 3rd person possessive suffix, which in turn was used as a definiteness marker (Майтинская 1979 : 102; Rédei 1988 : 382—383; Raun 1988; Kiss, Tánczos 2018). Given that definiteness is the original function of the accusative marker, it can be concluded that in the current stage of the language the system of DO marking is changing. The referential properties factor is being replaced by other factors, among which animacy is the strongest. With different animacy classes this process goes differently: with humans the functions of zero marking are narrowed from indefiniteness/non-specificity onto particular semantic types of situations (creation of a new object). With non-human animates the referential properties are still very relevant, with the only exception of specific lexical pairs. With inanimates the definiteness-based rule is weakened and turned into a mere tendency; its strictness can be "reinforced" by specific lexical devices signaling the referen-

tial status of the NP (demonstrative pronouns, indefinite pronouns etc.). It can be speculated that presence of restrictive relative clauses could also cause the choice of the accusative. The property of containing such reinforcing devices could be then reanalyzed as the heaviness of the NP, namely, in case of any modifier (except for indefinite pronouns) the speakers would prefer the accusative. The robustness of this explanation may be verified by the future direction of language change.

Address

Natalia Serdobolskaya
Institute of Linguistics, Russian Academy of Sciences
The Pushkin State Russian Language Institute
E-mail: serdobolskaya@gmail.com

Abbreviations

ACC — accusative; **AKK** — accusative; **ADD** — additive particle; **ATTR** — attributive; **AUTOCIT** — autocitation marker; **CAUS** — causative; **CIT** — citation marker; **CMPR** — comparative; **COLL** — collective numeral; **COND** — conditional; **COP** — copula; **CVB** — converb; **DAT** — dative; **DETR** — detransitive; **df** — degrees of freedom; **DO** — direct object; **DOM** — differential object marking; **EGR** — egressive; **EL** — elative; **EMPH** — emphatic particle; **EXP** — stem expansion; **F** — Fisher exact test of independence; **FUT** — future; **GEN1** — genitive (except with direct object); **HES** — hesitation marker; **ILL** — illative; **IMP** — imperative; **INDEF** — indefinite prefix; **INF** — infinitive; **ITER** — iterative; **LOC** — locative; **MULT** — multiplicative; **N** — non-marked DO; **NEG** — negation; **NMLZ** — nominalization; **NOM** — nominative; **NP** — noun phrase; **OBL** — oblique nominal stem; **ORD** — ordinal numeral; **p** — probability of obtaining the observed results under the assumption of independence of factors; **P.1/2/3SG/PL** — possessive markers; **PL** — plural; **POSS** — possessive markers; **PROL** — prolative; **PRS** — present; **PST** — past; **PTCL** — particle; **PTCP.ACT** — active participle; **Q** — question marker; **RECP** — reciprocal; **RES** — resultative; **RUS** — Russian word; **SG** — singular; **SMLF** — semelfactive; χ^2 — chi-square test of independence.

R E F E R E N C E S

A b b o t t, B. 2011, Reference. Foundational issues. — Semantics. An International Handbook of Natural Language Meaning. Vol. 1, Berlin—Boston, 49—74.
A i s s e n, J. 2003, Differential Object Marking. Iconicity vs. Economy. — Natural Language and Linguistic Theory 21, 435—483.
B a k e r, R. 1986, The Role of Animacy in Komi Direct Object Marker Selection. — UAJb. Neue Folge 6, 47—60.
C o r b e t t, G. G. 2000, Number, Cambridge.
C s ú c s, S. 1990, Chrestomathia Votiacica, Budapest.
d e S w a r t, H., Z w a r t s, J. 2009, Less Form — More Meaning. Why Bare Singular Nouns Are Special. — Lingua 119, 280—295.
d e S w a r t, P. 2007, Cross-linguistic Variation in Object Marking, Utrecht. https://www.lotpublications.nl/Documents/168_fulltext.pdf.
D é c s y, Gy. 1990, The Uralic Protolanguage. A Comprehensive Reconstruction, Bloomington (Bibliotheca Nostratica 9).
D o n n e l l a n, K. S. 1966, Reference and definite descriptions. — Philosophical Review 75, 281—304.
É. K i s s, K., T á n c z o s, O. 2018, From Possessor Agreement to Object Marking in the Evolution of the Udmurt *-jez* Suffix. A Grammaticalization Approach to Morpheme Syncretism. — Language 94, 733—757.

F r a u r u d, K. 2001, Possessives with Extensive Use. A Source of Definite Articles? — Dimensions of Possession, Amsterdam—Philadelphia (Typological Studies in Language 47), 243—267.
I e m m o l o, G. 2011, Towards a Typological Study of Differential Object Marking and Differential Object Indexation. University of Pavia PhD thesis, Pavia.
K a r t t u n e n, L. 1974, Presuppositions and Linguistic Context. — Theoretical Linguistics 1, 181—194.
K l u m p p, G. 2008, Differentielle Objektmarkierung und Informationsstruktur in Dialekten des Komi. Habilitationsschrift, München.
—— 2014, Identifiability, Givenness and Zero-marked Referential Objects in Komi. — Linguistics 52, 415—444.
K ü n n a p, A. 2006, Historically Problematic Morphosyntactic Features in Uralic Languages, München (LINCOM Studies in Uralic Linguistics 01).
P i ñ ó n, C. 2006, Definiteness Effect Verbs. — Event Structure and the Left Periphery. Studies on Hungarian (Studies in Natural Language & Linguistic Theory 68), Dordrecht, 75—90.
R a u n, A. 1988, Proto-Uralic Comparative Historical Morphosyntax. — The Uralic Languages. Description, History, and Foreign Influences, Leiden—New York—København—Köln, 555—571.
R é d e i, K. 1988, Geschichte der permischen Sprachen. — The Uralic Languages. Description, History, and Foreign Influences, Leiden—New York—København—Köln, 351—395.
R o u n d s, C. 1990, The Distribution of the Accusative Case in Komi Folktales. — CIFU VII. Vol. 3C, 227—232.
S c h l a c h t e r, W. 1960, Studien zum Possessivsuffix des Syrjänischen, Berlin.
S e r d o b o l s k a y a, N., U s a c h e v a, M., A r k h a n g e l s k i y, T. 2019, Grammaticalization of Possessive Markers in the Beserman Dialect of Udmurt. — Possession in Languages of Europe and North and Central Asia, Amsterdam (Studies in Language Companion Series 206), 291—311.
S t a l n a k e r, R. 1974, Pragmatic Presuppositions. — Semantics and Philosophy, New York, 197—214.
S t v a n, L. S. 2009, Semantic Incorporation as an Account for Some Bare Singular Count Noun Uses in English. — Lingua 119, 314—333.
S u i h k o n e n, P. 2005, On the Categories and Functions Developed from the Possessive and Deictic Suffixes in Udmurt. — Lihkkun lehkos! Beiträge zur Finnougristik aus Anlaß des sechzigsten Geburtstages von Hans-Hermann Bartens, Wiesbaden, 401—432.
v o n H e u s i n g e r, K. 2008, Verbal Semantics and the Diachronic Development of DOM in Spanish. — Probus 20, 1—31.
W i c k m a n, B. 1955, The Form of the Object in the Uralic Languages, Uppsala—Wiesbaden (UUÅ 6).
W i n k l e r, E. 2001, Udmurt, München (Languages of the World. Materials 212).
—— 2011, Udmurtische Grammatik, Wiesbaden (Veröffentlichungen der Societas Uralo-Altaica. Band 81).
W i t z l a c k-M a k a r e v i c h, A., S e r ž a n t, I. A. 2017, Differential Argument marking: Patterns of Variation. — The Diachronic Typology of Differential Argument Marking, Berlin (Studies in Diversity Linguistics 19), 1—40.
А л а т ы р е в В. И. 1970, Выделительно-указательная категория в удмуртском языке (Расширенный доклад на III Международном конгрессе финно-угроведов в г. Таллине), Ижевск.
Б и р ю к О. Л., У с а ч е в а М. Н. 2012, Дискурсивные факторы, влияющие на выбор между послеложной и послеложно-падежной конструкцией в бесермянском диалекте удмуртского языка (экспериментальное исследование). — Финно-угорские языки. Фрагменты грамматического описания. Формальный и функциональный подходы, Москва (Studia philologica), 607—646.
Грамматика современного удмуртского языка. Синтаксис простого предложения, Ижевск 1970.
Грамматика современного удмуртского языка. Фонетика и морфология, Ижевск 1962.

Едыгарова С. 2010, Категория посессивности в удмуртском языке, Tartu (Dissertationes Philologiae Uralicae Universitatis Tartuensis 7).
Кельмаков В. К. 1996, Формы субъективной оценки имен существительных в удмуртском языке. — CIFU VIII. Pars III, 131—134.
Кондратьева Н. В. 2000, Особенности выражения прямого объекта в удмуртском языке в зависимости от характера глагола и объекта. — Проблемы синхронии и диахронии пермских языков и их диалектов, Ижевск (Пермистика 6), 100—102.
—— 2002, Выражение прямого объекта в удмуртском языке (в историко-сопоставительном плане). Автореферат диссертации на соискание ученой степени кандидата филологических наук, Ижевск.
—— 2010, Межкатегориальные связи в грамматике удмуртского языка (на материале падежа прямого объекта), Ижевск.
—— 2011, Категория падежа имени существительного в удмуртском языке. Монография, Ижевск.
Кретов А. А. 1992, «Съедобное-несъедобное», или криптоклассы русских существительных. — Linguistica Silesiana 14, 103—114.
Кузнецова А. И. 2012, Кумуляция грамматических значений в агглютинативных показателях: дейктические функции посессива в уральских языках. — Финно-угорские языки. Фрагменты грамматического описания. Формальный и функциональный подходы, Москва (Studia philologica), 250—252.
Майтинская К. Е. 1979, Историко-сопоставительная морфология финно-угорских языков, Москва.
Сердобольская Н. В. 2017, Инвариант определенности и дифференцированное маркирование прямого дополнения в бесермянском удмуртском. — Исследования по теории грамматики. Вып. 8. Дифференцированное маркирование аргументов: материалы к типологии, Санкт-Петербург (Acta Linguistica Petropolitana. Т. XIII, ч. 3), 76—122.
Сердобольская Н. В., Толдова С. Ю. 2012, Дифференцированное маркирование прямого дополнения в финно-угорских языках. — Финно-угорские языки. Фрагменты грамматического описания. Формальный и функциональный подходы, Москва (Studia philologica), 59—142.
—— 2013, Лексические свойства глагола и оформление прямого дополнения в коми-зырянском языке (печорский диалект). — Лингвистический беспредел-2. Сборник научных трудов к 80-летию А. И. Кузнецовой, Москва, 164—175.
—— 2017, Оформление прямого дополнения в финно-угорских языках: между предикацией и дискурсом. — Ural-Altaic Studies. Урало-алтайские исследования, № 4 (27), 92—112.
Толдова С. Ю., Сердобольская Н. В. 2002, Намерения говорящего и референциальные свойства именных групп. — Computational Linguistics and Intellectual Technologies. Proceedings of the International workshop Dialogue 2002 (Protvino, June 6—11, 2002), Москва, 508—522.
Шматова М. С., Черниговская Е. А. 2012, Категория числа существительного в марийском и пермских языках. — Финно-угорские языки. Фрагменты грамматического описания. Формальный и функциональный подходы, Москва (Studia philologica), 221—250.

НАТАЛЬЯ СЕРДОБОЛЬСКАЯ (Москва)

КОРПУСНОЕ ИССЛЕДОВАНИЕ ВАРИАТИВНОГО ОФОРМЛЕНИЯ ПРЯМОГО ДОПОЛНЕНИЯ В БЕСЕРМЯНСКОМ УДМУРТСКОМ

В работе анализируется явление вариативного оформления прямого дополнения в бесермянском диалекте (наречии) удмуртского языка. Методом лингвистического анкетирования выявлено, что на выбор между аккузативом и отсутст-

вием оформления влияет референциальный статус прямого дополнения. Однако в исследованиях по литературному удмуртскому в качестве основного приводится фактор одушевленности. Для определения степени релевантности данных параметров был проведен анализ корпуса разговорных бесермянских текстов, который включает 10 539 предложения, содержащие 2187 прямых дополнений. Результаты исследования показывают следующее. Прямые дополнения, обозначающие людей, маркируются аккузативом всегда, за исключением нескольких лексических классов. Неодушевленные имена чаще выступают без аккузатива, но могут маркироваться при наличии определенного типа модификаторов.

NATALJA SERDOBOLSKAJA (Moskva)

UDMURDI KEELE BESSERMANI MURDE SIHITISKÄÄNETE KASUTAMISE KORPUSPÕHINE UURIMUS

Udmurdi kirjakeele puhul on peetud sihitiskäänete (akusatiivi ja nominatiivi) kasutamise peamiseks kriteeriumiks elusa ja elutu eristust. Et teha kindlaks, millest sõltub sihitise kääne bessermani murdes, on autor analüüsinud selle murde korpuses (http://beserman.ru/corpus/search/?interface_language=en) leidunud 2187 sihitist. Uurimistulemused näitavad, et mitmuslik sihitis on akusatiivis; inimest tähistav ainsuslik sihitis on akusatiivis, teiste elusolendite puhul oleneb akusatiivi või nominatiivi tarvitus põhiliselt sihitise referentsiaalsetest omadustest; elutut märkiv sihitis on enamasti nominatiivis, akusatiivi saab kasutada siis, kui sihitisel on laiend.