

*Pseudoincorporation analysis of unmarked Direct Objects in Mari*¹

Natalia Serdobolskaya

Russian State University for the Humanities and Moscow State University for the Humanities, Moscow

1. Introduction

The term “incorporation” is most commonly used to describe constructions where two lexical stems are compounded to form a new stem with more or less compositional meaning (e.g. *apple-picking*). In recent literature, more and more examples have been presented of constructions that resemble incorporation, but do not exactly involve stem compounding: for instance, where a caseless NP occurs adjacent to the verb and may not take any nominal inflectional markers, but may be modified by adjectives and numerals and conjoined with other NPs. Such constructions have been labelled ‘syntactic incorporation’ or ‘pseudoincorporation’ (cf. Sadock 1986 and Mithun 1984; see also Mithun 2010, where it is argued that incorporation is a continuum rather than a single clear-cut process).

Pseudoincorporation has been defined on the basis of both syntactic and semantic properties of NPs. For example, Massam (2001, 193) defines pseudoincorporation as a “grammatical phenomenon which is characterized by the existence in a sentence of a non-referential internal argument... an NP (not DP)” which is not morphologically incorporated into the verb. The criteria proposed by Dayal (2011) are absence of case marking, non-specificity, number neutrality with atelic predicates, discourse opacity and the phrasal structure of direct objects (DOs), while verb adjacency is not required. Farkas and de Swart mention non-specificity, absence of articles, discourse opacity, number neutrality, and strict verb adjacency, but allow the presence of case marking (cf. also the discussion in (Farkas and de Swart 2003, 1–6)). Espinal and McNally (2011) list absence of articles, non-specificity, and discourse opacity; for them, however, plural marking and phrasal structure in the nominal complex are acceptable.

In spite of the differences between these definitions, most authors consider non-specificity to be one of the diagnostic features of pseudoincorporation. In the present article, however, on the basis of data from Eastern Mari (Finno-Ugric, Uralic), I claim that pseudoincorporation can vary in its semantic properties (i.e. it does not always correlate with non-specificity). The constructions I consider are morphosyntactically close to other examples of pseudoincorporation discussed in the literature. Nonetheless, as I show below, their semantics is not restricted to non-specificity: pseudoincorporated NPs in Mari can be specific and even definite. Hence, I claim that either the notion of pseudoincorporation should be broadened (allowing for a wider range of semantic types) or another term (e.g. quasi-incorporation) is to be introduced; the former option appears to be preferable.

The data from Eastern Mari presented here were collected in the course of fieldwork in the village of Staryj Torjal (Novotorjalskij district of the Republic of Mari El, Russia; approx. 20 native speakers examined) in the years 2000, 2001 and 2004. All the examples given without references are drawn from this fieldwork.

The work is structured as follows. In the first section I consider the typical morphosyntactic and semantic properties of pseudoincorporation as they have been described in the literature. In the second section I provide the Eastern Mari data and show that in this language unmarked DOs possess the morphosyntactic properties of pseudoincorporation, while semantically they differ from other cases of pseudoincorporation described in the literature. The discussion and syntactic analysis are provided in section 3.

¹ I thank Irina Nikolaeva, Malte Zimmermann, and especially Svetlana Toldova for discussion of this phenomenon in Mari. This work has been supported by the RSUH Development program, and RHF grants No 12-34-01345 and 11-04-00282.

1. Unmarked Objects and (Pseudo)Incorporation

In this section I consider the morphosyntactic and semantic properties of pseudoincorporation as they have been described in the typological and theoretical literature. I compare the properties of constructions with Differential Object Marking, incorporation, and pseudoincorporation, and attempt to define diagnostic features of pseudoincorporation.

1.1. Morphosyntactic properties

The phenomenon of Differential Object Marking (where the language offers more than one means of encoding direct objects; DOM) has received a great deal of attention in the linguistic literature since Moravcsik (1978), Bossong (1985) (see also Aissen 1998, de Swart 2007, Dalrymple and Nikolaeva 2011 among others). Systems that demonstrate DOM include those where the direct object can bear a case marker or remain unmarked:

Mandarin Chinese (Thompson 1972, cited in Moravcsik 1978, 279)

(1) a. ta mài le tāde chē.

he sell ASPECT² his car

‘He sold his car.’

b. ta bă tāde chē mài le.

he ACC his car sell ASPECT

‘He sold his car.’ (‘car’ being the topic)

This type (also termed asymmetric DO encoding) is observed in Indo-European, Altaic languages, some Semitic and Malayo-Polynesian languages, and some Uralic languages, e.g. Mari (Galkin 1964, Tužarov 1998). For some languages it has been claimed that the unmarked Direct Object in the asymmetric type is incorporated into the verbal complex, cf. Pearce (2001) for Iaa and Fijian (Oceanic), Mithun (1984) for Mam and Kanjobal (Mayan), Lahu (Tibeto-Burman) etc. In most of these languages the unmarked DO must be verb-adjacent and does not take modifiers or nominal grammatical markers of any kind (Mithun 1984, 856).

Noun Incorporation is defined as “a construction where a nominal stem is compounded with a verbal stem to yield a larger, derived verbal stem” (Mithun 1984, 847), whence the inability to take nominal grammatical markers (including case) and modifiers and the requirement of verb adjacency. According to the analysis of incorporation by Baker (1988), it is the N⁰ that is incorporated in these cases. Hence the absence of case marking on the DO, and the resulting DOM phenomenon in these languages.

However, constructions with asymmetric DOM do not have the same morphosyntactic properties in all languages. In Mari and Turkish (see Muravyova 2008, von Heusinger, Kornfilt 2005 for Turkish) unmarked DOs must be verb-adjacent (in Mari the unmarked DO and the verb are also pronounced without a pause), but they can occur with adjectives, plural endings, numerative expressions (e.g. *two bottles of milk*) etc. Thus, they show phrasal structure. In Komi-Zyryan and Erzya-Mordvin, unmarked

² List of abbreviations in glosses: ACC – accusative; ADD – additive particle; ADV – adverbial suffix; ASPECT – aspectual marker; ATTR – attributive suffix; CAUS – causative; COMPL – complementizer; CONV – simple converb; CONV.ANT – converb of anteriority; DAT – dative; DESID – desiderative mood; EL – elative; EMPH – emphatic particle; GEN – genitive; INDEF – prefix of indefinite pronouns; INESS – inessive; INF – infinitive; LAT – lative; NARR – narrative tense; NEG – negation; NZR – nominalizer; OPT – optative; P.1/2/3SG/PL – possessive first/second/third person singular/plural; PL – plural; PRS – present tense; PRT – preteritum; PTCP.ACT – active participle; Q – question marker; REFL – reflexive; SG – singular.

DOs, as well as being able to take modifiers and nominal grammatical markers, can occur non-adjacent to the verb. Hence, the relevant DO construction in these languages cannot be described as incorporation: it fits neither of the types postulated in Mithun (1984).

For some languages with this type of DOM the term “pseudoincorporation” has been proposed, e.g. for Turkish (Muravyova 2008), Niuean (Massam 2001), Hindi (Dayal 2011), and Sakha and Tamil (Baker 2012). In these languages the construction with the unmarked DO shows some of the properties associated with incorporation; however, the DO can have nominal inflection and take modifiers. In Turkish, unmarked DOs must be verb-adjacent and do not take possessive markers; they must be non-specific and have low discourse relevance, as is characteristic of incorporation (Mithun 1984). However, they can take plural endings, adjectives, numerals, numerative expressions and even relative clauses. Note that Mithun (1984, 872–873) places the Turkish case, together with Hungarian, on the list of possible diachronic “origins” of pure incorporation. Muravyova uses the term “quasi-incorporation” for Turkish.

In Niuean, Sakha and Tamil strict verb adjacency is also observed, while the grammatical and syntactic features of pseudoincorporated DOs differ. The Niuean construction allows adjectives, numerative modifiers, prepositional groups and relative clauses in the pseudoincorporated NP; the construction has a non-specific interpretation (Massam 2001). In Sakha and Tamil, unmarked DOs take adjectives and are number-neutral (Baker 2012).

In Hungarian strict verb adjacency is observed, and unmarked DOs take no articles, show number neutrality and are non-specific and discourse opaque; however, case marking is present (Farkas and de Swart 2003). In Hindi verb adjacency is not required, unmarked DOs can take adjectives, and nouns with adjectives can occur conjoined in the DO position, which demonstrates that they have phrasal structure. However, unmarked DOs are number-neutral with atelic predicates, non-specific and discourse opaque, thus leading to the pseudoincorporation analysis (Dayal 2003; 2011).

The phenomenon under discussion, then, seems to be rather vaguely defined, since the morphosyntactic properties of pseudoincorporated DOs are quite different for Turkish, Hindi, Niuean, Sakha and Tamil. What all these constructions do have in common from the morphosyntactic point of view is asymmetric DOM, phrasal status of the unmarked DO (which makes them different from incorporation), which is, however, reduced to some extent (e.g. determiners are not allowed), and the absence of some nominal grammatical categories (e.g. possessive markers in Turkish, cf. Muravyova 2008, plural marking etc.).

Let us consider the properties of incorporation, pseudoincorporation, and unmarked DOs in those languages where they cannot be analysed as (pseudo)incorporation under any known definition of this term. In Komi-Zyryan and Erzya-Mordvin, the DO can occur unmarked for case; however, it can take nominal inflectional markers, does not require verb adjacency, allows conjoined groups, takes modifiers of various types, and is not necessarily non-specific or discourse opaque (cf. Toldova and Serdobolskaya 2012a).

Table 1. Morphosyntactic and semantic properties of incorporation, pseudoincorporation and asymmetric DOM

Features	Incorporated DOs	Pseudoincorporated DOs	Unmarked DOs (Komi-Zyryan, Erzya-Mordvin)
Verb adjacency	obligatory	obligatory/non-obligatory	non-obligatory
Adjectival modifiers,	–	+	+

participles, etc.			
Pronouns	–	–	–
Proper nouns	–	–	–
Plural marking	–	+/-	+
Possessive marking	–	–	+/-
Coordination	–	+/-	+
Semantics	non-specific	non-specific	no general restrictions

Table 1 shows that there are relatively few differences between constructions with pseudoincorporation and unmarked DOs. They comprise the following: obligatoriness of verb adjacency, acceptability of coordinated NPs and plural and possessive marking, and semantics. However, even where these properties are concerned the constructions may be very similar (there are no sharp contrasts between the columns, cf. +/- signs in the relevant rows). By contrast, morphological incorporation is very different from pseudoincorporation, since the incorporated DO is a head, not a phrase, and cannot have any dependents or host nominal inflectional markers. However, this only concerns prototypical incorporation; as shown in Mithun (2010) and van Geenhoven (1998), these restrictions may be violated.

Thus, establishing whether a given language has pseudoincorporation is a task with no straightforward solution. While analysing the constructions found in Mari, I will take as my diagnostics verb adjacency and the phrasal status of unmarked DOs (including whether their phrase structure is somehow reduced compared to that of case-marked DOs).

1.2. The semantics of unmarked DOs and pseudoincorporation

Those studies that define pseudoincorporation on the basis of semantics consider the limited referential potential of pseudoincorporated NPs as one of the basic properties of this construction: in Turkish, West Greenlandic, Niuean, Hungarian and Hindi, unmarked DOs are non-specific and cannot have a wide scope interpretation (Enç 1991³; van Geenhoven 1998; Massam 2001; Farkas and de Swart 2003; Dayal 2011). Many of the morphosyntactic properties given in section 1.1 are explained on the basis of this semantic restriction. For example, the impossibility of pseudoincorporating proper nouns, pronouns and nouns with determiners follows from the unacceptability of pseudoincorporated definite NPs. Possessive markers and restrictive relative clauses can also function as semantic restrictors.

The property of number neutrality may follow from non-specificity and scopal effects. Many contexts that contribute to a non-specific reading may also trigger plural neutrality. For example, in certain iterative and habitual contexts singular and plural in English have the same meaning, e.g. (Krifka et al., 1995, 24) argues that the sentence *A cat runs across my lawn every day* can be interpreted as *Cats in general run across my lawn every day*; see also Dayal's explication of the Hindi data: "due to the fact that they modify the meaning of the predicate itself and their scopal effects are local to the predicate, singular implicatures are obfuscated in iterative and habitual contexts" (Dayal 2011). A similar effect is observed in certain modal contexts (cf. Krifka et al., 1995, 24): *Shirts must be worn. (Sign at a restaurant entrance)* may be interpreted as *If you want to enter this restaurant, you must wear a shirt*. It has been argued in the literature on genericity that in English characterising sentences (Krifka et al., 1995, 78: *Dutchmen are good sailors / The Dutchman is a good sailor*) the semantic difference between singular NPs and plural NPs is often a subtle one, cf. Krifka et al. (1995).

³ However, see the critique in von Heusinger and Kornfilt (2005).

Hence, many of the morphosyntactic restrictions listed in 1.1 may be due to the non-specific semantics of pseudoincorporation. However, there are languages that show similar morphosyntactic restrictions, but do not display the same semantic properties of pseudoincorporation as in Hungarian, Hindi etc. One of these is considered in this paper: in Mari unmarked DOs can be specific and even definite, and the choice of marking is based on information structure and discourse strategies, cf. section 2.3.

It is noteworthy that even pure incorporation does not necessarily require non-specificity; in some languages specific or definite NPs can be incorporated. For example, Mithun (2000) gives examples from Southern Tiwa (Kiowa-Tanoan), Tuscarora (Iroquoian) and Mohawk (Iroquoian).

This paper is focused on a similar situation in Eastern Mari. I demonstrate that Mari unmarked DOs show the syntactic properties of pseudoincorporation, but do not necessarily have non-specific semantics. This means that we must either admit the existence of different semantic types of pseudoincorporation or postulate a new type of construction requiring a different label (e.g. quasi-incorporation).

2. Unmarked DOs or “pseudoincorporation” in Eastern Mari

Like many other Uralic languages, Eastern Mari displays the phenomenon of Differential Object Marking (DOM) (cf. Wickman 1955, Comrie 1988, 471–472, Havas 2008 etc.). Mari DOM belongs to the asymmetric type (marked vs. non-marked DOs): there is an accusative case marker which can be omitted in certain contexts:

- (2) **kütü** / **kütü-m** č'umər-aš məj pij-lan šüškalt-əš-əm.
 herd / herd-ACC gather-INF I dog-DAT whistle-NARR-1SG
 ‘To gather the herd I whistled to the dog.’

As shown in 2.1 and 2.2, the constructions with unmarked DOs in Mari have some of the features of incorporation (no accusative case marking, prosody, strict verb adjacency, a ban on pronouns and restrictive relative clauses, etc.). However, certain features prevent us from analysing this construction in terms of incorporation (morphological or lexical compounding), since unmarked DOs can take plural markers and possessive markers and modifiers of various types. It is noteworthy that the construction under examination is not limited to a fixed set of verbs, meaning that its acceptability does not depend on the individual selectional properties of the verb (unlike most examples of lexical compounding, cf. Mithun 1984).

In reference grammars and specialised articles on Mari DOs, definiteness is used to explain the choice of DO marking, cf. Galkin (1964, 45), Tužarov (1984), Anduganov (1991, 71). Indeed, in some examples definiteness seems to be the main factor:

- (3) **rvez-əm** šel-m-əž-lan pet'a-m č'ot vurs-en-ət.
 boy-ACC hit-NZR-P.3SG-DAT Peter-ACC a.lot scold-PRT-3PL
 ‘Peter was scolded a lot for hitting the boy.’

- (4) **rveze** šel-m-əž-lan pet'a-m č'ot vurs-en-ət.
 boy hit-NZR-P.3SG-DAT Peter-ACC a.lot scold-PRT-3PL
 ‘Peter was scolded a lot for hitting a boy.’

However, it is shown in Toldova and Serdobolskaya (2002) and Toldova and Serdobolskaya (2012a, 79–89) that definiteness is not the crucial factor regulating the choice of the DO marker. In

Mari the information structure of the sentence is more important. The semantics of DOM in Mari is considered in detail in section 2.3. In the next section I analyse the prosodic and morphological properties of DOM.

2.1. Prosodic and morphological properties of unmarked DOs

Unmarked objects in Mari must be verb-adjacent and cannot appear with some of the modifiers that can be used with case-marked DOs. This suggests an incorporation or pseudoincorporation analysis for these constructions. Mithun (1984) proposes to distinguish between four types of noun incorporation: (1) lexical compounding: the incorporated noun is non-specific, does not take any nominal inflectional markers or modifiers (a subtype of this class is morphological compounding, where the N + V complex behaves as a single word morphophonologically); (2) “manipulation of case”: the DO is incorporated, and the verb takes another DO; the original DO does not take any nominal inflectional markers or modifiers, but is not necessarily non-specific; (3) “manipulation of discourse structure”: this shows the morphosyntactic properties of one of the first two types, while its function is to present the DO as background information; (4) classificatory noun incorporation: this also shows the morphosyntactic properties of one of the first two types, but here the incorporated DO with more general meaning serves as a classifier for a more specific external DO.

Hence, in terms of the morphosyntactic properties of the incorporated DO, the following types can be distinguished: morphological compounding (incorporation at the N⁰ level; DO + V make up a single word both morphophonologically and prosodically), lexical compounding (composition by juxtaposition in Mithun (1984): incorporation at the N⁰ level; DO + V are two separate words) and NP incorporation (incorporated nouns take modifiers, numerals etc.), cf. Mithun (2010). The first two types involve incorporation at the N⁰ level. It will be shown below that the unmarked DO in Mari does not belong to either of these types.

The prosody of the DO + V complex seems to work as an argument in favour of the compounding analysis. Most often, there is no pause between the unmarked DO and the verb, or else a pause of only 0.01–0.02 seconds. DOs with the accusative, if adjacent to the verb (see the discussion of (16) and further for verb adjacency) are often followed by a pause. I analysed 6 spontaneous texts (dialogues and narratives) recorded from 7 native speakers, three of them about 20 years old (with faster speech) and the other four from 50 to 80 years old. Only cases of verb-adjacent DOs have been included, and only nouns have been analysed. From 104 examples, the following results are observed. For constructions with unmarked DOs the average pause between the DO and the verb is 0.01 seconds, while for constructions with the accusative DO it is 0.04 seconds. Hence, it can be argued that unmarked DOs are prosodically united with the verb, while accusative DOs are treated as separate words.

Thus, unmarked DOs and the verb are prosodically treated as a single unit. However, the unmarked DO does not form a compound with the verb. The first two arguments for this contention are based on the morphophonology of the DO + V complex.

1. Mari requires vowel harmony for some types of suffixes, including possessive markers. For example, the third person possessive suffix can take the form *-že/-žo/-žö* depending on the frontness and roundedness of the stressed vowel of the stem (Alhoniemi 1993, 41):

- (5) *ergə-že* – *kogəl'-žo* – *üdər-žö*
 son-P.3SG cake-P.3SG daughter-P.3SG
 ‘his/her son’ – ‘cake’ – ‘daughter’

This vowel harmony rule is not observed in the DO + V complex. The possessive marker on the verbal form is harmonized with the verbal stem, and not with the DO (6); the possessive marker on the DO is harmonized with the DO only (7).

- (6) **mənam brükö urg-əm-əžo kelš-a.**
 I.DAT trousers sew-NZR-P.3SG like-PRS.3SG
 ‘I like sewing trousers.’

- (7) **pet’a-n kid-še muš-m-əž-əm məj už-əm.**
 Peter-GEN hand-P.3SG wash-NZR-P.3SG-ACC I see-NARR.1SG
 ‘I saw Peter wash his hands.’

Hence, the verb and the DO are treated as two separate words with separate stress patterns.

2. If the DO + V complex were an example of morphological/lexical compounding, it would display the properties of compounds in Mari. Mari possesses several types of compounds, which differ in their morphophonology. Old compounds show assimilation of the final vowel of the first stem and the initial consonant of the second stem, cf. Lavrentev (1975). The mid vowels *e*, *o*, *ö* are reduced to *ə*, while the initial consonant of the second word becomes voiced in intervocalic position:

- (8) **šül’ö + peč’e → šül’əveč’e** (Lavrentev 1975, 247), cited in (Shibasova 2006)
 ‘oats’ ‘fence’ ‘a field of oats’

No such changes are attested in the verb + unmarked DO complex:

- (9) **poŋgo pog-aš**
 mushroom collect-INF
 ‘to collect mushrooms’

However, this argument is rather weak, since not all compounds obey this rule.

3. Unmarked DOs can be coordinated with a conjunction:

- (10) **poŋgo den saska pog-aš kaj-em.**
 mushroom and fruit collect-INF go-PRS.1SG
 ‘I go / I am going to collect mushrooms and fruit.’

This shows that the DO + V complex does not form a lexical unit.

4. The fourth argument concerns grammatical features of unmarked DOs. The Mari noun is inflected for number (singular vs. plural), case, and possessivity (by means of markers that agree in person and number with the possessor). The unmarked DO in Mari takes plural markers (11) and possessive suffixes (12).

- (11) **tuvər-vlak nal-aš**
 shirt-PL take-INF
 ‘to buy shirts’

- (12) **pet’a-n kid-še muš-m-əž-əm məj už-əm.**
 Peter-GEN hand-P.3SG wash-NZR-P.3SG-ACC I see-NARR.1SG
 ‘I saw Peter wash his hands.’

Examples with possessive suffixes are often rejected by native speakers; however, they are found in texts, elicited examples (not focused on the DO marking) and specialised linguistic studies (Tužarov 1984, 286).

Therefore, on the basis of prosodic and morphological properties of unmarked DOs, the Mari construction cannot be analysed as incorporation at the word level.

2.2. Syntactic properties

Omission of the DO marker in Mari is restricted by a number of rules. First, it is only possible in dependent non-finite clauses⁴ (cf. Galkin 1964, 85), cf.:

- (13) a. č'odəra-šte kaj-že da **poŋg-əm** / * **poŋgo** pog-əžo.
 forest-INESS go-OPT and mushroom-ACC mushroom collect-OPT
 'Let him go to the forest and collect mushrooms.'

- b. məj č'odəra-š **poŋgo** pog-aš kaj-em.
 I forest-LAT mushroom collect-INF go-PRS.1SG
 'I am going to the forest to collect mushrooms.'

The same DO + verb complex is observed in these examples. In (13a) omission of the accusative is unacceptable because the verb is in one of the oblique moods (optative) and heads an independent clause, while in (13b) it is acceptable because the verb occurs in the infinitival form and heads a dependent clause. Both the conditions of non-finiteness and dependent status of the clause are necessary: the DO marker cannot be omitted if the same infinitival form heads an independent sentence:

- (14) mo-lan **omsa-m** / ***omsa** petər-aš?
 what-DAT door-ACC door close-INF
 'Why (should I) close the door?'

Moreover, the accusative cannot be omitted in a finite dependent clause:

- (15) iza-m teŋgeč'e **pir-əm** / ***pire** už-ən-am manən ojl-en.
 brother-P.1SG yesterday wolf-ACC wolf see-PRT-1SG COMPL say-PRT(3SG)
 'My brother said that he had seen a wolf yesterday.'

A diachronic explanation for this fact, based on semantics, is given by specialists on the history of the Mari language (cf. Galkin 1964, 46). Tužarov (1998, 122) argues that constructions with caseless DOs predate those with the accusative. Therefore, the behaviour of non-finite clauses is more archaic than that of finite constructions. I do not intend to provide a synchronic explanation here. However, observe that the situation in which incorporation is limited to non-finite forms, or more specifically to nominalizations (cf. Koptjevskaya-Tamm 1993: 184–191), is typologically widespread.

There are a number of other syntactic properties that speak in favour of the pseudoincorporation analysis of unmarked DO constructions:

1. *Verb adjacency*. The unmarked DO must be verb-adjacent, cf.:

⁴ Tužarov (1986) gives exceptions to this rule from Standard Eastern Mari. However, he shows that they are limited to DOs with possessive suffixes of 1st and 2nd person, unlike the constructions analysed in this paper.

- (16) a. rvez-ən motor-ən **poč'elamut** lud-m-əž-lan
 boy-GEN beautiful-ADV poem read-NZR-3SG-DAT
 tun-əkt-əšo kuan-en.
 learn-CAUS-PTCP.ACT rejoice-PRT(3SG)
- b. rvez-ən **poč'elamut-əm** / *poč'elamut motor-ən lud-m-əž-lan ...
 boy-GEN poem-ACC poem beautiful-ADV read-NZR-P.3SG-DAT
- 'The teacher is glad that the boy recites poems well.'

This rule is strict and can never be violated.⁵ Basic word order in Mari is SOV, and in NPs all modifiers precede the head. The elements that could possibly intervene between the verb and direct object are particles, adverbs, case-marked arguments, postpositional phrases and non-finite components of complex predicates; the subject can follow the direct object if the latter is a topic. Adverbs, case-marked arguments and postpositional phrases are not allowed between the verb and a caseless DO, cf. (16b) with an adverb. The unmarked DO can, however, host enclitic particles, e.g. the question particle *mo*:

- (17) **kinde** mo nal-meke keč'əval koč'kəš-əm jamdəl-en-da?
 bread Q buy-CONV.ANT afternoon food-ACC prepare-PRT-2PL
- 'Did you start cooking lunch after you had bought (the) bread (or later, or earlier)?'

Complex verbs in Mari cannot be separated by the DO, regardless of the presence/absence of the accusative: the non-finite component is strictly adjacent to the second component. The DO is always left-adjacent to the former, cf. the following example with the complex verb 'to find', made up of the stem of the verb 'to look for' and the conjugated form of the verb 'to find':

- (18) **pərəs** kəč'al mu-mek, möngə-že kond-əš.
 cat search find-CONV.ANT home-P.3SG bring-NARR.3SG
- 'When he had found the cat, he brought it home.'

However, I do not consider this an example of separation of the DO + V complex, since complex verbs in Mari are lexicalized; the two parts of the complex verb cannot be separated by any material.

Postpositional phrases describing the resultative phase of the situation must also occur before the DO (unlike in Sakha and Tamil, cf. Baker 2012) (19a). The DO with the accusative may precede or follow the postpositional phrase (19b, c).

- (19) a. ava joč'a-n küvar ümbak **č'əla kniška** kəšk-əm-əž-əm už-ən.
 mother child-GEN floor on all book throw-NZR-P.3SG-ACC see-PRT(3SG)
- 'Mother saw that the child had thrown all the books on to the floor.'
- b. ava joč'a-n **kniga-m** küvar muč'ko kəšk-en pətar-m-əž-əm už-ən.
 mother child-GEN book-ACC floor on throw-CONV finish-NZR-P.3SG-ACC see-PRT(3SG)
- c. ava joč'a-n küvar muč'ko **kniga-m** kəšk-en pətar-əm-əž-əm už-ən.
 mother child-GEN floor on book-ACC throw-CONV finish-NZR-P.3SG-ACC see-PRT(3SG)
- a.=b.=c.

⁵ However, see exceptions for Standard Eastern Mari in Tužarov (1986).

Therefore, strict verb adjacency is observed: no material (except enclitic particles) can intervene between the head noun and the verb. This can be easily explained if it is assumed that unmarked Mari DOs are incorporated into the verb.

2. *Attributes of the DO.* The unmarked DO can only take a limited set of modifiers. Restrictive relative clauses (20), universal quantifiers (21) and demonstratives (22) are unacceptable.

(20) *teŋgeč'e nal-me ü-əm / *ü kond-aš küšt-əš-əm.*
 yesterday buy-NZR milk-ACC milk bring-INF order-PRS.3SG
 'I asked (you) to bring the milk that we bought yesterday.'

(21) *keč'-mogaj kajək-əm / *kajək kuč'-aš lij-eš.*
 every-what.kind.of bird-ACC bird catch-INF need-PRS.3SG
 'Every bird can be caught (lit. it is possible to catch every bird).'

(22) *joč'a-vlak ač'a-ž-ən tide pört-əm / *pört əšt-əm-əž-əm pal-at.*
 child-PL father-P.3SG-GEN this house-ACC house make-NZR-P.3SG-ACC know-PRS.3PL
 'The children know that their father built this house.'

It could be speculated that this restriction arises due to the semantics of demonstratives (if we assume that Mari unmarked DOs are indefinite); however, indefinite pronouns are also unacceptable with unmarked DOs:

(23) *vas'a ala-mogaj istorij-əm / *istorij kalaskal-aš šon-en...*
 Vasja INDEF-what.kind.of story-ACC / story tell-INF think-PRT(3SG)
 'Vasja wanted to tell a story.' [I do not know what kind of story.]

Genitive modifiers are also rejected, even if they do not serve as presupposition inducers, as in the following:

(24) a. *keč'əval koč'kəš-lan kijar pog-aš kül-eš.*
 afternoon meal-DAT cucumber collect-INF need-PRS.3SG
 '(We) need to collect some cucumbers for lunch.'

b. **keč'val koč'kəš-lan poškud-ən kijar pog-aš kül-eš.*
 afternoon meal-DAT neighbour-GEN cucumber collect-INF need-PRS.3SG

Intended meaning: '(We) need to collect some of the neighbour's cucumbers for lunch.'

The genitive possessor in this example does not make the DO definite, since only a part of the set would be involved in the situation; however, the accusative cannot be omitted.

Adjectives can sometimes function as presupposition inducers in a specific context:

(25) *təj de-č'-et joškar olma-m / olma pu-aš jod-ən-am.*
 you near-EL-P.2SG red apple-ACC apple give-INF ask-PRT-1SG
 {There are two apples on the table.} 'I asked you to give me the red apple.'

However, even in such contexts it is possible to omit the accusative.

Thus, the hypothesis of presupposition triggers should be rejected. Another possible analysis for Mari is incorporation: only a head noun can be incorporated and, hence, occur in caseless form.

However, the unmarked DO can be modified by adjectives, cf. (25). Adjectives that do not serve as presupposition triggers are allowed as well, cf. (26) with a non-derived adjective and (27) with an adjective derived via an attributive marker. Dependent NPs are also allowed (28).

(26) **təlat kande tuvər urg-aš kül-eš.**
 you.DAT blue dress sew-INF need-PRS.3SG
 ‘You need to sew a blue dress (for yourself).’

(27) **ač’a-že erg-əž-əm tač’-əse gazet nal-aš kolt-en.**
 father-P.3SG son-P.3SG-ACC today-ATTR newspaper take-INF send-PRT(3SG)
 ‘Father sent his son to buy today’s newspaper.’

(28) **paša-jeŋ-vlak-ən tud-ən oškəl-mo jük kol-m-əm vujlatəše uməl-en.**
 work-person-PL-GEN he-GEN step-NZR sound hear-NZR-ACC boss guess-PRT(3SG)
 ‘The boss guessed that the workers had heard his footsteps’. (lit. the sound of his coming)

Cf. the availability of modifiers with incorporated DOs in West Greenlandic (van Geenhoven 1998) and with pseudoincorporated DOs in Turkish (Muravyova 2008). However, in West Greenlandic the modifiers of the incorporated nominal take a special case marker and obviously do not form a constituent with their semantic head, cf. van Geenhoven (1998). Conversely, in Mari the modifiers of the unmarked DO preserve all the morphosyntactic properties they have with case-marked heads: head-adjacent position, no case/number/possessive marking.

It can be concluded that unmarked DOs in Mari are not heads, but phrasal categories. They can take adjectives of various kinds and dependent NPs, thus they are at least NPs. They apparently do not have the D level, since determiners of various kinds cannot be pseudoincorporated. The unacceptability of relative clauses is easily accounted for on the assumption that they are syntactic complements of the determiner head of DP (while all kinds of relatives are presented as head-internal, cf. Zwart 2000, partly based on Vergnaud 1974).

3. *Type of the NP head.* Presupposition-inducing nouns (cf. Geurts 1998), as well as pronouns and proper names with definite reference, cannot head unmarked DOs:

(29) **məjə təj-əm / *təj kəč’al-aš kaj-em.**
 I you-ACC you search-INF go-PRS.1SG
 ‘I’m going to search for you.’

Native speakers marginally accept such examples with proper names:

(30) **jəvan kəč’al-še rveze**
 Ivan search-PTCP.ACT boy
 ‘the boy who was looking for Ivan’

It has been argued that pronouns and proper names are raised to D (Longobardi 1994). This explains their unacceptability in pseudoincorporated DOs if these do indeed lack the D level.

Nouns with possessive markers etc. may head unmarked DOs (12). As discussed above, possessive markers can serve as presupposition inducers in some cases, though not always. The semantics of such examples will further be commented upon in 2.3.

The following shows that while unmarked DOs in Mari are not DPs, they seem to possess the NumP level.

4. *Quantifiers and numerals*. Unmarked DOs can host quantifiers and numerals:

- (31) **šuko olma-m / olma** kudalt-aš logal-eš.
many apple-ACC apple throw-INF have.to-PRS.3SG
'We had to throw away many apples.'

Hence, syntactic properties of unmarked DOs in Mari are incompatible with an incorporation analysis. However, they possess the morphosyntactic properties of pseudoincorporation: they must be verb-adjacent, and they have reduced phrasal structure. It can be assumed that they are NPs or NumPs, lacking the QP and the DP levels, since they cannot host universal quantifiers and determiners, including demonstrative, indefinite, and interrogative pronouns.

2.3. Semantics

Most reference grammars argue that definiteness is important for the choice of DO encoding in Mari (Galkin 1964, Anduganov 1991); see the critique in Tužarov (1984, 1986). Some specialists informally describe the semantic opposition of the accusative vs. caseless DO in terms of “singling out the DO” with the accusative vs. “singling out the action itself” without the accusative (Pengitov 1961, 67). Toldova and Serdobolskaya (2002) show that it is the information structure of the sentence that is crucial for the choice of DO encoding in Mari. In this section, I consider definiteness, specificity, and information structure in constructions with unmarked DOs, and compare the Mari constructions to pseudoincorporation with respect to the semantics.

It must be noticed that the unmarked DO construction is not limited to a fixed set of verbs. To the best of my knowledge, every transitive verb can take an unmarked DO under certain semantic and pragmatic conditions. Hence, the acceptability of the unmarked DO does not depend on the individual selectional properties of the verb (unlike in most examples of lexical compounding, cf. Mithun 1984). It is noteworthy that unmarked DOs do not serve as a regular base for creating idioms with the verb. This is another important point that differentiates Mari unmarked DOs from incorporated NPs (cf. Mithun 2000).

2.3.1. Referential properties of unmarked DOs

For many languages with (pseudo)incorporation, it has been shown that (pseudo)incorporated DOs cannot have a wide scope interpretation (cf. Enç 1991 for Turkish⁶; van Geenhoven 1998 for West Greenlandic; Massam 2001, type 1 in Niuean; Farkas and de Swart 2003 for Hungarian, Dayal 2011 for Hindi). However, in Mari, as well as some incorporating languages, (cf. Mithun 2000), this restriction does not hold.

The narrow scope interpretation is the most frequent:

- (32) a. **vaza šənd-əme üstel**
vase set-NZR table
'a/the table to put vases on'

⁶ However, see the critique in von Heusinger, Kornfilt (2005).

b. **vaza-m** šənd-əme üstel
 vase-ACC set-NZR table
 ‘a/the table a vase has been put on’

In (32a) the DO *vaza* is indefinite non-specific, and occurs without the accusative marker. On the other hand, in (32b) the DO with the accusative can be interpreted as specific or definite. The same indefinite non-specific interpretation of the unmarked DO is observed in the next example, where the DO is under the scope of a frequency adverb:

(33) a. rvez-ən ere **poč’elamut** tunem-m-əž-lan tun-əkt-əšo kuan-en.
 boy-GEN always poem learn-NZR-P.3SG-DAT learn-CAUS-PTCP.ACT rejoice-PRT(3SG)
 ‘The teacher is glad that the boy always learns poems well.’

It is well known that frequency adverbs may contribute to a narrow scope reading of bare NPs. Indeed, such an interpretation is observed in (33a). However, there is no strict ban on wide scope interpretation of the unmarked DO:

b. rvez-ən **poč’elamut** tunem-m-əž-lan...
 boy-GEN poem learn-NZR-P.3SG-DAT

‘The teacher is glad that the boy has learnt a poem.’ (a. some poem the boy has chosen himself. b. the one that the teacher told him to learn the day before).

Consider another example of a specific indefinite NP, unmarked in DO position:

(34) iza-m teŋgeč’e **pire** už-mə-ž nergen ojl-en.
 brother-P.1SG yesterday wolf see-NZR-P.3SG about say-PRT(3SG)
 ‘My brother said that he had seen a wolf yesterday.’

In (34) the NP *pire* ‘wolf’ is specific, since the sentence can be followed by ‘He (my brother) shot at **it** (the wolf)’. Strict identity reading is perfectly acceptable in ellipsis contexts:

(35) məj teŋgeč’e pet’a-n **üdər** šupšal-m-əž-əm už-ən-am, tač’-at už-ən-am.
 I yesterday Peter-GEN girl kiss-NZR-P.3SG-ACC see-PRT-1SG today-ADD see-PRT-1SG
 ‘I saw Peter kiss a girl yesterday, and today, too. (the same girl)’

The caseless DO in this example can be interpreted as referring to the same girl in both parts of the sentence, and hence, allows a specific reading. A non-specific reading would lead to sloppy identity (cf. Dalrymple et al. 1991), i.e. different girls every day.

The unmarked DO can also be definite, as in the following examples. For example, it can refer to an aforementioned discourse referent (the example contains two sentences taken from a narrative; the second sentence contains a DO referring to an antecedent which is in the first sentence):

(36) tide istorij-že takšəm ške odnoklassnik-šaməč’ kuze vašlij-me nergen da.
 this story-P.3SG so REFL.GEN classmate-PL how meet-NZR about then
 ‘This story is about a meeting of my classmates.’ {Two more sentences.}

vot... pogən-en-na əle kutər-en kelš-en-na **odnoklassnik-šaməč'** vašlij-aš.
 well... get.together-PRT-1PL be.3SG talk-CONV agree-PRT-1PL classmate-PL meet-INF
 'Well... we got together and arranged to meet our classmates.'

Unique objects can also be unmarked in DO position:

(37) **keč'-əm / keč'e** onč'-aš jörat-aše jeŋ-vlak er kən'el-ət.
 sun-ACC sun look-INF love-PTCP.ACT person-PL early get.up-PRS.3PL
 'People who like to meet the sun get up early.'

Some native speakers marginally allow proper names (30) to be unmarked in DO position. DOs with possessive markers may also occur without the accusative, although rather rarely. It is well known that possessive markers can function as semantic restrictors and require markers of definiteness in many languages. However, they do not make possible a unique identification of the referent in all cases, e.g. *my hand*, *my pencil* when there are many pencils belonging to the speaker (in contrast to contexts where the object is unique for the possessor, e.g. *my head*, *her mother*), cf. Poesio (1994). Assuming a purely semantic account for Mari, it would not be surprising to find unmarked DOs in cases where the possessive relation does not suffice to identify a particular object. However, it is also possible in examples like (12). The context of "Peter washing his hands" contributes to the unique interpretation, since both Peter's hands are referred to (although the DO does not bear a plural marker), and Peter refers to an anchored element in the discourse. However, the accusative is omitted even in this context.

On the contrary, the accusative-marked DO can also take narrow scope:

(38) **urem-ašte pərəs-əm** už-aš saj-lan ogəl.
 street-INESS cat-ACC see-INF good-DAT NEG.PRS.3SG
 'It is a bad sign to see a cat / cats in the street.'

It has been argued that a partitive context (like *two slices* in *Mother bought bread, and I ate two slices*) gives rise to a specific interpretation (Enç 1991). The following shows that the accusative marker may be omitted in such contexts, too:

(39) **kok banka kompot-əm / kompot** koč'-m-em-lan
 two pot stewed.fruit-ACC stewed.fruit eat-NZR-P.1SG-DAT
ava-m məj-əm vurs-en.
 mother-P.1SG I-ACC scold-PRT(3SG)
 'Mother scolded me for eating two pots of stewed fruit (from our reserve at home).'

Dayal (2011) proposes a number neutrality test for pseudoincorporation, giving examples from Hindi where morphologically singular DOs can be interpreted both as singular or plural when they are not marked for case. In other syntactic positions singular NPs cannot be interpreted as referring to plural entities. However, in Mari the nominal plural marker is not obligatory in all syntactic positions. In DO position, singular DOs are ambiguous in interpretation both with and without the accusative marker:

(40) sol'o-m-ən **modəš** pog-əm-əž-əm už-ən-am.
 younger.brother-P.1SG-GEN toy gather-NZR-P.3SG-ACC see-PRT-1SG
 'I saw my younger brother picking up his toy / toys.'

(41) ava-ž-lan **uškal-əm** pukš-aš kül-eš.
 mother-P.3SG-DAT cow-ACC feed-INF need-PRS.3SG
 [We have three cows.] 'Mother has to feed the cows.'

In (41) the context (there are three cows) forces the plural reading. In another context the singular interpretation would be equally appropriate.

I assume that this property of Mari is due to the non-obligatoriness of the plural marker in general. It is noteworthy that Mari has four plural markers, all slightly different in their distribution. In reference grammars it is noted that number is not fully grammaticalized in this language (Galkin 1964, 78–82). The omission of plural markers in Mari is regulated by definiteness and animacy of the NPs: if the NP is definite and animate, the plural marker is most often present; else it may be omitted (Shmatova and Chernigovskaya 2012).

In Hindi and other languages where pseudoincorporation is limited to non-specific NPs, plural neutrality is (along with aspect etc.) based on scopal properties of the NPs, i.e. non-specificity (cf. the proposal of Farkas and de Swart (2003) to treat the plural marker as a presupposition trigger). Hence, it is entirely expected that in Mari, where pseudoincorporation of specifics and definites is possible, plural neutrality is not characteristic of pseudoincorporation.

It can be concluded that the referential properties of the DO are not the only factor determining the way in which it is encoded. More precisely, the type of modifiers found in the NP imposes strict limitations: DOs that host presupposition-inducing expressions cannot occur unmarked. However, bare nouns and nouns with adjectives/numerals/quantifiers can occur with or without the accusative, independent of their referential properties. As shown in Toldova and Serdobolskaya (2002), for these types of NPs it is the information structure of the sentence that is crucial.

2.3.2. Information structure and DO encoding

In Mari information structure is expressed by word order and prosody (and sometimes special particles). In (42a) and (42b) the sentence appears in such a context where the DO and the verb have different information structure functions. The DO is focused in (42a), while the verb constitutes the background part. By contrast, in (42b) the verb is focused, while the DO is the topic. In both examples the accusative marker on the DO is obligatory.

(42) a. – tide materjal gəč' mo-m urg-aš lij-eš?
 this cloth out.of what-ACC sew-INF be.possible-PRS.3SG
 – **tuvər-əm** / ***tuvər** urg-aš lij-eš.
 shirt-ACC shirt sew-INF be.possible-PRS.3SG

'What can I make of this cloth? – You can make a shirt.' (Toldova and Serdobolskaya 2002, 117)

b. {What are you going to do with the underclothes?}

– **vurgem-əm** / ***vurgem** šakal-aš kül-eš.
 underclothes-ACC underclothes hang-INF need-PRS.3SG
 'I have got to hang the underclothes.'

However, if the whole VP with the DO and the verb is focused, the accusative marker is omitted (43). The same holds for cases when the whole VP is topical (44).

(43) – pet'a-lan [mo-m əšt-aš] kü-l-eš?
Peter-DAT what-ACC do-INF need-PRS3SG

– pet'a-lan [pareŋge erəkt-aš] kü-l-eš.
Peter-DAT potato peel-INF need-PRS3SG

‘What should Peter do? Peter should peel the potatoes.’ (Toldova and Serdobolskaya 2002, 115)

(44) təj mo ergə-č'-ən araka jü-m-əž-lan kuan-et mo?
you Q son-P.2SG-GEN vodka drink-NZR-P.3SG-DAT rejoice-PRS.2SG Q

(The speaker and the hearer are discussing the fact that the hearer’s son drinks.) ‘Do you approve of your son’s drinking vodka?’

Thus, it is relevant whether or not the DO and the verb form a single information structure unit. This claim is supported by the interpretation of particles that mark information structure. Consider the following examples with the additive particle *-at*:

(45) məlam ni-mo-m č'ij-aš. jupk-əm-at/ *jupk-at, šovər-əm-at/
I-DAT NEG-what-ACC put.on-INF skirt-ACC-ADD skirt-ADD waistcoat-ACC-ADD /

*šovər-at urg-aš kü-l-eš.
waistcoat-ADD sew-INF need-PRS.3SG

‘I have got nothing to wear. I have got to make a skirt and a waistcoat.’

(46) šür-at šolt-aš kü-l-eš.
soup-ADD cook-INF need-PRS.3SG

‘I have got to make soup [and sweep the floor etc.]’

In (45) the particle semantically modifies the DO, and the omission of the accusative is unacceptable. In (46), though the particle occurs on the DO, it semantically modifies the whole VP “to make soup”. In this case the accusative is omitted.

The generalization can be made that the DO is unmarked if both the verb and the DO bear the same status in the information structure of the sentence, thus forming a single unit, a topical or wide-focused VP. The DO takes the accusative marker if the verb or the DO bears narrow focus, or if the DO / V is topicalized, or is in contrastive focus.

The evidence for this generalization is considered in detail in Toldova and Serdobolskaya (2002), where it is confirmed by tests with constituent questions, alternative questions, contrastive focus context and different discourse particles. It must be specified, however, that the omission of the DO marker is never obligatory. For example, it is omitted in (44), but it is present in a similar example produced by another consultant as her first reaction to the same Russian stimulus (47) (this was also judged as grammatical by the speaker that produced (44)):

(47) təj mo ergə-č'-ən araka-m jü-m-əž-lan kuan-et mo?
 you Q son-P.2SG-GEN vodka-ACC drink-NZR-P.3SG-DAT rejoice-PRS.2SG Q

(Both the speaker and the hearer know that the hearer's son drinks.) 'Do you approve of your son's drinking vodka?'

The restrictions imposed by information structure on the absence of the accusative marker are necessary, but not sufficient: it can be omitted only if the whole VP belongs to one and the same information component, but omission is not obligatory. The unmarked DO is preferred if its referent has low discourse relevance in terms of Givón's notions of referential distance, topic persistence (cf. Givón 1983), number of mentions in the discourse (i.e. the number of times the referent is mentioned) etc. It has been shown in Toldova and Serdobolskaya (2012b) that it is the discourse rank of the DO that is crucial for omission of the accusative: DOs with low topic continuity and number of mentions are likely to appear unmarked, while DOs with high topic continuity and number of mentions tend to bear the accusative even when the information structure restrictions are respected.

Let us consider the following example taken from a narrative (a tale):

(48) ohotnik pünč'ö vuj-əšte šinč'-əše or'ol-əm už-eš da
 hunter pine branch-INESS sit-PTCP.ACT eagle-ACC see-PRS.3SG and
 pəč'al-ž-əm tud-ən ümbak vikt-a, lüj-ne-že. a or'ol-et
 gun-P.3SG-ACC he-ACC on aim-PRS.3SG shoot-DESID-3SG and eagle-EMPH
 ohotnik-ən pəč'al vikt-əm-əm už-eš-at,
 hunter-GEN gun aim-NZR-ACC see-PRS.3SG-ADD
 ajdeme jük de-ne tud-lan ojl-a:
 human voice near-INESS he-DAT say-PRS.3SG

'A hunter goes by, spots the eagle sitting on a pine branch, and points the gun at it, wants to kill it. The eagle sees the hunter pointing his gun and says with a human voice...'

In this text the main characters are a mouse and a sparrow, and in the course of the narrative they are changed into a hunter and an eagle. The noun "gun" in this episode is mentioned only twice, in adjacent sentences. It occurs firstly in the VP "points the gun at it" and secondly in the backgrounded non-finite clause, lit. "the hunter's pointing the gun". In this second case "gun" occurs within the topical domain together with the verb. Hence, the accusative marker can be omitted. Let us compare this to another example with the same type of information structure, where the NP is highly discourse relevant:

(49) pervəj paša-m mu-aš kül-eš, jol ümbak šogal-aš kül-eš.
 first job-ACC search-INF need-PRS.3SG foot on raise-INF need-PRS.3SG

'{I don't want to marry now.} First one needs to find a job, set oneself up.' (A dialogue.)

This example is from a discussion of the speaker's life plans: studying, finding a job and marrying. Hence, the generic referent "job" is mentioned several times in the VP "to find a job", and is highly discourse relevant. Thus, it occurs with the accusative marker, though the whole VP is not partitioned from the point of view of information structure.

This semantic type is somewhat close to Mithun's (1984) "manipulation of discourse structure" variety of incorporation. However, Mithun outlines cases where the noun and the verb denote

information already activated in the discourse, both of them therefore belonging to the topical domain. In Mari, DOs can also be pseudoincorporated if they belong to wide focus.

Very often, unmarked DOs are observed if the VP describes an “institutionalized activity” (wash the hands, sweep the floor, collect berries etc.), cf. Mithun (1984). In those cases, even nouns marked with possessive suffixes can appear without the accusative. Hence, both from the point of view of information structure and discourse structure, the verb and the unmarked DO form one and the same unit.

The importance of discourse structure can be supported by the following observation. The unmarked DO is discourse opaque (as in Hindi, cf. Dayal 2011). In spontaneous texts no examples have been found of an unmarked DO as an antecedent of an anaphoric pronoun. In elicited examples, the speakers prefer the accusative when the DO is referred to:

- (50) pet’a-n ala-mogaj **poŋg-əm** pog-əm-əž-əm da vara
 Peter-GEN INDEF-what.kind.of mushroom-ACC collect-NZR-P.3SG-ACC and then
tud-əm lukt-ən kudalt-əm-əž-əm už-əm.
 he-ACC take.out-CONV throw.away-NZR-P.3SG-ACC see-NARR.1SG
 ‘I saw Peter collect a lot of mushrooms and then throw them away.’

When asked about the possibility of omitting the accusative, they offer paraphrases like the following:

- (51) pet’a-n **poŋgo** pog-əm-əž-əm už-əm no vara č’əla
 Peter-GEN mushroom collect-NZR-P.3SG-ACC see-NARR.1SG but then everything
 lukt-ən kudalt-əž.
 take.out-CONV throw.away-NARR(3SG)
 ‘I saw Peter collect a lot of mushrooms and then throw everything away.’

Notice that in (50) an ordinary anaphoric pronoun is used, while in (51) the unmarked DO is referred to by the pronoun ‘everything’ with vague reference to the entities described before. Similar examples are given for lexical compounding in English, cf. *Mary went apple-picking. #They/The apples were delicious* from (Dayal 2011, ex. 5). In English examples, a usual anaphoric device is banned, while another referring expression (like the DP *the apples*) is acceptable.

It can be concluded that the unmarked DO in Mari most often has low discourse relevance. This interacts with information structure in the following way: if the DO has low discourse relevance, the whole VP is focused or topical (and the accusative is omitted). If it has high discourse relevance, then it can either remain in the same information component as the verb (taking the accusative), or constitute the topic/focus of the sentence (also taking the accusative).

Thus, the main factors that determine the DO marking in Mari are information flow and discourse structure, while the referential properties of the DO are less relevant. This contrasts with other cases of pseudoincorporation and many cases of incorporation, which require non-specific semantics of the DO.

3. Discussion and conclusions

The investigation of some cases of asymmetric DOM leads to their analysis in terms of incorporation or pseudoincorporation, as in Turkish (Muravyova 2008), Niuean (Massam 2001) Hindi (Dayal 2011), and others. I provide such an analysis for Eastern Mari. Constructions with unmarked DOs in Mari have the following features characteristic of incorporation: no accusative case marking, strict verb

adjacency and no prosodic pause between the DO and the verb, a ban on demonstrative, indefinite and interrogative pronouns as modifiers, a ban on restrictive relative clauses and a ban on pronominal unmarked DOs and unmarked proper names in DO position. However, an analysis of this construction in terms of incorporation (morphological or lexical compounding) is ruled out by the availability of coordinated DOs, plural marking, possessive markers, numeral and numerative expressions (e.g. *two bottles of milk*), quantifiers, and adjectival and nominal modifiers with unmarked DOs.

On the basis of these properties, I propose to analyse constructions with unmarked DOs in Mari in terms of the pseudoincorporation of a NumP. I define pseudoincorporation as a syntactic construction where a verbal argument is expressed via a nominal constituent smaller than DP (an NP, a NumP etc.).

I assume the DP structure proposed in (Longobardi 2001, 597). Here I give a fragment of his schema:

(52) [Q [D [GenS [Num [S-adj [M-adj [Arg-adj [NN-dep N]]]]]]]]]

where Q is a position for quantifiers, D is a determiner position, GenS is a position for prenominal genitive, Num is a base position for numerals and numerative groups (as in *two bottles of milk*), S-adj is a subject- or speaker-oriented adjective, M-adj is a manner adjective, Arg is an argument adjective, NN-dep is a position for unmarked dependent nouns, and N is the head noun.

Considering the position of Q, D, adjectives and dependent nominals, I follow the analysis of the Mari nominal complex proposed in Simonenko and Leontjev (2012). I use the label NN-dep for unmarked dependent nouns. The unmarked dependent nouns in the NP form a construction known as juxtaposition in the Finno-Ugric literature, e.g. Mari *kü pərdəž* ‘‘a stone wall’’. This construction is similar to English compounds like *stone wall*, *fish tail* etc. In Mari, the dependent noun in such constructions cannot take any nominal grammatical markers or modifiers. Thus, I follow Simonenko and Leontjev (2012, 325) in analysing such constructions as head adjunction under External Merge.

In terms of morphosyntax, my proposal is similar to the analysis developed by Massam for Niuean, Dayal for Hindi and Farkas and de Swart for Hungarian. They assume that the unmarked DO is incorporated on the NP level and, hence, does not have the DP projection. Lyutikova and Pereltsvaig, treating asymmetric DOM in Mishar Tatar, argue for the analysis of unmarked DOs as Small Nominals (cf. Lyutikova and Pereltsvaig *forthc.*). They claim that Small Nominals cannot be assigned structural case and do not have the D level, whence the restriction on various determiners and modifiers with unmarked DOs.

All these authors partly base their argumentation on the semantics of the pseudoincorporated DO. In these languages unmarked DOs must be non-specific, and this is easily explained by its lack of the D level, if we make the assumption that the referential potential of the NP is established at this level (following Stowell 1989, Longobardi 2001 among others).

However, the semantics of unmarked DOs in Mari differs from these cases. Pseudoincorporated DOs in Mari can be specific and even definite. In Mari omission of the accusative is determined by information flow and discourse structure: the DO and the verb must form a single unit. They must belong to one and the same information structure component, and the DO must have low discourse relevance, based on the parameters suggested in (Givón 1983).

Thus, Mari syntactic constructions with unmarked DOs do lack the D level; however, this does not affect their semantic structure⁷. I consider this to be an argument for differentiating between the

⁷ These facts are not easy to analyse if we take the usual assumptions regarding the semantic contribution of the D level. They are, however, captured by the approaches where the semantic and syntactic structure are independent, e.g. Culicover and Jackendoff (2005).

syntactic structure of DPs and their semantics. In a way this idea follows the approach taken by Farkas and de Swart (2003), who propose to differentiate between thematic arguments at the syntactic level and discourse referents at the level of discourse. Thus, a full-fledged DP does not necessarily introduce a discourse referent. I suggest that, vice versa, an argument can be a discourse referent, even if it is not a DP syntactically.

This solution does not seem surprising, since there are languages that allow specific and definite nouns to be incorporated even in lexical compounding constructions. Such examples are considered in Mithun (1984; 2000):

Huahtla Nahuatl

(53) A: kanke eltok kočillo? Na' ni-'neki amanci.

where is knife I I-it-want now

'Where is the knife? I want it now.'

B: ya' ki-kočillo-tete'ki panci.

he (he)it-knife-cut bread

'He cut the bread with it (the knife).' (Mithun 1984, 861)

Even if the DO is definite (as in (36)–(37), (48)), it can have low discourse relevance; as Mithun suggests for Mohawk and similar cases of incorporation of definites, their “identification is deduced rather than specified” (2000, 925). That is, the speaker presumes that the hearer can identify their referents; however, s/he does not require them to be identified, because these participants are not relevant to the discourse.

A similar explanation is offered for weak definites like *the hospital* in *Galya is in the hospital* (Klein et al. 2009). The authors assume that “weak definites seem to evoke the event or activity being described, rather than an individual discourse referent corresponding to the noun”. In Mari, as in English, weak definites allow distributive reading by ellipsis (*Jacob went to **the bank** today, and Carmen did too*), where the referent of *the bank* in the first clause does not necessarily coincide with the ellipsed NP in the second clause.

Therefore, I propose to analyse the unmarked DO in Mari in terms of pseudoincorporation at the NumP level. This solution is possible if the theoretical decision is taken to accept, semantically, the pseudoincorporation of specific and even definite NPs. I assume that Mari DOs are syntactically NPs and do not have the D level; however, the lack of D does not affect their semantics.

References

- Aissen, Judith, “Differential object marking: Iconicity vs. economy,” *Natural Language and Linguistic Theory* 21, № 3 (1998): 435–483.
- Alhoniemi, Alho, *Grammatik des Tscheremissischen (Mari): mit Texten und Glossar*. Hamburg, Buske, 1993.
- Anduganov, Jurij D., *Istoričeskaya grammatika marijskogo jazyka. Sintaksis*. Vol. 1: *Vvedenie. Substantivnye slovosočetačanija*. Joškar-Ola, Marijskoe knižnoe izdatel'stvo, 1991.
- Baker, Mark C., *Incorporation: A theory of grammatical function changing*. Chicago, University of Chicago Press, 1988.

- , “Pseudo-noun incorporation as covert noun incorporation: linearization and crosslinguistic variation.” Unpublished Rutgers University ms., 2012. (Presented at the Academia Sinica in Taiwan, 2011) <http://www.rci.rutgers.edu/~mabaker/PNI-adjacency-Taiwan.pdf>
- Bossong, Georg, *Differentielle Objektmarkierung in den neuiranischen Sprachen*. Tübingen, Gunter Narr, 1985.
- Comrie, Bernard, “General features of the Uralic languages,” *The Uralic languages. Description, history, and foreign influences*, ed. Denis Sinor, 451–477. Leiden – New York – København – Köln, E.J. Brill, 1988.
- Culicover, Peter W. and Ray Jackendoff, *Simpler Syntax*. Oxford, Oxford University Press, 2005.
- Dalrymple, Mary, Stuart M. Shieber and Fernando C. N. Pereira, “Ellipsis and higher-order unification,” *Linguistics and Philosophy* 14 (1991): 399–452.
- Dalrymple, Mary and Irina Nikolaeva, *Objects and information Structure*. Cambridge, Cambridge University Press, 2011.
- Dayal, Veneeta, “Hindi Pseudo-Incorporation,” *Natural Language and Linguistic Theory* 29, № 1 (2011): 123–167.
- , “A semantics for pseudoincorporation,” Unpublished Rutgers University ms., 2003. <http://semanticsarchive.net/Archive/WI2ZjZkM/pseudo-incorporation.pdf>
- de Swart, Peter, “Cross-linguistic variation in object marking.” Ph.D. diss., Radboud Universiteit Nijmegen, 2007.
- Enç, Mürvet, “The semantics of specificity,” *Linguistic Inquiry* 22 (1991): 1–25.
- Espinal, M. Teresa and Louise McNally, “Bare nominals and incorporating verbs in Spanish and Catalan,” *Journal of Linguistics* 47 (2011): 87–128.
- Farkas, Donka and Henriëtte de Swart, *The semantics of incorporation: from argument structure to discourse transparency*. Stanford, CA, CSLI Publications, 2003.
- Galkin, Ivan S., *Istoričeskaja grammatika marijskogo jazyka. Morfologija*. Vol. 1. Joškar-Ola, Marijskoe knižnoe izdatel'stvo, 1964.
- van Geenhoven, Veerle, *Semantic incorporation and indefinite descriptions: semantic and syntactic aspects of noun incorporation in West Greenlandic*. CSLI Publications, 1998.
- Geurts, Bart, “Presupposition and anaphors,” *Linguistics and Philosophy* 21 (1998): 545–601.
- Givón, Talmy, ed., *Topic continuity in discourse: A quantitative cross-language study*. Amsterdam, John Benjamins, 1983.
- Havas, Ferenc, “Unmarked object in the Uralic languages. A diachronic typological approach,” *Linguistica Uralica* XLIV, №1 (2008): 1–33.
- von Heusinger, Klaus and Jaklin Kornfilt, “The case of the direct object in Turkish: semantics, syntax and morphology,” *Turkic Languages* 9 (2005): 3–44.
- Klein, Natalie M., Whitney M. Gegg-Harrison, Rachel S. Sussman, Greg N. Carlson and Michael K. Tanenhaus, “Weak definite noun phrases: rich, but not strong, special, but not unique,” *Semantics and pragmatics: from experiment to theory*, ed. Ulrich Sauerland and Kazuko Yatsushiro (Palgrave studies in pragmatics, language, and cognition), 264–276. Houndmills, Basingstoke, Hampshire; New York, Palgrave Macmillan, 2009.
- Koptjevskaya-Tamm, Maria. *Nominalizations*. Routledge, London and New York, 1993.
- Krifka, Manfred, Francis J. Pelletier, Gregory N. Carlson, Alice ter Meulen, Gennaro Chierchia and Godehard Link, “Genericity: An introduction,” *The generic book*, ed. Gregory Carlson, Francis J. Pelletier, 1–125. Chicago – London, The University of Chicago Press, 1995.
- Kuznetsova, Ariadna I., Natalia V. Serdobolskaya, Svetlana Ju. Toldova, Sergey S. Saj and Elena Ju. Kalinina, eds., *Finno-Ugorskie jazyki: fragmenty grammatičeskogo opisanija. Formal'nyj i funkcional'nyj podxody*. Moscow, Jazyki slavjanskix kul'tur, 2012.

- Lavrentev, Guriĭ I., “Nekotorye voprosy obrazovania i oproščenia kompozit v marijskom jazyke,” *Sovetskoe finno-ugrovedenie* 11, № 4 (1975): 244–252.
- Longobardi, Giuseppe, “Reference and proper names: a theory of N-movement in syntax and logical form,” *Linguistic Inquiry* 25 (1994): 609–65.
- , “The structure of DPs: some principles, parameters and problems,” *Handbook of contemporary syntactic theory*, ed. Chris Collins and Mark Baltin, 562–603. Oxford, Blackwell, 2001.
- Ljutikova, Ekaterina A. and Asya Pereltsvaig, “Elucidating nominal structure in articleless languages: a case study of Tatar,” to appear in: *Berkeley Linguistic Society proceedings*. Forthcoming.
- Massam, Diane, “Pseudo noun incorporation in Niuean,” *Natural Language and Linguistic Theory* 19 (2001): 153–197.
- Mithun, Marianne, “The evolution of noun incorporation,” *Language* 60 (1984): 847–894.
- , “Incorporation,” *Morphologie/Morphology: a handbook on inflection and word formation/ein internationales Handbuch zur Flexion und Wortbildung*, ed. Geert Booij, Christian Lehmann, Joachim Mudgan, 916–928. Vol. 1. Berlin, Walter de Gruyter, 2000.
- , “Constraints on compounding and incorporation,” *Compounding*, ed. Irene Vogel and Sergio Scalise, 37–56. Amsterdam, John Benjamins, 2010.
- Moravcsik, Edith, “On the case marking of objects,” *Universals of human language*, ed. Joseph H. Greenberg, 249–289. Vol. 4: *Syntax*. Stanford, Stanford University Press, 1978.
- Muravyova, Irina A., “O traktovke neofornlenogo imeni v t’urkskix jazykax,” *Issledovanija po teorii grammatiki*. Vol. 4: *Diskursivnye aspekty grammatiki*, ed. Yuri Lander, Vladimir Plungian and Anna Urmanchieva, 321–421. Moscow, Gnozis, 2008.
- Pearce, Elizabeth, “Lexical feature-driven incorporation in two Oceanic languages: Iaaï and Fijian,” *The Linguistic Review* 18 (2001): 265–288.
- Pengitov, Nikolaj T., ed., *Sovremennyj marijskij jazyk*. Vol. 2: *Morfologija*. Joškar-Ola, Marijskoe knižnoe izdatel’stvo, 1961.
- Poesio, Massimo, “Weak Definites,” *Proceedings of the fourth conference on Semantics and Linguistics Theory, SALT-4*, 282–299. Ithaca, Cornell University Press, 1994.
- Sadock, Jerrold M., “Some notes on noun incorporation,” *Language* 62 (1986): 19–31.
- Shibasova, Natalia L., “Tipologija parnyx slov: na materiale nekotoryx finno-ugorskix jazykov.” Ph.D. diss., Moscow Lomonosov University, 2006.
- Shmatova, Maria S. and Evgenija A. Chernigovskaya, “Kategorija čisla suščestvitel’nogo v marijskom i permskix jazykax,” Kuznetsova et al. *Finno-Ugorskie jazyki: fragmenty grammatičeskogo opisanija. Formal’nyj i funkcional’nyj podxody*, 221–250.
- Simonenko, Alexandra P. and Alexey P. Leontjev, “Morfosintaksis imennogo kompleksa v finno-permskix jazykax: analiz v ramkax gipotezy minimalizma,” Kuznetsova et al. *Finno-Ugorskie jazyki: fragmenty grammatičeskogo opisanija. Formal’nyj i funkcional’nyj podxody*, 262–340.
- Stowell, Tim, “Objects, specifiers and X-bar theory,” *Alternative conceptions of phrase structure*, ed. Mark R. Baltin and Anthony S. Kroch, 232–262. Chicago, University of Chicago Press, 1989.
- Thompson, Sandra A., “Transitivity and some problems with the *ba* construction in Mandarin Chinese,” *Studies in East Asian syntax*, ed. George Bedell, 70–86. UCLA papers in syntax, 1972.
- Toldova, Svetlana and Natalia Serdobolskaya, “Nekotorye osobennosti ofornlenija prjamogo dopolnenija v marijskom jazyke,” *Lingvističeskij bespredel*, ed. Olga A. Kazakevich, 106–125. Moscow, Nauka, 2002.
- , 2012a. “Differencirovannoe markirovanie pramogo dopolneniya v finno-ugorskix jazykax,” Kuznetsova et al. *Finno-Ugorskie jazyki: fragmenty grammatičeskogo opisanija. Formal’nyj i funkcional’nyj podxody*, 59–142.

- , 2012b. “Information structure at odds with discourse factors: evidence from Finno-Ugric differential object marking,” Paper presented at the workshop “Categories of Information Structure across Languages”, Max Planck Institute for Psycholinguistics, Nijmegen, November 9-10, 2012.
- Tužarov, Gennadij M., “Problema nemarkirovannogo imeni v marijskom jazyke,” *Sovetskoje finno-ugrovedenie* 20, № 4 (1984): 282–289.
- , “Problema nemarkirovannogo akkuzativa v marijskom jazyke,” *Sovetskoe finno-ugrovedenie*, vol. 22, № 2 (1986): 99–107.
- , “Grammatičeskie kategorii imeni suščestvitel’nogo v marijskom jazyke.” Dokt. diss., Joškar-Ola, 1998.
- Vergnaud, Jean-Roger, “French relative clauses.” Ph. D. diss., M.I.T., Cambridge, Mass, 1974.
- Wickman, Bo, *The form of the object in the Uralic languages*. Uppsala: Lundequistska bokhandeln, 1955.
- Zwart, Jan-Wouter, “A head raising analysis of relative clauses,” *The syntax of relative clauses*, ed. Artemis Alexiadou, Paul Law, André Meinunger, and Chris Wilder (Linguistik Aktuell / Linguistics Today, Bd. 32), 349–387. Amsterdam, John Benjamins, 2000.