

FERENC HAVAS (Budapest)

## UNMARKED OBJECT IN THE URALIC LANGUAGES. A DIACHRONIC TYPOLOGICAL APPROACH

**Abstract.** The conventional wisdom is that unmarked object existed in Proto-Uralic. Of the syntactic environments I would only exclude the occurrence of unmarked object next to "passive" verbs from those that may have been inherited from the proto-language. For the remaining six environments inheritance from the proto-language is at least a possible alternative, in some cases the only one. Thus the efforts to derive unmarked object from the reanalysis of some other syntactic function of the nominative form are unjustified. The synchronic coexistence of the syntactic structures found in the world's languages is the depository of a historical relation in a diachronic-typological sense, which can be interpreted in terms of a schematogony. In these terms, the most recent syntactic structure is the nominative pattern with subject and object as the typical categories. Subject typifies the nominative pattern, but object is found already in triadic languages and is thus a pre-nominative category in itself. More clearly pre-nominative is its unmarked variant, whose origins go back to the proto-transitive and proto-active stages, where the category of object had not yet emerged. Consequently, unmarked object must be a pre-nominative inheritance in Uralic, one of those relics that make it possible to trace the prehistory of our languages in a diachronic-typological sense back into periods that are — justifiably — excluded from the scope of traditional comparative and historical research.

**Keywords:** Uralic, unmarked object, syntactic structures, diachrony, typology.

### 1. The notion of unmarked object

Problems of the unmarked object in Uralic are among the frequently revisited ones, discussed usually in the context of single languages or subgroups within the family, and usually from a descriptive point of view. Within Uralic linguistics the prime concern was always to prove that the category of marked object was present in the proto-language. These efforts were successful: as is known, an accusative ending *\*-m* can be reconstructed for the proto-language on the basis of the modern Uralic languages.<sup>1</sup> At the same time all those who

<sup>1</sup> The views that challenge the reconstruction of the Proto-Uralic *\*-m* suffix have remained a minority (see Майтинская 1974 : 244—245). Some suggest a second accusative ending *\*-t* besides *\*-m* for Proto-Uralic (see Bereczki 2003 : 51). For the purposes of the present paper this is immaterial.

have studied this question agree that the object was not always marked, hence Proto-Uralic must have possessed the category of unmarked object. This conclusion is based on the simple observation that, in some form, unmarked objects are present in all modern Uralic languages, and the current form of these objects, which thus appear in the nominative, cannot in most cases be thought to result from the apocope of the original *\*-m*.

In the present paper I shall attempt first to survey unmarked objects in the Uralic languages in terms of structural categories, then to analyse the parallel phenomena thus uncovered along the lines of diachronic typology. But before I embark upon this, the clarification of some important terms and the delimitation of the questions to be posed is in order.

I will make no attempt here at precisely defining the notion of object and transitive verb. In everyday descriptive practice these terms tend to be used on the analogy of their prototypical occurrences and it is left to intuition to decide to what extent a particular linguistic construction may deviate from the benchmark and still count as an object or a transitive verb.<sup>2</sup> It is nevertheless clear that the apparently self-explanatory definition "an object is the patient of a transitive verb (or at least a transitive verb in the active voice)" needs to be qualified from a typological perspective. I refer to something as an object if in the given language no other type of patient

---

<sup>2</sup> On a first approach object is nothing else but a kind of patient. The notion of patient contrasts primarily with that of the agent. The term agent refers to the intentionally acting instigator of the action, process or state expressed in the verbal predicate, whereas the term patient refers to the animate or inanimate entity that is in some way affected by the process, on which it unfolds, to which it is directed, which comes into being due to the event or which – without any active participation – is in or enters the state referred to by the verbal predicate. This definition of patient is, however, much broader than that of the object: the subject of a medial transitive verb – such as *die* – is also a patient. In a more restricted sense, an object is the patient of a transitive verb. At the same time, a transitive verb can hardly be defined without reference to the patient, which then renders our definition inherently circular. A verb is usually called transitive if it has at least two obligatory complements. In the default case one of these is a patient, the other normally an agent. (This is, however, not always true. In a sentence like *the boy hates spinach* neither noun can be categorised as an agent and, if we take the above definition of patient seriously, the actual patient in this sentence would be *the boy*, whereas the verb *hate* is categorised as transitive because of the complement *spinach*.) Furthermore, the definition "an object is the patient of a transitive verb" is not reversible. In classical passive structures we are faced with a dilemma: either we claim that the verbs found in them are not transitive, or we have to admit that the patient of the transitive verb in them is a subject rather than an object. What we end up with, then, is the formula "the patient of a transitive verb in the active voice", and we can hope that we will be able to define transitivity and active voice without reference either to each other or to the object and thus avoid the vicious circle. If we go further we note that the criterion of two obligatory complements can also get us into trouble. Several languages have impersonal transitive verbs that have only one obligatory complement which, however, is usually taken to be its object (e.g. Latin *taedet me* 'I detest'). To circumvent this problem one is tempted to abandon the semantic approach in favour of a morphosyntactic one and say that the object is what assumes the accusative case in the vicinity of the verb (or displays other formal features reserved for this function in any particular language). But even this will not get us far. For one thing, not everything that is in the accusative in the vicinity of a verb is an object (in several languages it can be adverbs of goal or time); for another, the object of a verb is not necessarily something that is in the accusative. At any rate, the case form called accusative could only be defined again with reference to transitivity, thus it may be reasonable to give up the idea of an axiomatic–deductive definition of object.

is characterised by the same form of expression. If, for instance, a language makes no formal distinction between the patients of intransitive vs. transitive verbs, the latter are not a category on their own, i.e. in such languages the patients of transitive verbs are grammatically not objects. The category of object in that sense is found only in nominative and triadic languages, as will be seen later. Since Uralic languages are all nominative languages, the use of this category is justified in their analysis.

In what follows the focus will be on objects expressed by nouns in the Uralic languages. This is because in several languages pronouns, especially personal pronouns display a behaviour different from that of nouns in object function. In Ostyak, for instance, only the former have marked accusative case, in Finnish they assume a marked accusative where nouns remain unmarked, and so on. Since I am primarily interested in primordial forms of object marking, I further restrict myself to the study of singular forms, which may have been either the only forms in the proto-language, or were at least primary and provided a model for the formation of non-singular forms.<sup>3</sup>

The term markedness has a variety of meanings. Something may be marked as object by case morphemes, clitics, postpositions, word order or some formative of topicalisation. In the present paper a marked object will always refer to an object marked by a case morpheme, and this morpheme will be referred to as accusative.<sup>4</sup>

A further qualification must be made regarding the markedness, or more precisely the unmarkedness, of objects. By unmarked object in this paper I only mean nominal forms that contrast in the same language with marked accusative forms for nouns. It is only in such a case that the unmarkedness of the object has grammatical value. In Ostyak, for instance, where, as said above, the accusative of nouns always coincides with the nominative, a noun, which is in base form even in object function, will not be referred to as an unmarked object, even though that is what it formally is.

One last point: in order for unmarkedness to be typologically interpretable, it has to be specific with respect to a particular structural category. If, for instance, the indefinite object is left unmarked next to a finite indicative verb in a particular language, the same unmarkedness in the context of participles, imperatives etc. is not specific (i.e. does not arise due to that particular grammatical environment).

## **2. Structures involving unmarked object in the Uralic languages**

### **2.1 With finite indicative forms of transitive verbs**

In at least half of the Uralic languages finite indicative forms of transitive verbs may be accompanied by an unmarked object. Interestingly, this unmarkedness indicates the indefiniteness of the object in the Finno-Ugric languages, whereas in the larger part of Samoyedic, its definiteness. I will

<sup>3</sup> L. Honti (2003) presents a different view; see also the references there.

<sup>4</sup> Accusative will thus not refer to the function of the objects, as in most Finnish grammars, where this term is often used in two senses, to denote (i) an abstract function whose morphological exponent is (barring personal pronouns) the nominative and the genitive, (ii) the accusative case of personal pronouns. To avoid equivocation I will henceforth use *unmarked object*, which of course coincides with the nominative, and I will not say *unmarked accusative*.

now illustrate this phenomenon through a handful of examples from each of these languages.

In Vogul the ending of the definite object is *-m* or *-mV* in all dialects. By contrast, the indefinite object that accompanies a verb in indefinite form is unmarked:

- (1) *l ə g n ə n ɔ h s ə g äli* 'he hunts squirrels and sables',
- (2) *t ē n e k a r tēs* 'he ate food' (Wickman 1955 : 68).

However, the differentiation between definite and indefinite object in Vogul is not entirely consistent. In principle a verb in definite form governs a definite object and a verb in indefinite form governs an indefinite object. In reality, counterexamples occur, especially with nouns bearing possessive suffixation (Px), where the general rule is valid in the first and second persons, but in the third person the accusative suffix is sometimes found next to Px or, contrary to what the rule would predict, is not found even on definite objects, as in 3:

- (3) *ń ā l ə m kwänlēwəliləm* 'I pull out my arrow' (Wickman 1955 : 69–71).

The situation in Permic is more straightforward. In Votyak, a definite object is always suffixed, an indefinite object is always suffixless, except when it bears the possessive suffix, since nouns with Px are always definite. 4 and 5 are examples of unmarked object:

- (4) *g o ž t e t gožto* 'I will write a letter',
- (5) *d a s k i l o m e t r pjdin minono* 'one must go ten kilometers on foot' (Csúcs 1990 : 63).

Zyryen shows essentially the same with the exception that unmarkedness is only found with nouns not denoting human beings. An unmarked object is indefinite, a marked object is definite (although in post-verbal position an unmarked object may have an accusative ending) (Лыткин 1955 : 142). Words denoting human beings and nouns with Px can only be definite objects. Examples for unmarked object are given in 6 and 7:

- (6) *Peder pjr k e ć kije* 'P. keeps hunting hare' (Rédei 1978 : 93),
- (7) *Me šeta tenjd s u m k a* 'I give you a knapsack' (Wickman 1955 : 61).

In Mordvin, the general principle is the same, so an indefinite object tends to be unmarked (and thus coincides with the nominative), as in the following sentences:

- (8) *свѧк семдонь, пади, кодама-кодама с т и р н я каванят* 'take all of it [scil. of the money], in case you host some girl' (Грамматика 107),
- (9) *tjndenĳ joutan tējk b o b a s k a* 'I will tell you a tale' (Wickman 1955 : 47).

Mordvin also has an indefinite and a definite verbal conjugation, and this is related to the choice between the two kinds of objects: indefinite objects accompany indefinite verbs, definite objects accompany definite verbs. Objects with Px are always accompanied by a definite verb, but where the nominative and the accusative of Px forms differ from each other, the case marking of nouns vacillates (as in Erza), thus unmarkedness occurs, like in 10:

- (10) *istuvtik ton nej a v a t?* 'have you forgotten your mother?' (Wickman 1955 : 53).

In Cheremis and in the Finnic languages unmarked objects next to indicative finite verbs are unattested, but in Lapp, the language most closely related to Finnic, they are found in the plural, as 11 shows:

(11) *vuostəgh d u ø l j ə j d' ə riddšiepəawh* 'first we tan the hides' (Wickman 1955 : 32).

The same is usually the case in possessive structures:

(12) *ō, mū ahčēn, etnen č ä l m e h wäinam* 'oh, I see my father's and mother's eyes' (Wickman 1955 : 35).

An indefinite object is always unmarked in the plural, but a definite object is either unmarked or marked for the accusative, or, in Southern Lapp — as opposed to the Finnic languages — for the partitive case (cf. Wickmann 1955 : 35 ff.). This does not appear to be an archaic feature; it is more likely that in Lapp the same form was used in the plural too to express the subject and the object next to an indicative finite verb, and thus the category of unmarked object (in the contrastive sense in which this term is used here) was unknown.

The behaviour of nouns with Px in Lapp is different. It is generally true that for such nouns the nominative and the accusative coincide in the first person in all three numbers (Singular, Dual, Plural), and this rule has been generalised in certain dialects to all persons in the plural. Since possessive suffixes have a determinative effect in all Uralic languages that actually have such suffixes, it is up to the grammarian to decide whether to regard such objects as unmarked or not.<sup>5</sup>

By contrast, all Samoyedic languages clearly have a category of unmarked object next to an indicative finite verb, and in three of these: (Tundra) Nenets,<sup>6</sup> Enets and Selkup this form expresses the definiteness or the focus function of the object. In other words, the marked object is used for topicalisation. This is what the two Nenets sentences in 13 and 14 show (the former with accusative, the latter with unmarked object).

(13) *тюку тым' намзодайна"* т э м д а в а 'we bought this deer [for food<sub>FOCUS</sub>]',

(14) *тюку т ы намзодайна"* т э м д а в а 'we bought [this deer<sub>FOCUS</sub>] for food' (Терещенко 1973 : 178).

Similar phenomena are seen in Enets, where the unmarked object is definite, while that in the genitive case is indefinite:

(15) *д я х а"* мотэза 'he crossed the rivers',

(16) *дяхү"* мотуна 'he crossed rivers' (Терещенко 1973 : 179).

With certain stem types, unmarked object is found in Nganasan as well, but here it is not related to the definite–indefinite distinction.

(17) *н̄ягə т ā хутруи"*эм 'I have caught a/the good deer',

(18) *хиримə к о л ы н̄əмса"ту"* 'they are eating cooked fish' (Терещенко 1973 : 75).

<sup>5</sup> From a strictly structural perspective it must be unmarked. One reason is that in these languages Px-es make a noun definite even in the nominative (often they have no other function, since they do not express possession), thus we cannot say that the Px used after transitive verbs is simultaneously the exponent of objecthood along with definiteness. The other reason is that in some languages, as we have seen (or indeed in different persons within the same language), a separate element appears beside the Px to indicate the definiteness of the object. Both facts point to a zero object morpheme on nouns with Px and thus to an unmarked object in Lapp.

<sup>6</sup> In Forest Nenets the *-m* accusative is unattested, object and subject without Px coincide, while Px sometimes assumes an object-marking function.

Kamass object marking used to be similar to that in the Finno-Ugric languages just seen: definite object was suffixed with *-m*, indefinite object was unmarked. The latter is shown in 19 and 20:

(19) *k ā m ž u kubiaŋ* 'I found a whip',

(20) *u j a' pa'bi* 'she cooked meat' (Wickman 1955 : 139).

Selkup, by contrast, follows the pattern of Enets and Nenets: an unmarked object is definite, a marked object is indefinite,<sup>7</sup> as shown in 21 and 22:

(21) *п и ч и т ы мицалтыты* 'he picked up his ax' (Терещенко 1973 : 179),

(22) *nagurmželi čemnjād a n d mannembad* 'the third brother sees the boat' (Wickman 1955 : 128).

Note that in Vogul, Lapp and Kamass the unmarkedness of the object is generalised also to personal pronouns, whereas in the rest of the languages surveyed up to this point personal pronouns appear obligatorily in the accusative case as objects.

## 2.2. With imperative forms

In some Uralic languages the form of nominal objects accompanying the affirmative imperative of a transitive verb coincides with the nominative (in some of these languages further conditions are necessary). As I have said earlier, the unmarked object in this context can only be regarded as specific where it differs from the form that appears next to indicative finite verb forms. Accordingly, not only those languages will be left out of consideration here which generally do not distinguish between marked vs. unmarked object, such as Ostyak, Forest Nenets or Livonian (though in the last so called passive structures are an exception),<sup>8</sup> but also those that show the opposite, i.e. which distinguish between the two object types in the indicative too, since in these languages this is a particular case of a general principle, the unmarked object not being governed by the imperative *qua* imperative.

Thus here I only list those languages in which the imperative is accompanied by an unmarked object — subject to the conditions that hold for the indicative —: Vogul, both Permic languages and — with the exception of Forest Nenets — all Samoyedic languages. Nganasan is, in a sense, a subset in itself because alternating stems there occur in the so called second stem when unmarked (also, incidentally, in prohibitions), so it is the analyst's choice whether to regard this as unmarked object. In Nenets, Enets and Selkup, unmarkedness is again the form of the definite object, and the accusative is that of the indefinite object. In second person imper-

<sup>7</sup> This is what N. Tereščenko consistently claims (e.g. Терещенко 1973 : 181); A. Kuznecova, E. Helimski, J. Gruškina find, on the basis of their own corpus, that this cannot be established definitively and that the criteria of the choice between the two ways of marking objects are unclear (Кузнецова, Хелимский, Грушкина 1980 : 382 ff.).

<sup>8</sup> In Lapp unmarked object is found only in the plural, so it is not surprising that it is not found with the imperative singular. In Livonian the nominative and the genitive — the latter the normal case of total object — coincide (except for demonstrative pronouns), but since genitive is the case of such objects in the singular as well, this also cannot be seen as unmarked.

atives, however, when Px-es are used, the object is always unmarked, and the same stands for the so called predestinative conjugation.

Strictly speaking then we may restrict the discussion to the Finnic languages within Uralic (except for Livonian).<sup>9</sup> In these languages the relevant distinction is not between definite vs. indefinite, but partial vs. total object. The unmarkedness of the total object in second person imperatives is a general rule in all of them, as the following examples show.

Finnish:

(23) *lue tämä kirja* 'read this book through',

(24) *ottakaa ehdotus huomioon* 'consider this suggestion'.

Estonian:

(25) *jutusta mulle üks muinasjutt* 'tell me a tale',

(26) *pange see kaart minu lauale* 'put this map on my desk'.

Votic:

(27) *mene ösa papilta lehmä* 'go (and) buy a cow from the priest' (Kont 1963 : 147),

(28) *tširjottagā mokomain papperi* 'write a document like that'.

Karelian:

(29) *ota šõapka* 'put the cap on',

(30) *keittäkeä kohvi* 'make the coffee' (Kont 1963 : 150).

Ingrian:

(31) *oda heppi ilma siutsida* 'lead the horse without the bit',

(32) *avatkas ovi* 'open the door' (Kont 1963 : 148).

Veps:

(33) *mäne soupta ukš* 'go (and) close the door',

(34) *anjkaat mil'ei sigäine* 'give me a piglet' (Kont 1963 : 150).

In all these sentences a case form different from the nominative would be found if the verb was in an indicative finite form, e.g. genitive in Finnish, as in 35 and 36, which correspond to 23 and 24, respectively:

(35) (*sinä*) *luet tämän kirjan* 'you read this book through',

(36) (*te*) *otatte ehdotuksen huomioon* 'you consider this suggestion'.

In those Finnic languages where the total object of a second person imperative verb is unmarked, the same stands for the object of a first person plural verb form, except for Votic, where in such a case a marked object is found. Estonian is exempt from this regularity in a different way on two counts: here the total object of an imperative verb is unmarked also in the third person, whereas it is marked in all the other Finnic languages.

### 2.3. With infinitive

Let us note again that where the unmarked object appears next to finite indicative verbs forms as well,<sup>10</sup> it is natural, i.e. non-specific next to infinitives. (This is the case in Vogul, in the Permic languages, in Mordvin and in the Samoyedic languages.) The converse is not true, however: the fact that in a given language unmarked object is not found next to an indicative verb form does not imply that unmarked object is also not used next to an infinitive. In Hungarian this is found in archaic and dialectal vari-

<sup>9</sup> Classic works on unmarked object in Finnic languages are Grünthal 1941 and Kont 1963.

<sup>10</sup> Lapp will not be included here for the reason explained earlier.

eties, in Cheremis the vicinity of an infinitive is actually the context in which unmarked object occurs most frequently, and it also generally typifies the Finnic languages (again with the exception of Livonian, which only has unmarked object in the "passive voice").

In modern standard Hungarian the nominal object is always marked even when governed by an infinitive. In archaic and dialectal varieties, however, such unmarked forms are encountered:

(37) *az uram oda van fa vágni* 'my husband is away cutting wood',

(38) *jó lesz ez a vödör víz hordani* 'this bucket will be good for bringing water'.

Even present-day Hungarian native speakers would not intuitively judge these structures as definitively ill-formed, and in certain restricted sorts of text (e.g. children's tales) they would be regarded appropriate. But the above spelling and the verb-object structure it reflects seem already dated, and the forms *favágni* 'to cut wood', *vízhordani* 'to bring water' would now be analysed by speakers rather as compounds. Historically, however, it is clear that these structures represent infinitives accompanied by unmarked objects. There must have been a stage in the history of Hungarian when the unmarked and the marked forms coexisted, but it is difficult to say if they systematically contrasted, as, for instance, indefinite vs. definite object. Intuitively these unmarked objects are felt to be indefinite rather than definite. But the question also arises to what extent this unmarkedness can be seen as conditioned by the infinitive (i.e. as specific), since in Hungarian similar structures are found with other kinds of deverbal nominal forms, although only in compounds (see later).

The unmarked object governed by an infinitive is unequivocally specific in Cheremis, since finite indicative verbs there never govern such forms. The object can be unmarked if it immediately precedes the infinitive, like the structures *sömari-šta-š* 'to hold a wedding feast', *üðâr nalaš* 'to buy a woman' in 39 and 40:

(39) *kuçura-k izažân s ö m a r i- šta-š kâle-š* 'the elder brother must hold a wedding feast' (Bereczki 1990 : 34),

(40) *Ološ, talat kuleš ü ð â r nalaš* 'Aljoša, you must get married' (Alhoniemi 1970 : 25).

The unmarked object governed by an infinitive<sup>11</sup> is a typical and well known feature of the majority of Finnic languages. In these languages this is the most frequently used form of total object nouns (partial objects, by contrast, are always marked). The qualification "most frequently used" above is warranted because in the absence of certain other conditions a total object may, and in certain cases must, be marked. Finnish presents a particularly complicated case. An infinitive may govern an unmarked object if the infinitive itself is the subject of the sentence 41, or if the infinitive is immediately the complement of a noun 42, or if the infinitive is not the complement of a finite verbal predicate (or of a noun dependent on it), which agrees with its subject (or if it does the subject is far enough from the predicate to behave as a separate phrase, in which case vacillation occurs 43).

(41) *Sinun olisi parasta mennä lukemaan englannin l ä k s y* 'It would be better for you to go and do your English lesson',

<sup>11</sup> More precisely this is the so called *t*-infinitive, the most general of the infinitives in all Finnic languages.



(42) *Jaakolla oli halu käydä kiskaisemassa joulupukilta p a r t a* 'Jaakko felt like going up to Santa Claus and tearing off his beard',

(43) *Tuulikki piti velvollisuutenaan pestä l a t t i a/lattian* 'Tuulikki regarded it as her duty to mop up the floor' (Leino 1987 : 126–127).

Estonian examples:

(44) *on tarvis ainult ü k s s õ n a ütelda* 'one only has to say a word',

(45) *isa ülesanne oli l a p s kooli viia* 'the father's task was to take the child to school'.

Votic:

(46) *piäb\_ antà t ü t ä r mehelè* 'the girl must be married off',

(47) *tyttäred nii vahtivad kuibõ saisõssi tältä vargassa s e l u s i k k a* 'the girls were watching out for a chance to steal that spoon from him' (Kont 1963 : 171).

Karelian:

(48) *nyt s e pitäy järveh upottua* 'this must now be sunk in the lake',

(49) *pitäv makšoa v e l k a* 'the debt must be paid' (Kont 1963 : 173).

Ingrian:

(50) *toimittaa t ä m ä k y s y m y s, pittää tehdä m o k o m a o o p ь t - t a* 'to solve this question, this trial must be made' (Kont 1963 : 171).

Veps:

(51) *tariz opeta k e l'* 'the language must be taught',

(52) *ol'ž, bašib, ne, tsę h ä n d' k a z rikta* 'one should, he says, kill this wolf' (Kont 1963 : 173).

#### 2.4. With other non-finite verbal and deverbal nominal forms

This phenomenon in a specific form is latent in Hungarian and overt in Cheremis. In Hungarian it underlies compound words that historically go back to object + deverbal (of various sorts) phrases, as in 53–56:

(53) *favágó* 'woodcutter',

(54) *favágás* 'woodcutting, logging' (noun),

(55) *(?)favágva* '(while) cutting wood' (adverb),

(56) *(?)favágtlan* 'without having cut wood' (adverb).

The occurrence of the last two examples is questionable, but they are not entirely alien to the general pattern of the language (cf. the widely attested adverb *dolgavégezetlen* 'without having succeeded'). There is no doubt that all these compounds go back historically to a structure involving unmarked object.

In Cheremis, by contrast, the unmarkedness of objects governed by participles displays the same regularity as that of objects governed by infinitives:

(57) *nənəlän mārəwläžəmät uwläm, k i t - j a l lültəšəwläm puaš keleš* 'they must be given new songs, uplifting ones' ('uplifting' literally 'lifting up their hands and legs') (Bereczki 1990 : 34).

#### 2.5. With "passive" verb forms

Although it is a phenomenon amply discussed in the literature and well known to learners just as much as to specialists, I note here that the grammatical form called passive in the Finnish grammatical tradition (including

works on Finnish in other languages) is unlike the passive structure seen in the Indo-European languages. This latter sort — i.e. the structure involving a transitive verb in passive voice, with a patient in the nominative, and an agent, if any, in an oblique case — is also attested in a number of Uralic languages, but these patterns are, in theory, incompatible with a grammatical object and *a fortiori* with an unmarked object. The "passive" structure that I want to look at in connection with the unmarked object is indeed "passive" in inverted commas only. Estonian grammars mention an impersonal conjugation and an impersonal structure, which is not necessarily incorrect, but it must be noted that in linguistic tradition the term impersonal is usually reserved for zero-argument verbs like Hungarian *hajnalodik* 'the day is breaking', Russian *дождёт* 'it rains', Latin *ninguit* 'it snows', but these have nothing to do with the Uralic structures to be discussed here. These structures are best described as transitive structures with a general subject. What this implies is that the exact translation of Finnish *laulu lauletaan* or Estonian *lauldakse laul* is 'one sings the song' rather than 'the song is (being) sung'. The word in the nominative here is not a subject but an unmarked object. This is clearly shown among others by the corresponding negative, in which it assumes the partitive case, something that the subject of a normal sentence<sup>12</sup> cannot do.<sup>13</sup> Furthermore, in all these languages intransitive verbs also have "passive" forms; this again shows that these are not passive but general-subject-forms, which naturally do not — indeed cannot — agree with any other constituent.

The unmarked object governed by a "passive" verb form occurs only in the Finnic languages (but in all of them) within the whole of the Uralic family (in Liv this is its only environment apart from compounds). Examples are given in 58–64.

Finnish:

(58) *a s i a jätettiin lepäämään yli vaalien* 'they put the matter aside until after the elections'.

Estonian:

(59) *a s i pandi valimistele järgneva ajani seisma* 'they put the matter aside until after the elections'.

Livonian:

(60) *l ō d a um pūstō tiedōd* 'the table is made of wood' (Kont 1963 : 161).<sup>14</sup>

Votic:

(61) *mentii kutsuttii s õ t a m e e s* 'one went (and) called a soldier' (Kont 1963 : 156).

<sup>12</sup> The term "normal sentence" contrasts with the category of existential sentences in the Finnish grammatical tradition. If the generally accepted analysis is correct and the subject of existential sentences is indeed a subject (which I, for one, doubt, see Havas 2006a), then this is a subject that assumes the partitive case in negation.

<sup>13</sup> Further evidence for this being an object rather than a subject comes from the behaviour of personal pronouns, which are always in the accusative in the vicinity of "passive" verb forms (except in Votic, where the rule for nouns is generalised), e.g. Finnish *poika* (Nom.) *kutsuttiin kylään* 'one invited the boy' vs. *hänet* (Acc.) *kutsuttiin kylään* 'one invited him'.

<sup>14</sup> The complex verb forms including the verb 'be' and a participle can be analyzed in two ways: (i) subject + copula + participle; in this case the noun in nominative is a subject; (ii) object + complex "passive" verb; in this case the the noun in nominative is an unmarked object. The partitive case appearing in the corresponding negatives makes the latter analysis more plausible.

Karelian:

(62) *da šoatih kiini se vora* 'and one caught that thief' (Kont 1963 : 158).

Ingrian:

(63) *vêtti se poiga kuningahalle* 'one led the boy to the king' (Kont 1963 : 156).

Veps:

(64) *noŋtsja noŋdaze ka ŋgas* 'one makes linen on a loom' (Kont 1963 : 159).

## 2.6. Numerals as unmarked object

When the structure of a numeral + NP phrase is parallel to that of ordinary adjectival phrases, and thus the numeral is syntactically subordinate to the adjacent noun (i.e. the noun is the head of the phrase that includes the numeral), the formal behaviour of the numeral is analogous to that of adjectives and consequently cannot be regarded as specific to the object phrase. This is the case with those NumP's in the Uralic languages (including object phrases) where the case form of the NP following the numeral depends only on the syntactic function of the NumP (or its head), and not on the numeral itself. In such cases the markedness or unmarkedness of the numeral that functions as a modifier in an object phrase follows from the general principles of the language: is there object morphology at all? if yes, then do adjectives agree with the modified noun? if yes, then is the nominal head of the object phrase marked or unmarked? If, then, in such cases a numeral is unmarked in object function (or as part of the object), there is nothing specific about it (i.e. specific to numerals and their behaviour as object).

By contrast, if the head of the NumP is the numeral and the NP is the subordinate element, the NP (or its head) assumes the non-nominative case governed by numerals. This is seen in the Finnic languages, where a noun governed by a numeral other than 'one' in (subject or) object function assumes the partitive case.<sup>15</sup> When it heads an object phrase, a numeral remains unmarked in all these languages. The following examples are from Finnish 65, 66 and Estonian 67, 68.

(65) *Näin vain yhden pienen sorsan* 'I only saw one duckling',

(66) *Näin ka k s i pientä sorsaa* 'I saw two ducklings',

(67) *Ostsin ühe vihiku* 'I bought one exercise book',

(68) *Ostsin ka k s vihikut* 'I bought two exercise books'.

---

<sup>15</sup> If the NumP is in a case other than nominative or accusative, its structure parallels that of ordinary adjectival phrases with all numbers, i.e. the noun is the head and the form of the numeral is a function of the case form of the noun (full or partial agreement). The same stands for the numeral 'one', which behaves as an ordinary adjective in all respects. Lapp provides an illusory parallel: in that language the accusative of numerals other than 'one' coincides with the nominative also when the numeral is used a noun. Furthermore, the singular genitive of nouns (used also after higher numerals) coincides with the accusative, and thus if such a phrase is in the accusative it is impossible to decide which constituent is the head, and the unmarkedness of the numeral cannot be regarded as the consequence of its object function.

## 2.7 Unmarked object in compounds

The traditional division of compounds into coordinating and subordinating may be right formally, but the underlying hypothesis that these originate historically in coordinating and subordinating structures, respectively, does not always stand up to scrutiny. It is true especially of coordinating compounds that even the theoretical outlines of the reconstruction of such a putative proto-structure are hard to arrive at.<sup>16</sup> The reconstruction is difficult also for object compounds whose head is a deverbal nominal category. In such cases two kinds of problems may emerge. One is the unclear diachronic relation between participles and deverbal nouns;<sup>17</sup> whether there is any such relation at all, and if there is, what is its direction. (If participles provided the pattern for deverbal nouns, the appearance of a complement — in this case the object — with the latter is not necessarily the result of syntagmatic contraction with the noun. The other problem is that these compounds may be based on semantic analogies and mechanisms of a more abstract kind, without syntactic underpinning. Nevertheless, in the case of those subordinating compounds where the head is a verb or a participle, it appears to be a plausible hypothesis that at least a core subset of such compounds which provided a basis of analogy had emerged from the contraction of syntactic constructions.

Naturally, it has to be borne in mind that the unmarkedness of the object within compounds is specific only if the same element would be marked for the accusative when governed by the finite indicative forms of the same verb. What this means is that of the languages discussed here the unmarkedness of the object in compounds is specific only in Hungarian, Cheremis and the Finnic languages.

The Hungarian examples 53–56 could be repeated here; strictly speaking they actually belong here, along with a few, mostly fictitious, deverbal adjectives such as *favágós* 'one who likes cutting wood or cuts wood habitually' and their derivatives such as *favágóskodik* 'makes his living (temporarily or by necessity) by cutting wood'. Since, however, *favágós* is already a derivative of *favágó*, we may as well rely on our earlier examples when discussing the object. I would find it hard to argue against the position that the first member of these compounds is historically an unmarked object.

Cheremis shows the exact opposite of Hungarian. It is questionable if Cheremis actually has subordinating object compounds.<sup>18</sup> The following pair of examples suffices to demonstrate this:

<sup>16</sup> Consider for instance the Hungarian neologisms *látványpékség* lit. 'spectacle' + 'bakery', i.e. 'a bakery where bread is baked in front of the patrons' and *szendvics-generáció* lit. 'sandwich' + 'generation', i.e. 'a generation that has to care for aging parents as well as children (a generation between two other generations like the contents of a sandwich between the two slices of bread)', which are totally transparent compounds, but it is difficult even to paraphrase them, let alone reconstruct a source sentence in which the members of the respective compounds would be autonomous but consecutive with the same meaning. The factors that gave rise to these compounds are obviously analogy, semantics, even free association. On the frequent syntactic problems in deriving compounds, see Fejes 2005.

<sup>17</sup> I also do not claim to know what the relation is between participles and deverbal adjectives, since I do not accept the commonly held opinion that participles were originally forms with an attributive function. For this question see Havas 2006b.

<sup>18</sup> G. Bereczki's outline of Cheremis grammar does not even list such a category among the types of subordinating compounds (Bereczki 1990 : 70).

- (69) *колым кучымаш*,  
 (70) *к о л кучымаш*.

Both expressions mean 'fishing', literally 'fish-catching' (cf. *кол кучаш* 'to catch fish'). The fact that deverbal nouns can be accompanied both by the unmarked and marked accusative forms implies that these are likely to be object constructions rather than compounds.<sup>19</sup> Cf. also 71—72:

- (71) *к о л кучышо* 'fisher = fish-catcher',  
 (72) *пйчым ончышо* 'deer breeder' (*пйчö* 'deer', *ончылташ* 'to oversee'),  
 where one of two deverbal nominal elements formed in the same way is accompanied by an unmarked, the other by a marked, object. The conclusion to be drawn is not only that the use of unmarked object in Chereemis is not governed by categorical rules, but also that the behaviour of objects governed by participles and other deverbal nominal elements is essentially the same. Thus the occurrences of unmarked object in 2.4 and 2.7 do not form two disjunct categories.

Finnic languages will be here represented by Finnish. As the survey of the "Iso suomen kielioppi" (2004 : 401) clearly shows, the large majority of the relevant deverbal suffixes (16 out of 18) derives nouns that can govern the original object of the verb in an unmarked as well as a marked form, often with the same lexical items, and no systematic dichotomy can be observed between the meaning or the use of the two kinds of compounds — even if native competence occasionally distinguishes between the use of the two variants. Examples are given in 73—78.

- (73) *tiedonhaku* 'data-gathering (lit. information-seeking)' (*tieto + n + haku*),  
 (74) *h e n k i l ö haku* 'casting (lit. person-seeking)' (*henkilö + haku*),  
 (75) *levontarve* 'need for rest' (*lepo + n + tarve*),  
 (76) *a s u n t o tarve* 'demand for flats' (*asunto + tarve*),  
 (77) *metsänhakkuu* 'forest-clearing' (*metsä + n + hakkuu*),  
 (78) *m e t s ä hakkuu* 'forest-clearing' (*metsä + hakkuu*).

The appearance of *-n-* between the two members of the compounds can be interpreted in a variety of ways. Since example of the kind 73, 75, 77 are traditionally classified as subordinating object compounds, it seems reasonable to equate this *-n-* with the accusative suffix *-n* (from Proto-Uralic *\*-m*). Since, however, the non-nominative marker of (total) object in modern Finnic languages is identical to the genitive ending (i.e. the suffix *-n-* can be analysed as a genitive marker of a different provenance), these compounds could just as well be seen as possessive compounds.<sup>20</sup> In view of the fact that this *-n-* element may also appear in Finnish subject compounds — just as unpredictably as in the previously seen types (see for instance *auringonnousu* 'sunrise' from *aurinko + n + nousu* vs. *kurssilasku* 'exchange rate sinking' from *kurssi + lasku*) — one would be justified in asking the question whether this *-n-* has anything to do with the accusative or the genitive, or whether it is simply

<sup>19</sup> For a full demonstration of the validity of such a claim further tests would be necessary to see e.g. what can appear between the object and the noun, whether they are separable etc. This is not germane to the concerns of this paper, since the question I am seeking to answer is not whether these structures are compounds but whether the object in these structures is unmarked or not, and the answer clearly is that it *may* be unmarked.

<sup>20</sup> Similarly to the Hungarian *favágó, favágás* type, where the first member is often described (in my view erroneously) as an unmarked possessive form: 'the cutter/cutting of wood' etc.

an optional infix of compounds. On the first interpretation an unmarked object appears in the *n*-less variants of subordinating object compounds, on the third interpretation in all instances of such compounds.<sup>21</sup> *Mutatis mutandis* the same stands for the rest of the Finnic languages.<sup>22</sup>

### 3. The origins of unmarked object

#### 3.1. Hypotheses concerning the emergence of unmarked object

As I pointed out at the beginning of this paper, it is a commonly held view in Uralic linguistics that in Proto-Uralic (and consequently in Proto-Finno-Ugric as well as in Proto-Samoyedic) there were objects marked with the accusative case ending *\*-m* and there were also unmarked objects. What this means is that the object was sometimes marked, sometimes unmarked. In spite of the fact that in some of those Uralic languages that have retained this bifurcation the conditioning factor of the choice between the two forms is not clear, linguists tend to unanimously assume that in Proto-Uralic their distribution was rule-governed. The general opinion is that markedness reflected the definiteness of the object, while unmarkedness its indefiniteness.<sup>23</sup> This virtually unchallenged view is surprising at least with respect to Proto-Uralic since, as I pointed out above (2.1), the evidence of Samoyedic languages could just as well prompt the opposite conclusion, i.e. that unmarked object was definite and marked object was indefinite. It must be noted that there are hypotheses competing with the idea of the definite vs. indefinite distribution; it has been proposed that *\*-m* was originally an identificatory element (Grünthal 1941 : 278), it may also be surmised that it served the purpose of topicalisation,<sup>24</sup> or to distinguish affected vs. effected object, or perhaps internal object vs. *accusativus obiectivus*.<sup>25</sup> The common denominator of all these hypotheses is that — whatever its logic — the contrast existed in the proto-language and perhaps unmarkedness was the dominant form of objects.

In view of this almost complete uniformity of opinion, it is all the more surprising that when discussing the unmarked object in the individual languages or subgroups it is only a small minority of linguists who derive this phenomenon — or at least some type of it — from the Proto-Uralic unmarked object.<sup>26</sup> The general feeling seems to be that it is an internal

<sup>21</sup> On the second interpretation these compounds do not include objects.

<sup>22</sup> Livonian compounds invite further investigation as do Lapp (a non-Finnic language) compounds, where the unmarkedness in the singular would be specific. The literature known to me is silent about this question.

<sup>23</sup> Cf. Wickman 1955 : 146; Collinder 1960 : 239; Hajdú 1966 : 65; Майтинская 1974 : 243 ff, among others. H. Katz (1980 : 398) excludes this interpretation (albeit without detailed argumentation) on the basis of typological considerations and also because it is not in harmony with his ergative hypothesis.

<sup>24</sup> An argument for this point could be based on the analogy of the putative derivation of Finnish and Ostyak pronominal accusative *-t* from a demonstrative pronoun (Kulonen 1999 : 63–71; Havas 2006a), or of Hungarian accusative *-t* from a topic marker (*expressis verbis* in Marcantonio 1988 : 156 ff).

<sup>25</sup> This latter hypothesis is (or could be) based on J. Zsilka's notion of transformational group (1973 : 44 ff; 1981 : 72 ff).

<sup>26</sup> Exceptions include E. N. Setälä (1926 : 23); W. Grünthal (1941 : 291–292); E. Itkonen (1971–1972 : 184); to this list K. Korompay (1991 : 284–285) may be added, who appears to take it for granted that the Pre-Hungarian unmarked object derives trivially from Proto-Uralic.

development of the individual languages/subgroups, and the various types of unmarked object are not given uniform treatment.

At one end of the spectrum of explanations we find its occurrence next to a "passive" in Finnic, a pattern derived by virtually everyone from syntactic reanalysis. It is generally believed that the verb form now seen as "passive" or impersonal used to be an active factitive or reflexive — for some linguists, finite — form, and the noun in question was its subject. The reanalysis of the subject as (unmarked) object took place in the common Finnic period at the earliest.<sup>27</sup> At the other end of the spectrum we find the unmarked object of compounds based on participles, which is not normally explained with reference to a subject or to an object whose ending has eroded. At the same time, one can always resort to derivation from an (unmarked) possessive.<sup>28</sup>

The explanations of all the other types of this phenomenon can be placed between these two endpoints. All authors strive to demonstrate, albeit with different arguments, that the unmarked object is either not completely an object, or is completely a non-object. B. Wickman, for one, says the following on its occurrence next to an infinitive: "This is probably a heritage from ancient times, when the present object of a verbal noun was not a real object but an attribute or the first part of a compound" (1955 : 73). V. Kiparsky demonstrates that this phenomenon is an archaism found not only in Finnic but also in northern Russian dialects, Latvian, Lithuanian, Old Czech, Old Bulgarian and Sanskrit, but assumes that it appeared independently in Indo-European and Finnic (allowing for borrowing from Finnic in the Baltic languages). He agrees with J. Endzelins, whose Latvian grammar he cites: "Als der Infinitiv noch eine lebendige Kasusform war, konnte das Objekt der durch den Infinitiv ausgedrückten Handlung als grammatisches Subjekt und Nominativ erscheinen, und auch nachdem der Infinitiv zu einer rein verbalen Form geworden ist, kann sein Objekt noch immer als grammatisches Subjekt im Nominativ erscheinen, wenn der Infinitiv durch ein modales Adverb bestimmt ist." (Kiparsky 1959—1960 : 334). It remains unclear to me how the infinitive's being a living case form could explain that its object is a subject, especially if one assumes that the infinitive was not a verb (as V. Kiparsky claims); thus this explanation is more of a *petitio principii*. The same is true of the unmarked object governed by the imperative. J. Janhunen, for instance, writes the following: "Since the marked accusative (in \*-m) is the normal case [sic!] of the object in Uralic, the nominative object of imperative sentences immediately suggests the possibility of prehistorical syntactic restructuring. There is no need to assume that the synchronic nominative "object" has always been an object. Rather, since it is in the unmarked nominative case of the *subject*, it is likely to have been the grammatical subject of the sentence." (Janhunen

<sup>27</sup> See Lehtinen 1984; Keresztes 1996 and the references there.

<sup>28</sup> Not that this last one can be taken seriously as a diachronic explanation. Although the first member of *favágó* 'wood-cutter' can be analysed as a *genitivus obiectivus*, but only grammatically; in semantic terms, it definitely does not represent a possessive relation. This is why a marked genitive renders the structure highly unnatural (*a fának a vágója* 'the cutter of the wood'), and a possessive sentence is outright absurd (*a fának van egy vágója* 'the wood has a cutter'). Also bear in mind that the noun *favágó* has, and can have, no Px in its base form, which would be compulsory in a Hungarian possessive structure.

2000—2001 : 78—79). Since the premiss introduced by *since* in the quotation is clearly unfounded (why would the marked accusative have been a “more normal case” of the object in Proto-Uralic than the unmarked form?), the conclusion is spurious.<sup>29</sup> The unmarked accusative governed by transitive finite verbs in Mordvin and Samoyedic can hardly be explained with reference to a reanalysis of the subject, and it would make no sense to talk about an apocopated accusative suffix in languages where the accusative suffix did not actually apocopate. It would be even harder to explain the unmarkedness of numerals as objects with reanalysis.

In fact, there is no reason to believe that the occurrences of unmarked object in Uralic — apart from the Finnic “passive” — are not archaic structures attributable to the proto-language. The fact that other families show very similar phenomena does not, of course, imply common origin, but this is compatible with the assumption that they are archaic features in all of them. Unmarked object ultimately emerges from a prenominal stage — this is what I shall now trace with the help of the outlines of what I call a schematogony.

### 3.2. On the typology of sentence patterns

#### 3.2.1 The transitivity pattern

At the beginning of this paper I discussed the definition of transitive verb and object. With reference to these categories, let us take a closer look at the essence of nominative structure and compare it to other sentence patterns.

I shall begin with the so called transitivity pattern. This can perhaps be schematically represented in the simplest way as follows (tacitly assuming an SVO language):

(79)	I.    intransitive verb	
	II.    transitive verb	III.

The Roman numbers in the chart refer to the compulsory complements of intransitive and transitive verbs, respectively. In nominative languages like Hungarian, German or English etc., I and II correspond to the subject of verbs, whereas III to objects.<sup>30</sup> These entities are syntactic positions or functions (and

<sup>29</sup> J. Janhunen entertains two possible explanations for the reanalysis. One, which he considers less likely, involves original ergative structures, but no serious arguments are adduced, and it is also a fact that in ergative structures the — mostly unmarked — absolutive is not a “subject” at all; the other explanation refers to a reanalysis that parallels the “passive” — though it is again unclear how.

<sup>30</sup> Other abbreviatory conventions for the same items are S<sub>i</sub>, A and O, respectively (cf. Mallinson, Blake 1981 : 41, where the terms are intransitive subject, transitive subject and object), or simply S, A, O, as in Dixon 1994 : 6, where the terms are literally the same, although Robert M. W. Dixon later condemns the subordination of the first two functions to the notion of subject. Neither book explains why they use precisely these abbreviations, but the source quite clearly appears to be Comrie 1978 : 330 ff., though with P instead of O. Bernard Comrie, for his part, draws attention to the tentative and imprecise nature of these terms (subject, agent, patient). On closer scrutiny the terms cannot possibly correspond to the functions they are supposed to represent. Because of this, some authors resort to the letters X, Y, Z (Itkonen 2001 : 211 ff.); I have decided in favour of neutral numbering.



not arguments as logicians are wont to call them). Distinctions between them can be manifested through case marking (a variant of which is the use of appositions like the postpositions in Japanese or Korean), verbal morphology (whether it be autonomous or concord-marking), word order or any combination of these. For a simple illustration at this point case marking is invoked.

Since no language meets the basic requirements of communication without being able to distinguish between *The hunter killed the tiger vs. The tiger killed the hunter*, functions II and III always differ in a transitive sentence. There are then three possibilities:

- a) I ≠ II ≠ III — triadic construction,
- b) I = II (but ≠ III) — nominative construction,
- c) I = III (but ≠ II) — ergative construction.

### 3.2.2 The nominative pattern

Nominative sentence construction is obviously so called because of the nominative case, the case of the subject (I = II). (It is also frequently referred to as accusative, an unjustifiable practice, as we shall see). In fact, nominative is the most abstract member of the case system and therefore the most difficult to define semantically. It encompasses an embarrassingly wide range of content, but it is not designated as the sole exponent of any of the semantic categories. Let us look at the following examples:

(80) *The boy is cold* (patient of state)

*The boy coughed* (agent [or patient?] of unintentional action)

*The boy ran* (agent of intentional intransitive action)

*The boy is eating bread* (transitive affecting agent)

*The boy is baking bread* (transitive effecting agent)

*The boy is washing* (agent of reflexive action)

*A bottle of wine was given to the teacher* (patient of transitive action)

*The teacher was given a bottle of wine* (recipient of transitive action)

*This violin is easy to play the sonata on* (locative of transitive action) etc.

Let us also note that, apart from the subject, nominative is also the case of nominals functioning as predicate or as part of a predicate.

With transitive verbs the nominative case contrasts with the accusative case (= III). If the constituent order is fixed, both nominative and accusative may be unmarked morphologically (as in English or French with lexical NP-s), otherwise — even with fixed constituent order — the following two types are observed:

(a) Subject unmarked, object marked:

(81) Hungarian *A lány fut* 'The girl- $\emptyset$  runs'

*A lányok futnak* 'The girls- $\emptyset$  run'

*A lány látja őket* 'The girl- $\emptyset$  sees they-ACC'

*A lányok látják őket* 'The girls- $\emptyset$  see they-ACC'

(b) Both subject and object marked:

(82) Latin *Domin-us agricola-m laudat* 'Master-NOM<sub>SG</sub> praises (the) farmer-ACC<sub>SG</sub>'

*Domin-us agricola-s laudat* 'Master-NOM<sub>SG</sub> praises (the) farmer-ACC<sub>PL</sub>'

*Domin-i agricola-m laudant* 'Master-NOM<sub>PL</sub> praises (the) farmer-ACC<sub>SG</sub>'

*Domin-i agricola-s laudant* 'Master-NOM<sub>PL</sub> praises (the) farmer-ACC<sub>PL</sub>'

In nominative languages, if there is any agreement involving the verbal predicate, it either relates the verb to the subject only (as in Indo-European

languages) or to the subject if the verb is intransitive and to both the subject and the object if the verb is transitive (as in Hungarian). Agreement with the object is found only in a minority of nominative languages, and no nominative language has agreement exclusively with the object.

Nominative languages usually have passive construction as well. Though this tends to be described both traditionally and in more recent frameworks as secondary, in terms of its origin (and also in terms of speaker intention) it is actually independent of what is called active construction. Its most salient feature is the representation of the patient of the active sentence as subject. Since it is also based on the nominative case, the passive construction is nominative just like its active counterpart.<sup>31</sup>

The overwhelming majority of Indo-European languages are purely nominative, and so are many that belong to the Uralic, Altaic and other families.

### 3.2.3 Ergative constructions

As shown in the chart above, in these constructions functions I and III are equivalent and they contrast with II.<sup>32</sup> Where this is expressed formally on the nominal elements, the corresponding grammatical cases cannot be identified either as nominative or as accusative. After a period of hesitation the literature settled about fifty years ago on the following terms: I = III (also for nominal predication) — absolutive; II — ergative. To take an example from Yalarnnga (Pama-Nyungan, Australia) (see Mallinson, Blake 1981 : 49—50):

(83) *kupi waya kunu-ŋka* 'The fish is in the water'

fish the water-LOK

*kupi-ŋku miŋa taca-mu* 'The fish swallowed the fly'

fish-ERG fly swallow-PAST

*na-tu kupi wala-mu* 'I killed the fish'<sup>33</sup>

I-ERG fish kill-PAST

The word meaning 'fish' is in the absolutive case (unmarked in this language) when it is the "subject" of an intransitive verb (first example) or the "object" of a transitive verb (third example), and in the ergative case when it is the "subject" of a transitive verb (second example).<sup>34</sup> The verb does not agree with any of the three positions.

<sup>31</sup> Active sentence structures and active languages will be discussed later. It is an infelicitous terminological coincidence that the active construction in nominative languages is called by the same name; the two are in fact unrelated.

<sup>32</sup> In theory, the contrast could be expressed purely through constituent order (e.g. I = III always before the verb, II always after the verb), but I know of no such case of morphologically unmarked ergativity. Therefore here only cases of morphologically marked ergativity will be discussed.

<sup>33</sup> Note that in phonetically transcribed examples I left out some of the typographically problematic diacritics, which are irrelevant to the morphological and syntactic phenomena under discussion.

<sup>34</sup> We leave for later discussion the erroneous idea of the universality of the terminology that has become general in the description of nominative languages. Here we only note that the terms "subject" and "object" are definitely inadequate (as inherently nominative), and are only used for comparison here. The quotation marks indicate that here we are talking about constituents that would be translated into a nominative language as subject and object.

In some instances verbal morphology also displays ergative agreement. The following examples are from Avar (Caucasian) (see Mallinson, Blake 1981 : 56):

- (84) *či v- ač-ula* 'The man comes'  
 man MASC come-PRES  
THINKING  
*ebél-alda či v- at-ula* 'The mother finds the man'  
 mother-ERG man MASC find-PRES  
(SUPÈSS) THINKING

In this language too absolutive is not marked by a case ending, but is cross-referenced by agreement — as can be seen, the verb agrees with 'man', the element in absolutive function — while the ergative case has its dedicated ending (which happens to coincide with that of superessive), but no agreement cross-references it. Agreement involves a prefix that refers to the noun class: *v-* for thinking male beings, *y-* for thinking females and *b-* for everything else.

To take a language in which absolutive is marked, let us look at Georgian:

- (85) *dzagh-l-i bagh-ši da-i -mal -a* 'The dog hid in the garden'  
 dog-ABS garden-DAT he-INTR-hide-AOR3SG  
*bic'-ma dzagh-l-i bagh-ši da-mal -a* 'The boy hid the dog in the garden'  
 boy-ERG dog-ABS garden-DAT he-hide-AOR3SG  
*bic'-i bagh-ši da-rch- a* 'The boy stayed in the garden'  
 boy-ABS garden-DAT he-stay-AOR3SG

That ergative is a purely syntactic position is highlighted by the fact that it is not confined semantically — as one would expect — to intentional agents (or to agents of any kind, for that matter). Compare the use of the word meaning 'axe' in Yidin (also Pama-Nyungan) (Dixon 1994 : 59—60):

- (86) *wagu:ja-ŋgu jugi gunda-l (galba:n-da)* 'The man is cutting the tree  
 man-ERG tree cut-IMPF axe-INSTR (with an axe)  
*galba:n-du wagu:ja gunda-:ji-ŋ* 'The axe cut the man' (sc. accidentally)  
 axe-ERG man cut

In the description of nominative languages functions I = II are summarily called subject, and III is called object (and the cases that express them are called nominative [or subjective] and accusative [or objective], respectively). For ergative languages there are no comparable terms for I = III vs. II, though the terms denoting the cases that express them could be borrowed for this purpose (cf. Palmer 1994 : 15).<sup>35</sup>

Just as nominative languages have passive construction, many of the ergative languages have a "secondary" type of sentence structure called antipassive. In antipassive constructions the nominal element in function II assumes the form reserved for I = III in an ergative construction (absolutive), while the nominal element "originally" in function III either assumes an oblique case (mostly instrumental) or is left out entirely (as is compulsory in Avar). In the latter case its absence is sometimes indicated by a dedicated marker (or, in certain languages, the absence of an otherwise

<sup>35</sup> In which case, though, it would be desirable to consistently distinguish between ergative qua case vs. ergative qua syntactic function, perhaps in the form of ergative vs. ergate (along the lines of subjective and objective vs. subject and object). Since, however, such terminological rigour is unlikely to be imitated, and since it is typologically irrelevant, I will not insist on it here.

expected marker) normally in function II, as in Bandjatang or Archi. The antipassive construction usually has two semantic functions: to express incomplete or habitual action on the one hand, or unspecified patient on the other. Yalarnnga examples show this (first sentence ergative, second antipassive):

- (87) *matyumpa-yu kukapi taca-mu* 'The kangaroo ate the grass'  
kangaroo-ERG grass eat-PAST  
*matyumpa kukapi taca-li-ma* 'The kangaroo eats grass'  
≈ 'Kangaroos' are grass-eating animals'

Antipassive is diachronically as well as logically independent of the standard ergative construction, from which descriptive practice tends to derive it (just like passive is independent of active), and since it is based on the absolutive it does not go beyond the basic sentence construction principle of ergativity.

Ergative languages (i.e. languages that are exclusively or predominantly ergative in terms of sentence structure) include Basque, many of the Caucasian languages (Avar, Chechen, Lak, Mingrelian-Chan, Abkhaz-Adige, Nakh-Dagestanian, partly Georgian etc.), of the Pamir languages Burushaski, many of the Australian (Papuan) languages, of Amerindian the Algonquian languages (among others Mohican, Delaware, Cree, Illinois [now extinct]), some of the Paleo-Siberian languages like Chukchi. In the perfect tenses Hindi and Urdu (if taken as a distinct language; Indo-Aryan) transitive verbs enter into an ergative construction (these are thus split systems). Note that in these languages the ergative is marked with a post-position.

### 3.2.4 Triadic systems

There are very few examples of triadic, or three-way, systems (also called tripartite in the literature). The following pair of sentences is from Wankumara (also Pama-Nyungan, Australia) (Mallinson, Blake 1981 : 51).

- (88) a. *kana-ia palu-ya* 'The man died'  
man-STAT die-PAST  
b. *kana-ulu kalka-ya titi-nana* 'The man hit the bitch'  
man-AGNT hit-PAST dog-ACC (FEM)

This configuration differs from both of those discussed so far. The case of the "subject" (function I) of 88a is not nominative, because it should then coincide with that of the "subject" (function II) of 88b. But then it is also not absolutive, because it differs from the "object" (function III) of 88b. Note, however, that the inverted commas are superfluous in the last case. If the patient of a transitive verb that also has an agent shows specific marking, then it is the same as what is called object or accusative in a nominative construction. We shall return later to a discussion of the important fact that triadic languages have objects in the same sense and the same way as nominative languages have; but we note here that for the other two case forms we must give up the use of the previously introduced terms. The case that exclusively typifies function I may be called stative. As for the case of function II, in spite of appearances it cannot be fully equated with the ergative case. Although the ergative also meets the definition (obligatory non-patient complement of a transitive verb), it must

be borne in mind that in ergative languages function II contrasts with the absolutive case (I = III), whereas in triadic languages with two other cases (I ≠ III). Therefore I shall use the term agentive for function II in triadic languages (ergative is clearly out of question in non-ergative languages).

Since case-marking governed by transitive verbs differs at two points from that governed by intransitive verbs (as opposed to the nominative and the ergative constructions), the marking of transitivity is redundant in triadic languages. This may be one reason for there being so few such languages in the world.

Further examples of triadic languages include Galali (also Pama-Nyungan) and Dyalany (West Australia). In the latter, stative is unmarked, agentive and accusative are marked with two different suffixes, as one would expect, apart from the 1Sg personal pronoun, which has identical forms for functions I and II. It is noteworthy that a triadic system appears sporadically in some Indo-European languages too, for instance in the past tenses in Yasgulami (South-West Iranian).

### 3.2.5 Active constructions

Active constructions<sup>36</sup> are important to an understanding of transitivity, though in the present argument they serve rather as background information. Those typological studies which actually discuss this structure tend to derive it from the ergative construction though, in reality, it is based on entirely different principles. It is generally held that the only difference between active and ergative construction is that in the former function I bifurcates into an inactive and an active variant (i.e. intransitive verbs fall into two groups, one which has unergative or agent subjects and one which has unaccusative or patient subjects). This means that there is a function I<sub>active</sub> (I = II) and a function I<sub>inactive</sub> (I = III). An example is here taken from Dakota (a well known representative of Sioux), where activeness is marked on the verb (Климов 1977 : 128–129 and *passim*).

(89) *ma-ta* 'I die'

*ma-wašte* 'I am good'

*ma-kaška* 'he binds me'

*ma-ya-kaška* 'you bind me'

*wa-ti* 'I dwell'

*wa-kaška* 'I bind (it)'

As can be seen, the prefix *ma-* indicates a first person singular patient on both the intransitive and the transitive verb, whereas the prefix *wa-* indicates a first person singular agent in all environments (and *ya-* indicates a second person singular agent).

However, the claim that this dichotomy of functions is only typical of intransitive verbs (i.e. in function I) is erroneous. First, as opposed to intransitive verbs, most of which are probably medial, transitive verbs mostly express actions, and so they are obviously less likely to be accompanied by patient subjects. Thus the implicit assumption that if they had a patient subject it would nevertheless receive the same form as transitive agent subjects do is less than self-explanatory. Second, there are languages

<sup>36</sup> Also called agentive, though the term used here is much more current in the literature (see Palmer 1994 : 65 ff.).

in which the choice of the agent—patient pattern is not bound to lexemes but rather ad hoc: the speaker decides and grammatically marks on each occasion (for each occurrence of a verb) whether he views the event described by the predicate as an intentional act or as a "spontaneous" event. This is what we see in the Muskogean family (Algonquian), where e.g. Choctaw *čokma* '(to be) good' means 'I do good' with an agent variant of the first person singular pronoun, 'I am good' with the patient variant and 'I am well' with the dative variant of the same (Dixon 1994 : 33–34). A third argument is that in languages that have active syntactic constructions (or simply active languages) it is a conspicuous universal that if there are finite verbs (as in all currently spoken active languages), the person markers of the active and the stative verbal paradigms form two distinct sets. In some active languages there is even a third paradigm (or set of person markers) for involuntary actions like verbs of perception and mental states (*verba sentiendi et affectuum*) or 'laugh', 'sleep' and the like (this is found, for instance, in Assiniboin (Sioux family, Sioux phyle)). Of these, verbs of perception traditionally count as transitive, but their "subject" is not marked in the active way (unless it denotes voluntary perception, e.g. when two distinct paradigms of one and the same lexeme express 'see' and 'look'). What follows from all this is that in active languages function II also "bifurcates" in terms of the agent—non-agent parameter. If this "bifurcation" is typical of both function I and II, then in these languages functions I and II do not actually exist, since sentence constructions depend purely on the agent—patient dichotomy irrespective of whether the verb is transitive or intransitive. Of course, a language can always be described in terms of categories that were worked out for other language types (or categories that are held to be universal), but an understanding of how the language works can only be based on the categories that analysis can actually find in it. On this assumption active construction cannot be placed into the pattern we started out in 3.2.1, since *t r a n s i t i v i t y a s a n o r g a n i c p r i n c i p l e s i m p l y d o e s n o t o p e r a t e* in these languages.<sup>37</sup> What does operate is an active—stative dichotomy, furthermore, since it is here that this dichotomy essentially determines sentence construction, in a syntactic sense it only exists in these languages. (In any other case, including the language types previously discussed, it is merely an abstraction resulting from semantic analysis.)

In active languages the lexicalisation of verbs also reflects the active—stative rather than the transitive—intransitive dichotomy (active vs. stative verbs). In some cases a verb can be assigned to either category (in terms of inflection and potentially also case assignment). In a different terminology the semantic differentiation of such verbs is to be understood grammatically, as a difference between immanent and extroverted voice, one of the most salient and peculiar feature of active languages. (Thus pairs

<sup>37</sup> It is important to make it clear that we do not claim that transitivity relations cannot be *i n t e r p r e t e d* in such languages (are not extant in them). In a logico-semantic sense the verb *die* is intransitive and *kill* is transitive in any of the world's languages. The typological question is whether the difference between the sentence patterns that involve such verbs (apart from the trivial difference in the number of complements) is based on this particular difference or something else. It is only in the former case that transitivity relations can be said to be operative in the syntactic constructions of the language.

of meanings like 'set fire to' vs. 'catch fire', 'drag' vs. 'crawl', 'carry' vs. 'go' are expressed not via different lexemes but via voice, i.e. verbal inflection.) In any case, these "double" verbs in themselves demonstrate the absence of the transitivity principle: in nominative, triadic and ergative languages they correspond to different lexemes.

What follows from all this is that it is clearly a mistake to regard the active sentence pattern as a variant of the ergative pattern. The difference between the active and the ergative systems is the same as that between the nominative and the triadic systems: the absence of the transitivity principle.

Although in active language sentence construction crucially depends on the active or stative nature of the verb, nominal cases, where they exist, also form a specific category. The most obvious terms for the main cases could be active and inactive.<sup>38</sup> The active case denotes the agent, the inactive case denotes the non-agent or patient (though patient can be expressed by other means too, as in the Choctaw example above). These languages lack not only the nominative and the accusative cases (uninterpretable in this system anyhow) but also the genitive case (which is not needed because of Px-es).

It is noteworthy that active languages can now be found only on the American continent, but there they are quite well represented. Some of the bigger Indian language groups such as Na-Dene, Tupi-Guarani, Sioux, Iroquois-Caddo belong here. In sum there are more than a hundred such languages, and although many of them have very few speakers, at least one of them (Guarani) is spoken by a population of more than three million. Some linguists assume, on the basis of sporadic evidence, that such languages may yet be found in the Indonesian-Malay archipelago.

### **3.2.6 Other patterns**

The above discussion does not exhaust the list of sentence patterns in the languages of the world. Further types include the thematic (focus-oriented), the sporadically attested affective, locative, possessive types etc. For our present concerns, a survey of these is dispensable, since they do not impinge on the problem of unmarked object (or of object in general) (for more details see Havas 2003).

### **3.2.7 The synchronic system of sentence patterns**

From a purely descriptive point of view the above four sentence patterns are not equally different from each other but can be grouped on the basis of their similarities and dissimilarities. The most closely related two patterns are the nominative and the triadic pattern in that the use of the accusative case is common — and restricted — to them. Thus, these two can be collectively called *a c c u s a t i v e* types in pointed contrast to the erroneous tradition which treats this term as a synonym of nominative type. The type that is then closest to these two is the ergative type in that it is in these three — and only in these three — that sentence patterns are

<sup>38</sup> F. R. Palmer, who calls active languages agentive (see above), proposes the terms "agentive" and "patientive" for these grammatical forms (1994 : 14).

primarily determined by transitivity relations. Thus, they can be referred to as *t r a n s i t i v e* types. The sentence pattern in active languages, as we have seen, is based on a *n o n-t r a n s i t i v e* principle.

### **3.3. Schematogony**

#### **3.3.1 The question of diachronicity**

One of the most intriguing questions of historical typology, and the one most germane to the present discussion, is whether this synchronic description can be "translated" into the language of diachrony. Is it possible to say that the non-transitive sentence pattern is more primitive than the transitive, which is in turn more primitive than the accusative, which is more primitive than the nominative pattern? Apparently the same question can be asked in the following way. Is the list of sentence patterns: active, ergative, triadic and nominative, a linear series of diachronic development? In what follows I will attempt to answer these questions (with an emphatic yes for the former, and, for the latter, with yes in a certain sense, but with no in another sense).

#### **3.3.2 Previous approaches**

First let us briefly survey what attempts have been made in language typology to tackle with the issue of the diachronic relation between the various sentence patterns. As I see it, there are two discernible tendencies.

One is what can be referred to as the *g l o t t o g o n i c* assumption, the chief representative of which was stadial typology, an approach cultivated most vigorously by N. Marr's school in the 1930's and 1940's (the name itself was also popularised by him). It must be borne in mind that this approach had its roots to a certain extent in the language typology of the early nineteenth century (the "Old Grammarian" approach) and some of its elements survived into the late twentieth century as well. In this approach it is claimed that language types (in a syntactic sense) represent a line of diachronic development whose steps presuppose each other in a linear and unidirectional fashion. In other words, the language types to be found are the remnants of the various indispensable stages in the genesis and development of language (i.e. the glottogonic process). The nominative type is the endpoint of the development, immediately preceded by the ergative stage. Numerous attempts were made in the Soviet Union of the 1930's and 1940's, but also in the West both prior to and following that period, to demonstrate that the nominative pattern derives from the ergative pattern. How many and what sort of stages preceded that type is a moot question, but the initial stage is usually conceived of as a language having an amorphous grammar.

The other approach may be provisionally called sceptical. It claims either that language typology being a necessarily synchronic (or perhaps panchronic) branch of science it cannot, by its taxonomic nature, have any historical relevance, or that although the transition from one type into another is recognised, such processes are characterised neither by a determinate direction nor by any historical necessity: any type may or may not at any time

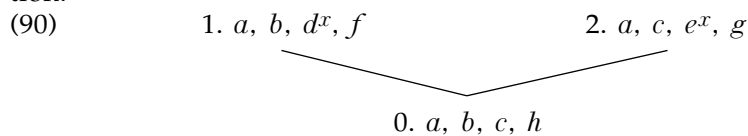


change into any other type. (This is most often claimed in connection with nominative and ergative languages, see Dixon's book cited earlier at several points.) Also, the argument goes, most languages are of a mixed type in one sense or another, so the possibility of type change is inherent in them anyhow.

### 3.3.3 The essence of schematogony

The argument I wish to elaborate here shares neither the historical-typological determinism of the former approach, nor the agnosticism of the second. It actually does not take the question that is answered in different ways by the two approaches to be a valid question of historical typology. As for the untenability of the glottogonic approach, it was amply demonstrated by the end of the 1940's. In fact all cases of a transition e.g. from ergative to nominative described in the literature involved languages of a mixed type ("split ergativity"), i.e. languages that had both ergative and nominative syntactic constructions, for instance depending on tense. A language of such a mixed type can naturally be affected by levelling in either – and truly either – direction, but this is not organic development but analogical generalisation (one of the existing variants simply replaces the others). By contrast, if a language is *p u r e l y* of one type syntactically, there definitely is no necessity, perhaps even no possibility, for it to go over into another type (a nominative language does not become ergative or *vice versa*). Thus any stadial succession of language types is fictitious. However, there is also no need for agnosticism: there are diachronic correlations, but between sentence types or sentence patterns rather than language types. Thus there is no glottogony, only "schematogony", the theory of the diachronic development of sentence patterns.

Before going into details we need to be clear about what we mean by the before vs. after (or archaic vs. innovative, ancestor vs. descendant) relation in this approach. Given any hypothetical prehistorical formation (physical object, living creature, language or anything) from which we derive two objects that exist in the present, the features of the three objects can be represented in the following (abstract and simplified) configuration:



The lower line shows the features of the earlier object, the upper line those of the two present objects. Seen from the perspective of the latter

$a$  means the common inherited features of the derivatives;

$b, c$  mean the inherited features which manifest themselves only in one of the derivatives respectively;

$d^x, e^x$  mean features that developed in an independent but parallel fashion;

$f, g$  mean independent and dissimilar features;

$h$  means demonstrable or conjectured features in the prehistoric formation which have no continuation in (these) derivatives.

Let us start from the — most likely — position in which only 1 and 2 are factually known and 0 is a hypothetical prehistoric form. Let us assume that the observable features *c* in 2, which are not found in 1, were nevertheless present in a historically earlier form of 1. Logically this means that the features under *c* in 2 are the remnants of a historically underlying form; in other words, that *c* was also present in 0. Very often, however, the conclusion drawn is 2 itself historically underlies 1.<sup>39</sup> Needless to say, such a conclusion is unwarranted.

I think this is the key to the understanding of the diachrony of sentence patterns. For instance, the fact that a historical analysis of Indo-European languages points to archaic structures of an ergative nature should not lead one to conclude that Proto- or Pre-Indo-European was an ergative language, while, of course, such observations are also not to be discarded. The hypothesis that may lead us toward a solution is that some of those features that manifest themselves in full-fledged form in ergative languages point to an archaic stage that historically underlies nominative structures in some languages and ergative structures in others. At this point again it must be borne in mind that we are not talking about languages or language types but sentence patterns. We do not make up a pristine language or language type but instead hypothesise that the very formation of the structural possibilities of sentence organisation is a diachronic process with its own historical logic and that tracing backwards the separate histories of the nominative and ergative constructions we may reach a point in the evolution of sentence patterns where the two lines of descent meet.

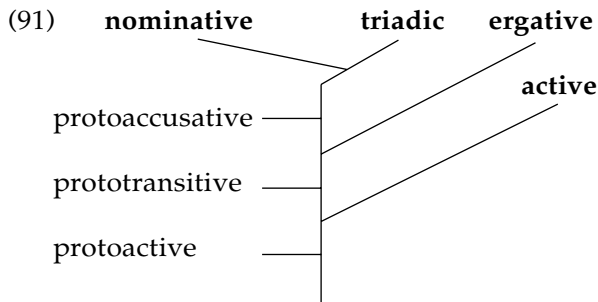
This approach then makes it possible to establish a *v i r t u a l* diachronic succession of the various sentence patterns without positing a *f a c t u a l* sequence of transitions among them. As in the case of a tree one can claim that a higher branch is younger (was formed later) than a lower branch without claiming that the higher branch sprouted from the lower, so in the case of sentence patterns one can make valid claims about their diachronic succession without also deriving one from the other in their full-fledged, self-contained forms.

#### **3.3.4. The outlines of schematogony**

The chart below (91) represents the — naturally hypothetical — diachronic (or, perhaps, it would be more accurate to say pandiachronic or "pale-ontological") aspects of the four basic sentence patterns descriptively characterised and classified above. (The remaining sentence patterns not discussed here would be placed in the continuation of the chart towards the bottom, i.e. earlier in time. We also do not discuss here the hypothetical forms relevant to the emergence of the sentence as such.)

---

<sup>39</sup> Cf. the commonly made mistake of deriving humans from apes on the basis of the hominid features of the latter. The correct hypothesis of course is that humans and apes had a common ancestor and certain features of this common ancestor are now directly observable only in apes and not in humans.



**3.3.4.1.** Let us begin our explanation of the tree at its top. The nominative and the triadic patterns are connected by the presence of the accusative and it is this feature that sets them in contrast — primarily — to the ergative pattern. What this implies for diachrony is that the emergence of the accusative (the object) is independent of that of the nominative and in all likelihood predates it. (The “common ancestor” of the nominative and the triadic patterns may have involved the accusative, but not the nominative, case.) The accusative, as was said earlier, is nothing else than the case of the patient — as opposed to the agent — that accompanies a transitive verb. It was probably in that capacity that it separated from the originally undifferentiated (i.e. of the same category when accompanying transitive and intransitive verbs) patient function (I = III). If we accept this line of reasoning it must be seen that the resulting pattern was more like the triadic than like the later nominative pattern in that the three functions all differed morphologically (this is what the broken line in the tree suggests: nominative derives from an already triadic line of development). As was said earlier, in the triadic pattern the transitivity of verbs is marked redundantly. After function III was differentiated, i.e. the category of the object emerged, there was no obstacle to the merger of functions I and II, that is, the emergence of the nominative case. (This is indeed what mostly happened, though not by necessity; where it did not happen, the triadic system remained — and obviously acquired further features.) In principle, the levelling could take one of two forms: either through the disappearance of agentive (= the subsumption of function II under the unmarked stative) or through the disappearance of the stative (= the subsumption of function I under the marked agentive).<sup>40</sup>

**3.3.4.2.** A comparison of the triadic and the nominative patterns in itself proves the important claim that object is a more archaic category than subject. There is further empirical evidence that supports this, viz. what could be called semi-accusative structures. In Russian, for instance, a sentence meaning ‘The wave carried away the boat’ can be construed in three ways:

<sup>40</sup> The morphological transfer of the agentive (i.e. the marked transitive agent) into the orbit of intransitive verbs has long been known, much discussed and interpreted in a variety of ways in the context of Indo-European. It seems clear that the suffix *-s*, already understood by the “Old Grammarians” as the marker not of the nominative but of the intentional agent (or, on other interpretations, of animacy) — cf. Greek *ti* ‘what’ vs. *tis* ‘who’, or common masculine and feminine forms for two-form adjectives — may well have been the suffix of the agentive case. First it may have spread to structures involving intransitive verbs accompanied by animate agents, later to all animate subjects (with possible further differentiation, e.g. nearly all masculines). This is how the marked nominative case came into being.

- (92) a. *Волна унесла лодку*  
 wave-NOM carried(FEM) boat-ACC  
 b. *Лодка (была) унесена волной*  
 boat-NOM (was) carry-PSSPART wave-INSTR  
 c. *Волной унесло лодку*  
 wave-INSTR carried(NEUTR) boat-ACC

The first variant has a nominative structure: 'the wave carried away the boat', the second is passive: 'the boat was carried away by the wave'; the third, however, is neither active, nor passive. It cannot be passive because the patient (the semantic object) is in the accusative case (*лодку*), but it also cannot be active because the agent (the semantic subject) is in the instrumental case (*волной*). It looks as though it was the contamination of 92a and 92b, but this, of course, is out of question, since then the sentence would be either semantically anomalous or elliptical ('carried away the boat by the wave'). One of the interpretations of such structures, as has been proposed, could be just that: *волной* is an instrumental and the subject is missing from the sentence. This mysterious subject could then be a supernatural force — unnamed perhaps for taboo reasons — which carried away the boat (or had the boat carried away) with the help of the wave as its instrument.<sup>41</sup> However imaginative these explanations may seem, they are completely untenable. First, it requires only a modicum of analytical sense to see that *волной* in 92b, an obviously passive sentence, is functionally exactly the same as *волной* in 92c, but in the former it would not occur to anyone to analyse it as an *a d v e r b o f i n s t r u m e n t* (least of all because that sentence has a subject: the boat, whose instrument the wave cannot be). Second, the predicate of the sentence is neuter — which makes it impossible to regard either of the two nouns (in this case both feminine) as its subject — so the putative "external" subject has to be neuter, which is again unlikely in view of what we know of the ancient mythological outlook on the world (a natural force would obviously be animate, thus masculine or feminine). But it is wrong in the first place to suppose that the neuter form *унесло* results from agreement. In Russian — just like in other languages, cf. Latin *pluit* 'it rains', *ningit* 'it snows' — there are zero-argument verbs (*дождит* 'it rains', *кажется* 'it seems'). With such verbs it makes no sense to claim that they have a hidden subject. In Russian the neuter is the default form of the verb in the past tense (*дождило, казалось*), which requires no explanation; just the other way round: it is for the other two genders that a triggering factor is needed. In sum, sentences of the type 92c are neither contaminated, nor elliptical.

One more note is in place here. School grammars and a widespread transformationalist tradition have made it look as a matter of course that the subject of an active sentence will be in the instrumental case in the passive counterpart of the sentence. But the question has to be asked: why

<sup>41</sup> This interpretation may have been suggested by the fact that such sentences usually describe events involving spontaneous natural forces: *дорогу занесло снегом* 'the road was covered in snow', *лес заволокло туманом* 'the forest was deep in fog', *кепку сбilo с головы* (scil. ветром) 'the wind blew my cap off my head' etc. In the case of sentences like *глаза заволокло слезами* 'my eyes got tearful' it would be difficult to conceive of pristine forces as instigators of the event.

instrumental? If one says: *книга пишется автором* 'the book is being written by the author', any other solution would be more plausible than taking the author as an instrument (and whose instrument, anyway?). The ultimate cause of the confusion lies in the name of the instrumental case, which suggests that the function of this form is *par excellence* the expression of an *i n s t r u m e n t*. The form, however, with which instruments are expressed — i.e. the specific case form of the instrumental or the case form that is used for that purpose beside others (like instrumental—comitative in Russian and Hungarian or the adessive in Finnish) — is used for a wide variety of adverbs in a host of languages (the following examples are Hungarian, Russian and Finnish). It is used for adverbs of time: Hu *tavasszal* / Ru *весной* / Fi *kevällä* 'in spring', Hu *reggel* / Ru *утром* / Fi *aamulla* 'in the morning', for adverbs of manner: Hu *joggal* / Ru *(с) правом* / Fi *oikeudella* 'with justice', in Russian also for adverbs of place: *мы шли лесом* approximately 'we walked on wooded land'. Traditionally these are referred to as instances of metaphoric use, but if one comes to think of it, what is more likely: an instrument metaphorised as circumstance (e.g. time) or the opposite, an instrument being interpreted as a certain circumstance of the event? It is clear that only the latter explanation is tenable: what is called instrumental (function) is originally an all-encompassing adverb of circumstance, which can quite naturally be used to denote place, time, manner etc., and also, of course, to denote instrument.

Thus the structure found in 92с, *волной унесло лодку*, is one in which the noun in instrumental case resembles the agentive in the triadic pattern much more than the adverb of instrument in the nominative pattern, let alone the subject in the latter. Of course this element cannot be a real agentive, since the word in instrumental expresses circumstance rather than agent (i.e. intentional, volitional actor). We interpret this syntactic form as the relic of an archaic sentence pattern whose organising principle involves the object but does not involve the subject. That is why I have suggested, as long as a more accurate term is not found, to call these semi-accusative structures. They are found in other Indo-European languages too (e.g. in Hindi) as well as in one Uralic language (Ostyak) (see Havas 2006a). The point is that this pattern is a pre-nominative syntactic pattern, the late remnant of a stage in which accusative was already there but nominative was not yet. To repeat this as seen from another angle: the nominative case — and thus the nominative pattern — is a linguistic entity that came into being historically in a relatively recent period (more recently even than the accusative).

**3.3.4.3.** Since it is not germane to the topic, the diachronic interpretation of the earlier stages of schematogony will not be pursued here. But it must be stressed again that transitivity, which obviously played a crucial role in the emergence of the category of object, among others, is itself a historically emerging category. There is no doubt that the most important development of the "post-proto-active", i.e. the proto-transitive, stage was the appearance of transitivity as a principle of syntactic construction and its rise to a prominent function. Both the (proto-)ergative and the (proto-)triadic patterns are in harmony with this principle. Of the two, the ergative pattern is the conservative, the triadic pattern the innovative (the latter can actu-

ally be thought of as an originally ergative system where the occurrences of the absolutive were differentiated next to an intransitive vs. a transitive verb).

### **3.3.5 The diachronic-typological status of unmarked object**

Let us recall the claim made above that object is an older category than subject. In other words: object is a prenominal category already in itself. At the same time, as has been pointed out, object itself has its own prehistory. In the apparent fact that object is the patient of a transitive verb — and only of a transitive verb — there are two aspects that point to the historical emergence of object: transitivity and patienthood. As schematogony demonstrates, these two aspects, in turn, stand in a historical relation to each other. Transitivity is a younger category than the agent vs. patient distinction; as we have seen, the former is unknown in active systems, which are based on the latter. What follows from this is that the ultimate prehistoric origins (prenominal on several counts) of object must be sought where patient as a category emerges.

Patient is clearly a reflexive category, i.e. it is only together with the agent, not in itself, that it can be interpreted synchronically and can emerge historically. Pristine two-member sentences may have emerged as amalgamations of one-member sentences, i.e. word-sentences — this is how analyzable sentences in the modern sense appeared, and this is how word-sentences were degraded to words. The "common ancestor" of the agent and the patient was obviously the sentence element(-to-be) with a nominal content, which became the subordinated complement of the predicate on the amalgamation of the original word-sentences, i.e. on the emergence of two-member sentences. This constituent may originally have had a highly general meaning or function: it denoted the "close semantic complement", which was immediately connected to the event described by the predicative constituent. The doubling of this undifferentiated nominal entity then came about via the separation of a marked kind of "close semantic complement" from it, namely the initiator of the event or state expressed by the predicate, i.e. the agent. An event expressed with the help of an agent encompasses a broader range of pragmatic relations than the original two-member sentences, and this was one of the first steps towards the (linguistic) recognition of causal relations. But with the differentiation of agent the original, undifferentiated variant of the "close semantic complement" also became, as non-agent, a separate semantic category, viz. that of patient.

Regardless of how the morphological structure of languages developed later, in terms of their original functional/semantic roles markedness is characteristic of the agent, and unmarkedness of the patient. This has been preserved in the active construction and its traces are still to be found in ergative languages in that the ergative case is marked in all of them, whereas the absolutive is mostly unmarked. When in the proto-accusative period the patient bifurcated into forms expressing the patient of intransitive vs. transitive verbs, i.e. it developed into original stative vs. object, the newly emerging category of object became the exponent of the transitivity of the verb on the nominal element. As said above, object included two aspects: transitivity and patienthood. Of the two, the latter is clearly

the older. The object continues to be the patient, the closest complement of a transitive verb.<sup>42</sup> Thus it is not its markedness that makes the object but its status as the closest complement of a transitive verb. This is primarily a semantic–functional relation, which may just as well be unmarked morphologically.<sup>43</sup>

Since case endings are usually the grammaticalised “depositories” of syntactic–semantic functions, these functions must obviously precede them both temporally and logically. If we add to this that in a diachronic–typological sense object is heir to patient and if we bear in mind that patient is originally unmarked by nature, it will appear as self-explanatory that although both forms of object are prenominal, the roots of unmarked object go even deeper than those of marked object (namely, to the proto-transitive stage, as opposed to the “merely” proto-accusative origins of marked object). Thus object is originally a *par excellence* unmarked category. It follows that what demands a linguistic explanation is not why the object is unmarked in certain cases but just the opposite: it is its markedness that must be historically explained. This is so even if the markedness of object has become dominant or exclusive in a large set of languages by now.

#### **4. Conclusion**

I have surveyed the occurrence of unmarked object in twenty Uralic languages. I have found that grammatically relevant unmarked object occurs in all of the Uralic languages with the exception of Ostyak (and Forest Nenets) in at least one, and in many languages several, of the seven syntactic contexts I have looked at. The conventional wisdom is that unmarked object existed in Proto-Uralic. Of the syntactic environments here surveyed I only excluded the occurrence of unmarked object next to “passive” verbs from those that may have been inherited from the proto-language. For the remaining six environments inheritance from the proto-language is at least a possible alternative, in some cases the only one. Thus the efforts to derive unmarked object from the reanalysis of some other syntactic function of a nominative form — while at the same time accepting the idea that unmarked object existed in the proto-language — are unjustified.

The synchronic coexistence of the syntactic structures found in the world’s languages is the depository of a historical relation in a diachronic–typological sense, which can be interpreted in terms of a schematogony. In these terms, the most recent syntactic structure is the nominative pattern

---

<sup>42</sup> The diachronic–semantic relation between the object and the transitive verb is thus the same close relation as that between the transitive verb and the absolutive element in an ergative language, except that here the transitive nature of the verb is highlighted.

<sup>43</sup> There are no historical clues as for whether in the emergence of the triadic pattern the object or the stative was the new motif. Following the logic of the theory of schematogony, in terms of morphology, it should have been the stative that separated from the undifferentiated category of the patient, and what was “left behind” (the patient only of transitive verbs from then on) was the object. This could be proved by demonstrating that in triadic languages the stative is more frequently marked than the object. Unfortunately the triadic languages known and the data they provide are too few to establish this.

with the typical categories subject and object. Subject typifies the nominative pattern, but object is found already in triadic languages and is thus a prenominal category in itself. More clearly prenominal is its unmarked variant, whose origins go back to the proto-transitive and the proto-active stages, where the category of object had not yet emerged. Consequently, unmarked object must be a prenominal inheritance in Uralic, one of those relics that make it possible to trace the prehistory of our languages in a diachronic-typological sense into periods that are — justifiably — excluded from the scope of classical comparative and historical research.

Address:  
Ferenc Havas  
Eötvös Loránd University, Budapest  
E-mail: fhavas@ludens.elte.hu

#### REFERENCES

- Alhoniemi, A. 1970, Über die Form des zu Verbalnomina treffenden Objekts im Tscheremissischen. — FUF XXXVIII, 1—136.
- Bereczki, G. 1990, Chrestomathia Ceremissica, Budapest.
- 2003, A magyar nyelv finnugor alapjai, Budapest.
- Collinder, B. 1960, Comparative grammar of the Uralic languages, Stockholm.
- Comrie, B. 1978, Ergativity. — Syntactic Typology. Studies in the Phenomenology of Language, Austin—London, 329—394.
- Csúcs, S. 1990, Chrestomathia Votiacica, Budapest.
- Dixon, R. M. W. 1994, Ergativity, Cambridge.
- Fejes, L. 2005, Összetett szavak finnugor nyelvekben, Budapest (Unpublished PhD dissertation, Eötvös Loránd University, Budapest).
- Grünthal, V. 1941, Itämerensuomalaisten kielten yksikön nominatiivi objektin edustajana aktiivin yhteydessä, Helsinki (SKST 218.).
- Hajdú, P. 1989, Bevezetés az uráli nyelvtudományba, Budapest.
- Havas, F. 2003, A tárgy tárgyában. Mondattipológiai fontolgatások. — Budapesti Uráli Műhely III, Budapest, 7—44.
- 2006a, Die Ergativität und die uralischen Sprachen. — FUF 59, 81—131.
- 2006b, Die Pränominativität der uralischen Partizipien. — UAJb., Neue Folge, Band 20, 16—50.
- Honti, L. 2003, Volt-e az uráli/finnugor alapnyelvben nem egyes számban accusativus? — Budapesti Uráli Műhely III, Budapest, 76—83.
- Iso suomen kielioppi, Helsinki 2004.
- Itkonen, E. 1971—1972, Über das Objekt in den finnisch-wolgaischen Sprachen. Finnisch-Ugrische Forschungen 39 : 153—213.
- Itkonen, E. 2001, Maaailman kielten erilaisuus ja samuus. Turku.
- Jahnunen, J. 2000—2001, The Nenets Imperative Sentence and Its Background. — Finnisch-Ugrische Mitteilungen 24/25, 71—86.
- Katz, H. 1980, Das Ururalische — eine Ergativsprache. — CIFU V. Pars VI, 393—400.
- Keresztes, L. 1996, Gondolatok a finn passzívumról. — Finnisch-ugrische kontrastive Untersuchungen, Jyväskylä (Hungarologische Beiträge 6), 19—30.
- Kiparsky, V. 1959—1960, Über das Nominativobjekt des Infinitivs. — Zeitschrift für slavische Philologie 28, 333—342.
- Kont, K. 1963, Käändsõnaline objekt läänemeresoome keeltes. Tallinn.
- Korompay, K. 1991, A névszóragozás. — A magyar nyelv történeti nyelvtana. I. kötet: A korai ómagyar kor és előzményei. Budapest.
- Kulonen, U.-M. 1999, Object Marking in the Ugric Languages. — Budapesti Uráli Műhely I. Ugor műhely, Budapest. 63—71.



- Lehtinen, T. 1984, *Itämerensuomen passiivin alkuperästä*, Helsinki.
- Leino, P. 1987, *Hyvää suomea*, Helsinki.
- Mallinson, G., Blake, B. J. 1981, *Language Typology. Cross-Linguistic Studies in Syntax*, Amsterdam.
- Palmer, F. R. 1994, *Grammatical Roles and Relations*, Cambridge.
- Rédei, K. 1978, *Chrestomathia Syrjaenica*, Budapest.
- Setälä, E. N. 1926, *Suomen kielen lauseoppi*, Helsinki.
- Wickman, B. 1955, *The Form of the Object in the Uralic Languages*, Uppsala—Wiesbaden.
- Zsilka, J. 1973, *A nyelvi mozgásformák dialektikája*, Budapest.
- 1981, *Dialectics of the Motion Forms in Language*, Budapest.
- Грамматика мордовских (мокшанского и эрзянского) языков. Часть I. Фонетика и морфология, Саранск 1962.
- Климов Г. А. 1977, *Типология языков активного строя*, Москва.
- Кузнецова А. И., Хелимский, Е. А., Грушкина, Е. В. 1980, *Очерки по селькупскому языку. Тазовский диалект. Том I*, Москва.
- Лыткин В. И. 1955, *Современный коми язык*, Сыктывкар.
- Майтинская К. Е. 1974, *Сравнительная морфология финно-угорских языков. — Основы финно-угорского языкознания (вопросы происхождения и развития финно-угорских языков)*, Москва, 214—382.
- Терещенко Н. М. 1973, *Синтаксис самодийских языков*, Ленинград.
- 1975, *Нганасанский язык*, Ленинград.

*ФЕРЕНЦ ХАВАШ* (Будапешт)

#### **НЕМАРКИРОВАННЫЙ ОБЪЕКТ В УРАЛЬСКИХ ЯЗЫКАХ. ДИАХРОНИЧЕСКИЙ ТИПОЛОГИЧЕСКИЙ ПОДХОД**

Грамматически релевантный немаркированный объект обнаружен во всех уральских языках, за исключением хантыйского, по меньшей мере в одном, но большей частью в нескольких из рассмотренных семи контекстов. Принято считать, что немаркированный объект существовал уже в уральском праязыке.

Автор статьи приходит к выводу, что относящийся к пассивному глаголу немаркированный объект не берет свое начало в праязыке, но во всех остальных шести синтаксических контекстах современные языковые данные действительно указывают на уральское происхождение немаркированного объекта.