

## The Obligatory Coding Principle in diachronic perspective

**Abstract.** The *Obligatory Coding Principle* accounts for the inventories of possible coding frames in languages that, according to the current terminology, can be characterized as consistently accusative or consistently ergative in their system of argument coding. In coding frame inventories fully consistent with the Obligatory Coding principle, every coding frame includes a given type of coding, either A (in *obligatory A coding languages*) or P (in *obligatory P coding languages*). However, languages with coding frame inventories violating this principle are not exceptional. This paper examines the types of evolutions that may result either in global shifts affecting the Obligatory Coding Principle, in systematic violations of the Obligatory Principle, or in the gradual spreading of non-canonical coding frames. The idea underlying this study is that, before discussing the theoretical status of this kind of generalization, it is crucial to clarify its involvement in diachronic processes.

### 1. Introduction

This paper is about the possible consequences of some diachronic processes for argument coding systems, i.e. for aspects of the organization of languages commonly dealt with in terms of morphological *accusativity* vs. *ergativity*.<sup>1</sup> The possible links between the diachronic processes affecting argument coding systems analyzed here and so-called ‘deep/syntactic ergativity’ are not discussed in this paper, although a plausible connection is mentioned in Section 4.2.<sup>2</sup>

The paper is organized as follows. After putting forward some terminological clarifications (Section 2) and introducing the Obligatory Coding Principle (Section 3), I discuss markedness reversals between the basic transitive construction and one of its variants (either passive or antipassive) leading to global shifts from obligatory A coding to obligatory P coding and vice-versa (Section 4). In Section 5, I discuss TAM grammaticalization processes that introduce systematic violations of the Obligatory Coding Principle. The following two sections are devoted to changes that may be responsible for gradual shifts: emergence of isolated exceptions to the Obligatory Coding Principle in languages that initially keep strictly to this principle, or increase in the proportion of verbs with non-canonical coding frames. Section 6 deals with the conventionalization of argument ellipsis, and Section 7 with the univerbation of light verb constructions. Section 8 summarizes the conclusions.

### 2. Some terminological clarifications

#### 2.1. Coding frames

Each individual language has an inventory of possible coding frames through which verbs express their argument structure. Formal contrasts between arguments may involve flagging, indexation, and linear order. A coding frame is considered non-canonical if it is found with a restricted set of verbs. The distinction between

---

<sup>1</sup> Some of the questions dealt with here were already discussed in a slightly different perspective in Creissels (2008).

<sup>2</sup> A detailed discussion of this question for the types of changes dealt with here in Section 4 can be found in Queixalós (2013).

canonical and non-canonical coding frames is gradient rather than categorical, and may evolve in the history of a language.

## **2.2. Transitivity**

### *2.2.1. Core transitive verbs*

In this paper, verbs encoding events involving one, two, or three essential participants are designated as *monovalent*, *bivalent*, and *trivalent*. *Transitive* and *intransitive* do not refer to the number of essential participants in the events denoted by verbs, but to the fact that their construction coincides or not with that of verbs encoding a particular type of event. The delimitation of the set of transitive verbs is language-specific and relies on formal criteria, but the sets of transitive verbs of the individual languages are universally defined as including a particular semantic class of verbs, the *core transitive verbs*, defined as bivalent verbs expressing meanings compatible with a maximum degree of semantic transitivity.

In other words, a core transitive verb is a bivalent verb that has the ability to refer to two-participant events involving two well-individuated participants, a typical agent (i.e. a human participant consciously and willingly controlling an activity oriented towards the other participant), and a typical patient (i.e. a participant undergoing a change of state or position triggered by the activity of the agent).<sup>3</sup>

I assume that, in the languages of the world, the set of the verbs recognizable as core transitive verbs according to the restrictive semantic definition posited above shows a high degree of formal homogeneity, in the sense that, in each individual language, all core transitive verbs, or almost all, assign the same coding characteristics to their agents and patients. By contrast, cross-linguistically, as discussed among others by Tsunoda (1985) and Lazard (1994) and confirmed by Hartmann et al. (2013), no other class of verbs defined in terms of semantic role assignment shows a comparable propensity to group together into the same valency class. This suggests a cognitive prominence of this semantic class of verbs, and justifies giving it a privileged status in a typology of argument coding.

### *2.2.2. The transitive type of argument coding*

The coding of agents and patients in uses of core transitive verbs involving a maximum degree of semantic transitivity is designated as *transitive type of argument coding*, abbreviated as *transitive coding*. Like the notion of core transitive verb, this notion is a comparative concept in the sense of Haspelmath (2010).

All languages extend transitive coding well beyond the limits of the set of core transitive verbs. The term *transitive verb* without further specification refers to verbs whose construction includes two terms coded like the two arguments of core

---

<sup>3</sup> *Break* is a good example of a core transitive verb. By contrast, hitting verbs are not core transitive verbs, because hittees are not typical patients (and in many languages, they are coded differently from typical patients). Similarly, eating verbs are not core transitive verbs, because eaters are not typical agents (since the primary motivation of their activity is the satisfaction of a physiological need), and this explains why many languages have two totally different translational equivalents of English *eat*, one of them transitive and the other intransitive, whereas this never occurs with core transitive verbs.

transitive verbs, whatever their semantic roles. For example, English *see* is not a core transitive verb, but the coding it assigns to its arguments identifies it as transitive. Basque *ikusi* ‘see’ is also a transitive verb, since its coding frame <ERG, Ø> is the same as that of *puskatu* ‘break’ – Ex. (1). By contrast, Akhvakh *harigurula* ‘see’ is not transitive, since its coding frame <DAT, Ø> is different from the coding frame <ERG, Ø> selected in Akhvakh by *biq’ōrula* ‘break’ – Ex. (2).

(1) Basque<sup>4</sup>

a. *Haurr-ek ispilu-a puskatu dute.*  
child-PL.ERG mirror-SG break.CPLPRS.3SG.3PL  
‘The children have broken the mirror.’

b. *Haurr-ek ispilu-a ikusi dute.*  
child-PL.ERG mirror-SG see.CPL PRS.3SG.3PL  
‘The children have seen the mirror.’

(2) Akhvakh<sup>5</sup>

a. *Mik’i-de istaka biq’wāri.*  
child-ERG glass break.CPL  
‘The child broke the glass.’

b. *Mik’i-La istaka harig’wari.*  
child-DAT glass see.CPL  
‘The child saw the glass.’

There is cross-linguistic variation in the size of the set of bivalent verbs whose arguments are treated differently from the agent and patient of core transitive verbs – see for example Creissels and Bassène (2013), but transitive coding is universally the default type of coding for bivalent verbs.

In this paper, A and P refer to arguments that, in a given language, have the same coding characteristics as agents and patients of core transitive verbs, irrespective of their semantic roles.<sup>6</sup>

---

<sup>4</sup> The Basque examples quoted in this paper have been checked by Céline Mounole.

<sup>5</sup> Unless otherwise stated, the Akhvakh examples quoted in this paper come from the author’s field notes and have been checked with the help of Indira Abdulaeva.

<sup>6</sup> Not all authors use the terms of agent and patient and the corresponding abbreviations consistently, as rightly observed by Alice Harris in her review of Dixon’s *Ergativity* – Harris (1997). In this paper, the terms agent and patient without further specification, and the abbreviations A and P, consistently refer to arguments that, irrespective of their semantic role, are coded exactly like typical agents and patients of core transitive verbs in constructions compatible with the maximum degree of semantic transitivity – and not for example to the most agent-like and most patient-like participants in the argument structure of bivalent verbs, a notion which is crucial in other frameworks but plays no direct role here.

### 2.2.3. *Variation in the construction of transitive verbs and basic transitive coding*

In many languages, variation can be observed in the coding of the arguments of core transitive verbs, and this variation may lend itself to various types of analysis.

It may happen that the variation in the coding of the arguments is conditioned by the TAM or polarity value of the clause, commonly (but not necessarily) expressed through verb morphology. This phenomenon, to which I will refer as *conditioned transitive coding*, includes situations commonly described in terms of ‘split-ergativity’.

Another well-known phenomenon is the differential coding of agents or patients.

It may also happen that the variation in the coding of the arguments of core transitive verbs is best analyzed as bound to an alternation between a construction that qualifies as basic transitive coding and one or more constructions involving detransitivization. This is particularly obvious in the case of constructions (irrespective of whether they involve morphological coding on the verb or not) that semantically imply that the agent is removed from the event structure: anticausative constructions, P-oriented resultatives. There are also less obvious cases in which the event structure is not affected, but the status of one of the alternative constructions as the basic transitive construction can nevertheless be established on the basis of the following two criteria: the basic transitive construction is less marked than the other(s) in terms of discursive or semantic conditioning (and consequently much more frequent in texts), and the morphosyntactic properties of the alternative construction(s) of core transitive verbs provides evidence of demotion of either the agent (passivization) or the patient (antipassivization).<sup>7</sup>

There are, however, problematic situations in which no obvious candidate for the status of basic transitive construction emerges. I will refer to them as *multiple transitive coding*. The case of the languages with the Philippine-type of voice system has been widely discussed in the literature. A similar problem arises with other languages that have alternative constructions of transitive verbs expressing alternative perspectivizations of the event comparable to those expressed by passive or antipassive derivations, without however clear evidence that one of the alternative constructions could be considered basic, and the other(s) detransitivized variants. This includes in particular the inverse systems with a direct/inverse alternation for interactions between 3rd persons – Haude and Zúñiga (to appear).

## 2.3. *Alignment*

The usual definition of ‘ergative alignment’ and ‘accusative alignment’ refers to properties shared by S (sole argument of monovalent verbs) and one of the core terms of the basic transitive construction. It implies a more general notion of alignment that can be formulated as follows:

---

<sup>7</sup> Note that languages with both a passive and an antipassive variant of the basic transitive construction are not uncommon, contrary to the widespread opinion (originating in early work on ergativity) according to which passives are reserved to obligatory A coding languages, and antipassives to obligatory P coding languages. Janic (2013) provides a survey of antipassive constructions in obligatory A coding languages.

A TERM  $T_1$  OF A CONSTRUCTION  $C_1$  AND A TERM  $T_2$  OF A CONSTRUCTION  $C_2$  ARE ALIGNED WITH RESPECT TO A GIVEN FEATURE IF THEY SHARE THE SAME VALUE OF THE FEATURE IN QUESTION.

This general definition of alignment is retained here, to the exclusion of any other possible interpretation of this term. This must be emphasized, since some uses of ‘alignment’ in the typological literature are not consistent with this definition. In particular, in the term of ‘hierarchical alignment’ as introduced by Nichols (1992), ‘alignment’ does not refer to properties shared by terms belonging to different constructions, but to the mapping from the semantic roles of agent and patient to morphosyntactic slots. The misnamed ‘hierarchical alignment’ is rather a *type of transitive coding* in which the coding characteristics of A and P are determined by their relative ranking with respect to the indexability hierarchy. It is true that this type of transitive coding raises specific problems for alignment typology, since from a strictly logical point of view, it is difficult to compare the coding of S to that of A or P in languages in which it is impossible to define types of coding assigned to A and P independently from each other. It should, however, be clear that considering this situation as a particular type of ‘alignment’ makes no sense.

#### **2.4. Zero case**

In languages in which nouns are inflected for case, I designate as *zero case* (represented as  $\emptyset$  in the schematization of coding frames) the case form of nouns that coincides with the form used in isolation for quotation and labeling, whatever the distribution of this form in syntactic contexts. The notion of labeling includes for example the case of nouns accompanying a picture representing a possible referent, nouns written on signal boards, nouns describing the content of a box on which they are written, etc.

In most languages, the zero case is characterized by the absence of an overt case marker, but there are exceptions, and the absence of an overt marker is not essential in the notion of zero case. What is essential is the ability to be used, not only as the quotation form of nouns in elicitation contexts, but also as a pure label in the absence of any syntactic context.<sup>8</sup>

The term zero case can be understood as a cover term for the case forms currently labeled nominative or absolutive. There are two main reasons for preferring it.<sup>9</sup> On the one hand, the distinction between nominative and absolutive is not really useful, since a nominative case in an unproblematic ‘accusative’ language is simply a zero case contrasting with an accusative case, and an absolutive case in an unproblematic ‘ergative’ language is simply a zero case contrasting with an ergative case. On the other hand (and this is crucial), the usual definition of nominative and absolutive can

---

<sup>8</sup> For example, the Zero case of Latin (traditionally called Nominative) has a zero ending with some nouns (*puer* ‘child’), but an overt ending with some others (*domin-us* ‘master’). In Russian, nouns like *devuška* ‘girl’ have an ending *-a* in the Zero case in the singular, but a zero ending in the Genitive plural. Similarly, in Icelandic, *hatt-ur* ‘hat’ has an overt ending *-ur* in the Zero case in the singular, but a zero ending in the Accusative singular (*hatt*). With such nouns, flagging may involve deletion of morphological material present in the syntactically unmarked form of nouns.

<sup>9</sup> See Creissels (2009) for a more detailed discussion.

only lead to inconsistencies in the description of languages with less common patterns of alignment and/or case marking, for example, in ‘split-ergative’ languages like Georgian or Kurmanji Kurdish, in which the same morphological form of nouns meets the definition of ‘absolutive’ or ‘nominative’ depending on the tense value expressed by the verb heading the clause – for an illustration, see Ex. (5) in Section 5.1.

## **2.5. Ergative case, ergative alignment, ergative languages**

I basically subscribe to the idea vigorously expressed by DeLancey (2004) that the use of the term ‘ergativity’ has evolved so as to encompass a heterogeneous set of phenomena whose interrelations are much less simple and direct than commonly assumed, and are sometimes even inexistent. I also believe that no further progress in our understanding of the phenomena for the analysis of which this term is commonly used can be expected in so far as this question is not clarified. The solution I adopt in this paper is restricting the use of ‘ergative’ and ‘accusative’ to case terminology, and coining transparent and non-ambiguous terms for the other meanings with which these terms are used.

As regards case terminology, my proposal is to generalize the use of accusative / ergative case according to the following definition: if a form of nouns different from the quotation / labeling form is used to encode P but not A, it is labeled *accusative case*, and if a form of nouns different from the quotation / labeling form is used to encode A but not P, it is labeled *ergative case*.<sup>10</sup> Note that this definition allows using the label ‘ergative case’, not only for languages in which a marked case form is assigned exclusively by transitive verbs to their agent, but also in the following two types of situation:<sup>11</sup>

- *extended ergative case marking*, in which a substantial class of verbs whose coding frame includes no P term assign to one of their arguments the same marked case form as that assigned by transitive verbs to their A argument;
- *generalized ergative case marking* (commonly designated as the ‘marked-nominative’ type of case-marking), in which all coding frames must include a term in the same marked case form as the A argument of transitive verbs.

By contrast, I will avoid using ‘accusative’ and ‘ergative’ as labels for types of alignment. The main reason is that the extension of ergative marking to arguments of intransitive verbs results in situations in which ergative case marking is assigned by a

---

<sup>10</sup> *Oblique case* is a possible label for marked case forms involved in the coding of both A and P, depending on factors such as verb inflection. Such case forms can be found in some ‘split-ergative’ languages, for example Kurmanji Kurdish – see Section 5.1, Ex. (???)

<sup>11</sup> The terms of extended / generalized ergative case marking as I use them here must be understood, in a strictly synchronic perspective, as describing the syntactic distribution of ergative case marking. They do not imply that this distribution results from a historical process of extension of an ergative case whose use was initially limited to the A argument of transitive verbs. There is clear evidence that such a historical process was responsible for the extended ergative case marking found in Basque (as will be discussed further in this paper), or in Kartvelian languages, but other scenarios can be imagined, and for the languages of East Africa whose argument coding systems involve generalized ergative case marking, the extension of the use of an ergative case initially limited to the A argument of transitive verbs is certainly not the most plausible historical explanation.

class of intransitive verbs commonly characterized as ‘unergatives’, and that the terminology currently used by typologists leads to characterize as ‘aligned accusatively’. Such a terminological mess can only result in misunderstandings and analytical errors.<sup>12</sup> This is the reason why I propose the unambiguous terms of *A-alignment* and *P-alignment* for the types of alignment currently designated in the typological literature as accusative alignment and ergative alignment, respectively.

For similar reasons, I reject the use of accusative and ergative as possible labels for a global characterization of argument coding systems. According to the definitions found in the recent literature, a morphologically ergative language should be identified as such with reference to its alignment properties only. However, in current practice, it is clear that for many linguists, the notion of ergative language refers to a bundle of features that tend to co-occur cross-linguistically, but are nevertheless logically independent, and are dissociated in some languages, which leads to inconsistencies in the characterization of languages in which such a dissociation occurs:<sup>13</sup>

- (a) FLAGGED AGENTS, i.e. the coding of the agents of core transitive verbs by means of either an adposition or a case form (commonly termed *ergative case*) distinct from the zero case used in isolation for quotation or labeling;
- (b) UNFLAGGED PATIENTS;
- (c) EITHER NO INDEXATION AT ALL, OR INDEXATION OF PATIENTS ONLY;
- (d) OBLIGATORY P CODING, i.e. the selection of P coding as the default type of argument coding that must be included in the coding frame of all verbs (and is consequently the only possible coding of sole arguments of monovalent verbs).

The solution I propose is to replace ergative and accusative as labels characterizing systems of argument coding by the unambiguous terms of *A-unmarked* vs. *P-unmarked* systems of argument coding, conceived as referring to prototypes conflating features that tend to co-occur cross-linguistically but can nevertheless be dissociated in individual languages.

### 3. The Obligatory Coding Principle

Morphological accusativity / ergativity as it is usually defined can be viewed as a particular case of a more general principle, the Obligatory Coding Principle, accounting for a cross-linguistically common type of limitation on possible coding frames. In coding frame inventories fully consistent with this principle, every coding frame includes a given type of coding. Given the definition of A and P, this leaves two logical possibilities: in *obligatory A coding languages*, every coding frame includes a

---

<sup>12</sup> For example, one cannot be satisfied with a terminology according to which the only possible characterization of the situation of a language like Basque is that the increase in the proportion of intransitive verbs assigning ergative marking triggers the gradual disappearance of ergative alignment.

<sup>13</sup> For example, in recent publications on Basque, this language is often characterized as an ‘ergative language of the active type’. This formulation is nothing else than a pure and simple *contradictio in terminis*, if ergative and active are taken with their current definitions. It can only be consistent with an understanding of ‘ergative language’ according to which the overt flagging of agents is more important than the alignment properties of intransitive verbs.

term with coding properties identical to those of A in transitive coding, whereas in *obligatory P coding languages*, every coding frame includes a term with coding properties identical to those of P in transitive coding.

However, many languages have inventories of possible coding frames hardly compatible with the Obligatory Coding Principle. For example, Basque has two subsets of monovalent verbs, some of them assigning P coding to their sole argument – Ex. (3c), and the others assigning A coding – Ex. (3b).

### (3) Basque

a. *Haurr-ak ur-a ekarri du.*  
child-SG.ERG water-SG bring.CPL PRS.3SG.3SG  
'The child brought the water.'

b. *Ur-ak irakin du.*  
water-SG.ERG boil.CPL PRS.3SG.3SG<sup>14</sup>  
'The water boiled.'

c. *Haurr-a etorri da.*  
child-SG come.CPL PRS.3SG  
'The child came.'

A formal elaboration of the Obligatory Coding Principle is found in the generative literature under the name of Obligatory Case Parameter (Bobaljik 1993, Laka 1993, 2000, Rezac 2008a, 2008b). A question that has been particularly discussed, mainly with reference to Basque, is how to deal with the violations of the Obligatory Case Parameter in a formal syntactic framework. I will not discuss this issue further, since this paper is not devoted to an elaboration of the formal aspects of the question, but to an examination of diachronic processes likely to affect the status of a language with respect to the Obligatory Coding Principle.

Another important issue that I will not try to discuss here is the sense the Obligatory Coding Principle may have for argument coding systems in which the indexation of A and P depends on the choice of the co-argument (and in which it is consequently not possible to fully identify the coding of arguments of non-transitive verbs to that of A or P) and for systems with multiple transitive coding.

## 4. Markedness reversals between the transitive construction and its variants

### 4.1. *Introductory remarks*

As already commented in Section 2.2.4, the basic transitive construction may coexist with one or more alternative constructions of transitive verbs implying no change in the event structure referred to. In the simple cases, in the alternative constructions, one of arguments is straightforwardly encoded like the sole argument of monovalent verbs, whereas the other is either absent or encoded as an oblique. The construction

---

<sup>14</sup> See Footnote ???.

is identified as passive if the participant encoded like the sole argument of a monovalent verb corresponds to the P term of the transitive construction, as antipassive if it corresponds to the A term of the transitive construction.

In this section, I discuss possible evolutions by which a construction that was initially a marked variant of the basic transitive construction tends to become less marked and more frequent, the outcome of such an evolution being the decay of the construction that was initially the basic transitive construction in the language in question, and its replacement by a construction whose initial status was that of a derived intransitive construction of the passive or antipassive type.

For a proper understanding of the questions discussed in this section and in the following one, the distinction between passive constructions (which refer to an event structure identical to that encoded by the basic transitive construction) and anticausative or resultative constructions (which refer to an event structure including no agent) is crucial.<sup>15</sup> The difficulty is that the distinction is not always easy to draw, since diachronically, resultatives (and anticausatives) are a common source of passives, and many languages have constructions that are synchronically ambiguous between resultative and passive (or anticausative and passive) readings. This explains why much of the discussion on alignment changes has been flawed by a widespread confusion between the notions of passive and resultative which has its roots in the traditional grammar of European languages.

#### ***4.2. Shift from obligatory A coding to obligatory P coding resulting from the reanalysis of a passive construction as the basic transitive construction***

It has long been observed that obligatory P coding (with in particular the S argument of monovalent verbs encoded like the P argument of transitive verbs) is typically found in languages in which A is flagged and P unflagged. In other words, obligatory P coding is typically found in languages in which the basic transitive construction resembles the pattern found in the passive variant of the transitive construction in languages that have obligatory A coding and unflagged A's.

It is therefore tempting to imagine that, in obligatory P coding languages, the transitive construction might be the reflex of a former passive variant of the transitive construction reanalyzed as the basic transitive construction. This reanalysis can be conceived as the result of an evolution by which the former passive construction gradually loses its marked character and becomes more and more frequent, so far as to eliminate the former transitive construction, or to relegate it to the level of a mere variant whose use is bound to more or less restrictive conditions.

This is undoubtedly a plausible scenario, since among obligatory A coding languages that have a passive variant of the transitive construction, there are important differences in the frequency of passive constructions in texts, and there may even be conditions in which the basic transitive construction cannot be used, and the passive construction is obligatory. Queixalós (2013) provides a well-informed discussion, with some new elements, of the factors that may motivate the systematization of agent backgrounding, resulting in the obsolescence of the active

---

<sup>15</sup> On resultative constructions, see Nedjalkov and Jaxontov (eds.) (1988).

construction and the reanalysis of the former passive construction as the basic transitive construction.

The problem is, however, that no absolutely uncontroversial case of transitive constructions resulting from the reanalysis of a passive construction has been proposed so far. This possibility has been repeatedly evoked for many languages, but once the distinction between passive and resultative is duly acknowledged, one must conclude that, for most of the languages in question, no convincing proof of the passive origin of the transitive construction has ever been presented.

For example, Gildea (1997) analyzes six Cariban languages in which a participle with stative-resultative semantics has variously evolved to give an inverse voice, “some sort of pragmatically-marked active ergative construction”, and a split-ergative pattern with P-alignment in the past tense. He argues that, in all cases, the participle has evolved “through an eventive passive stage”, but at the same time he clearly recognizes that this eventive passive stage is “unattested in any synchronic Cariban language”, and that two steps in the evolution he postulates, agentless passive and agentive passive, “must be inferred from the further evolution of the construction”. In other words, the only reason why he posits the development of a passive in the evolution leading from the Proto-Cariban participle to the split-ergative pattern of Tiriyo and Wayana is the common (but erroneous) belief that an intermediate passive stage is obligatory in the process converting resultative forms of transitive verbs into plain transitive forms P-aligned with the corresponding intransitive forms.

In fact, relatively convincing (although indirect) evidence that P-alignment results from the reanalysis of a passive construction as the basic transitive construction can only be found in ‘deep-ergative’ languages, since the most plausible explanation of the pervasiveness of P-alignment in the syntax of the languages in question is that, diachronically, A’s are former adjuncts that have been reanalyzed relatively recently, and still maintain behavioral properties typical for adjuncts.

In this connection, I would like to emphasize that the reanalysis of a true passive construction as the basic transitive construction is a plausible scenario for a global shift from obligatory A coding to obligatory P coding, but not for the emergence of ‘split-ergative’ systems of argument coding. A true passive construction involves no modification of the event structure referred to, displays the same TAM paradigm as the basic transitive construction, and in languages in which passive constructions are obligatory in certain conditions, TAM is not relevant to the obligatoriness/optionality of passives. There is consequently no reason why a markedness reversal converting a passive construction into the basic transitive construction could be conditioned by TAM. ‘Split-ergative’ systems with a TAM-driven alternation between A-alignment and P-alignment are better explained as resulting from the type of evolution analyzed in Section 5, i.e. the grammaticalization of new TAM forms via reanalysis of various types of constructions in which, for some reason, the coding of agents and/or patients was different from their coding in the basic transitive construction.

#### ***4.3. Shift from obligatory P coding to obligatory A coding resulting from the reanalysis of an antipassive construction as the basic transitive construction***

Antipassive constructions in obligatory P coding languages typically involve unflagged agents and flagged patients (since the agent in an antipassive construction

is encoded like the sole argument of a monovalent verb, and the patient like an oblique), and consequently resemble the pattern typically found in the basic transitive construction of obligatory A coding languages. Consequently, it is not unreasonable to think that the basic transitive construction of at least some obligatory A coding languages might result from the reanalysis of an antipassive construction in an obligatory P coding system as the basic transitive construction.

Interestingly, contrary to the reanalysis of a passive construction as the basic transitive construction discussed in Section 4.2, this is not only a speculation supported by more or less convincing indirect evidence. The markedness reversal leading to the reanalysis of a former antipassive construction as the basic transitive construction is indeed documented in the Inuktitut dialect of Inuit.

In the Eskimo languages (Yupik and Inuit), core transitive verbs have three possible constructions. In the construction considered basic, the agent is in the Ergative case (a syncretic case form also used in genitive function, traditionally called ‘Relative case’), the patient is in the Zero case, and the verb agrees with both the agent and the patient – Ex. (4a). In the passive construction, the patient is in the Zero case, in the same way as in the basic construction, and the agent is in a marked case form (distinct from the Ergative), but the verb agrees with the patient only – Ex. (4b). In the antipassive construction, the patient is in a marked case form, the agent is in the Zero case, and the verb agrees with the agent only – Ex. (4c). The passive and antipassive alternations are morphologically coded on some verbs only.

(4) Baffin Island Inuktitut (Spreng, 2005: 2-3)

a. *Anguti-up arnaq kunik-taa.*

man-ERG woman kiss-3SG.3SG

‘The man kissed the woman.’

b. *Arnaq kunik-tau-juq anguti-mut*

woman kiss-PASS-3SG man-ABL.SG

‘The woman was kissed by the man.’

c. *Anguti kunik-si-vuq arna-mik.*

man kiss-ANTIP-3SG woman-MOD.SG

‘The man is kissing a woman.’

However, it has been observed that some varieties of the Inuktitut dialect of Inuit (the dialect spoken in the North Eastern part of Canada) tend to reanalyze the former antipassive variant of the transitive construction as the basic transitive construction: the conditions that limit its use in Yupik and in more conservative Inuit varieties are not active anymore, whereas severe restrictions have been introduced in the use of the former basic transitive construction. The former basic transitive construction, illustrated in (4a) above, tends to be used only with agents that are not represented by noun phrases and are expressed through indexation only, which may lead to the disappearance of Ergative flagging, and in some Inuktitut varieties, the former antipassive construction has become much more frequent than the other two variants of the transitive construction. For example, an Itivimiut narrative text analyzed by

Carrier (2012: 75-76) includes only 12 occurrences of the former basic transitive construction, all with agents expressed through indexation only, against 18 occurrences of the passive construction and 117 occurrences of the construction traditionally designated as antipassive.

## 5. Grammaticalization of TAM forms and the Obligatory Coding Principle

### 5.1. Introductory remarks

The changes examined in this section, like those examined in Section 4, are global changes that affect at the same time the construction of all transitive verbs and change the status of the argument coding system with respect to the Obligatory Coding Principle. The difference is that the types of changes examined in Section 4 convert obligatory A coding systems into obligatory P coding systems and vice-versa, whereas those examined in this section explain the emergence of TAM-driven split-alignment systems in which an alternation in the coding properties of A and P has no equivalent in coding frames other than the basic transitive construction. In such systems, the types of argument coding available for intransitive verbs cannot coincide globally with the coding of either A or P, hence a systematic violation of the Obligatory Coding Principle.

For example, in Kurmanji Kurdish – Ex. (5), the sole argument S of monovalent verbs such as *hatin* ‘come’ invariably shows coding properties typical of core arguments (no flagging and obligatory indexation on the verb), whereas the coding of A and P is characterized by an alternation between a coding identical to that of S and an oblique-like coding.

#### (5) Kurmanji (Blau and Barak 1999)

- |   |  |
|---|--|
| a. <i>Ez Sînem-ê dibîn-im.</i><br>1SG Sinem-OBL see.INCPL-1SG<br>‘I see Sinem.’ | e. <i>Min Sînem dît-Ø.</i><br>1SG.OBL Sinem see.CPL-3SG<br>‘I saw Sinem.’    |
| b. <i>Sînem min dibîn-e.</i><br>Sinem 1SG.OBL see.INCPL-3SG<br>‘Sinem sees me.’ | f. <i>Sînem-ê ez dît-im.</i><br>Sinem-OBL 1SG see.CPL-1SG<br>‘Sinem saw me.’ |
| c. <i>Ez tê-m.</i><br>1SG come.INCPL-1SG<br>‘I am coming.’                      | g. <i>Ez hat-im.</i><br>1SG come.CPL-1SG<br>‘I came.’                        |
| d. <i>Sînem tê-Ø.</i><br>Sinem come.INCPL-3SG<br>‘Sinem is coming.’             | h. <i>Sînem hat-Ø.</i><br>Sinem come.CPL-3SG<br>‘Sinem came.’                |

Interestingly, many languages in which the grammaticalization of a new TAM form resulted in such a situation have undergone a subsequent evolution that can be characterized as regularization under the pressure of analogy, and this regularization

may occur in two different ways: either the particular coding of agents and patients found in (a group of) tense(s) as the result of changes in TAM inflexion aligns with the coding found in the other tenses, or the alternation in transitive coding is extended to intransitive coding. The first variant of the regularization process occurred in many Iranian languages which had at some point in their history an argument coding system of the type illustrated in Ex. (5), but subsequently aligned the coding of agents and patients in all tenses with the type of coding found in the present. The second variant occurred in the history of the North Russian Perfect: after a transitive perfect construction with adessive marking of the agent and no indexation of either A or P was created – Ex. (6a), North Russian has extended the adessive marking to the subject of intransitive constructions in the Perfect – Ex. (6b).

(6) North Russian (Seržant 2012: 371-372)

a. *U menja ruka poraneno.*  
at 1SG.GEN hand.SG injure.PRF  
'I have injured my hand.'

b. *U cvetov sovsem zasoxnuto.*  
at flower.PL.GEN totally dry\_up.PRF  
'The flowers are totally dried up.'

This change re-established A-alignment through the whole TAM paradigm, with just an alternation between the type of argument coding commonly associated with A-alignment and a less common variety of A-aligned argument coding, in which the zero case is reserved for P, and the same marked case form is used for A and for the core argument of intransitive verbs.

## **5.2. *Reanalysis of a resultative construction as a perfect and split-alignment***

Among the languages with TAM-driven alternations between A-alignment and P-alignment, the configuration illustrated above by Kurmanji Kurdish, with P-alignment in past tense or perfective aspect, is particularly widespread, and at least in many cases, there is evidence that it arose with the emergence of a perfect that may subsequently have evolved toward a perfective aspect or past tense.

Perfects with a coding of A and P distinct from that found with other TAM forms have long been considered as having a 'passive' origin, but once the notion of passive is restricted to alternative constructions of transitive verbs with the same event structure and TAM semantics as their active counterpart, the passive theory is difficult to maintain, since it could only be accepted on the basis of evidence that a former passive construction ceased to be used in all tenses except the perfect, whereas in the perfect, the active construction was eliminated in favor of its passive counterpart. If such evidence is lacking, the passive scenario must be abandoned as purely speculative and needlessly complicated, since the reanalysis of P-oriented resultative constructions as transitive perfects is a very plausible evolution, widely attested across languages, and there is no difficulty in explaining why, in languages

with obligatory A coding, perfect tenses created by such an evolution may select a coding frame that violate the principle of obligatory A coding.

Benveniste (1952) argued that the evolution responsible for the emergence of perfects assigning oblique-like coding to A and S-like coding to P in Indo-Iranian languages was not the reanalysis of passive constructions, as had been traditionally assumed, but the creation of a perfect tense according to a scenario basically identical to the formation of Romance or Germanic *have*-perfects.

The first stage in the evolution that gave rise to *have*-perfects in Romance and Germanic languages is the development of a possessive-resultative periphrasis, i.e. a complex construction in which a resultative clause is embedded in a possessive clause. Originally, the term coded like the possessor in plain possessive clauses is interpreted in this periphrasis as a person concerned by the result of an event, as was the case in Late Latin when the periphrasis that subsequently became the Romance Perfect started developing – Ex. (7).

(7) Late Latin

a. *Littera scripta est.*

letter written be.PRES.3SG  
'The letter is written.'

b. *Habeo pecuniam.*

have.PRES.1SG money.ACC  
'I have money.'

c. *Habeo [litteram scriptam].*

have.PRES.3SG letter.ACC written.ACC

lit. 'I have (that) a letter (is) written.', with the meaning 'I am concerned by the fact that a letter is written.'

Later, the NP encoded like a possessor is reinterpreted as representing the A argument of the transitive verb and the possessive-resultative periphrasis becomes the expression of perfect with transitive verbs.

Starting from that, it is tempting to think that, in languages with an oblique-like coding of possessors in predicative possession, the same scenario may lead to a split alignment pattern with P-alignment in the perfect, since in such languages, a possessive-resultative periphrasis would assign oblique-like coding to the possessor subsequently reanalyzed as the agent of a transitive perfect. This was precisely the explanation put forward by Benveniste for the development of Indo-Iranian perfects.

After the publication of Benveniste's paper, some authors like Cardona (1970) argued the case for the traditional theory of the passive origin of Indo-Iranian perfects. Crucially, the agent in the construction of the Old Indic Perfect was in the Instrumental case (and not in the Genitive or the Dative), which casts a serious doubt on Benveniste's theory, according to which the agent should be marked by a case typically used for possessors in predicative possession. However, this observation does not constitute a proof in favor of the passive scenario, and more recently, on the basis of a careful examination of Old Indic data, Peterson (1998) and Bynon (2005)

have concluded that the traditional explanation must be rejected, without however accepting all the details of Benveniste's theory.

The point is that the reanalysis of a possessive-resultative periphrasis is not the only alternative to the traditional passive scenario. In languages other than the few European languages that have uncontroversial *have*-perfects, it is much more plausible that the crucial stage in the development of transitive perfects from P-oriented resultatives is not the embedding of a resultative clause in a possessive clause, but simply the emergence and routinization of a construction in which a person concerned by the result is encoded as an adjunct added to the resultative clause, as in Ex. (8b). This construction cannot be described as a possessive clause with an embedded resultative clause, since it has nothing in common with the expression of predicative possession in German, but semantically, the adjunct encoding a person concerned by the resultant state lends itself to the same reanalysis as the possessor in a possessive-resultative periphrasis.

(8) German (Bynon 2005:46)

a. *Die Kartoffeln sind angebrannt.*

DEF potato.PL be.PRES.3PL PREV.burn.PP

'The potatoes are / have got burnt.'

b. *Mir sind die Kartoffeln angebrannt.*

1SG.DAT be.PRES.3PL DEF potato.PL PREV.burn.PP

'I have been and gone and burned the potatoes.'

(lit. 'To me the potatoes are burnt.')

Similarly, in his analysis of the history of the North Russian Perfect and other constructions resulting from the evolution of P-oriented resultatives in various Slavic, Baltic and Uralic languages spoken in the same area, Seržant (2012) argues that there is no need to postulate either a passive construction or a possessive-resultative periphrasis as an intermediate stage in the evolution by which North Russian acquired a Perfect construction with a non-canonical argument coding.<sup>16</sup> On the basis of a careful examination of the available historical data, he convincingly shows that, in spite of the possessor-like coding of the agent in the North Russian Perfect, the scenario that best explains the whole of the data is not the development of a possessive-resultative periphrasis, but rather the addition of an adjunct initially referring to a person concerned in some way or other by the resultant situation, subsequently reanalyzed as referring specifically to the agent.

To summarize, in languages with obligatory A coding, a non-canonical coding of agents and patients leading to a violation of the principle of obligatory A coding in past tense or perfective aspect may develop as an automatic consequence of the evolution of P-oriented resultatives, if a construction of the type illustrated in (8b) is reanalyzed as a transitive construction with a new tense form expressing a perfect

---

<sup>16</sup> As already mentioned above, in North Russian, the violation of the Obligatory Coding Principle that could have resulted from this evolution has been eliminated by the extension of the non-canonical coding of the agent in the Perfect to the subject of intransitive verbs in the same tense.

value. If the same resultative construction is available with monovalent verbs (as in English *The man is gone* / *The mirror is broken*), and if the resultative construction of monovalent verbs undergoes the same reanalysis as a perfect without any change in its form, the construction of the perfect resulting from this reanalysis will automatically show P-alignment.

#### **5.4. Progressive periphrases and split alignment**

##### *5.4.1. Introductory remarks*

Progressive aspect is often expressed by complex constructions in which the phrase headed by the auxiliated verb in some non-finite or derived form is treated as a non-verbal predicate, as in English *Mary is [buying gifts for the children]* (to compare with *Mary is [in the garden]*), or Spanish *María está [comprando regalos para los niños]* (to compare with *María está [en el jardín]*). The tendency of such periphrases to evolve towards a more general meaning of present, as attested by the ongoing evolution of the progressive periphrasis of English, is a well-known phenomenon.

The motivations of this type of periphrasis and its further evolutions have been largely discussed. What I would like to draw attention to is that, in obligatory P coding languages, if no readjustment occurs, the development of such periphrases may lead to a split alignment pattern with A-alignment in the present tense.

Non-verbal predications generally involve an argument encoded like the S argument of monovalent verbs. Consequently, in languages in which A-alignment is canonical, the A argument of a transitive verb treated as S in a progressive periphrasis shows the same coding characteristics as in non-periphrastic constructions, and the grammaticalization of such a periphrasis cannot induce a change in alignment. By contrast, in languages in which P-alignment is canonical, the A argument of transitive verbs receives a different treatment in the progressive periphrasis, since it is treated as the S term of an intransitive predication.

##### *5.4.2. The ari izan construction of Basque*

Basque has a progressive periphrasis in which the intransitive compound verb *ari izan* ‘be engaged in’ combines with nominal complements marked typically locative – Ex. (9a), or with clausal complements headed by a so-called ‘Incompletive Participle’, used also to form the non-periphrastic Present of the verbs that do not have synthetic finite forms – Ex. (9b) and (9d). The construction with a clausal complement is a raising construction in which S in the construction of *ari izan* represents the S/A argument of the auxiliated verb. With transitive verbs, this results in coding characteristics different from those of the same argument in non-periphrastic constructions – Ex. (9d-e).

#### (9) Basque

a. *Jon lanean ari da.*

Jon work.SG.LOC engaged be.PRS.3SG

lit. ‘Jon is engaged in work.’ → ‘Jon is working.’

b *Jon paseatzen da.*

Jon walk.INCPL PRS.3SG<sup>17</sup>

‘Jon is walking.’ (non-periphrastic Present)

c. *Jon [[paseatzen] ari ]] da.*

Jon walk.INCPL engaged be.PRS.3SG

lit. ‘Jon is engaged in walking.’ (progressive periphrasis)

d. *Jonek berriak ikusten ditu.*

Jon.ERG news.PL see.INCPL PRS.3SG.3PL

‘Jon is watching the news.’ (non-periphrastic Present)

e. *Jon [[berriak ikusten] ari ] da.*

Jon news.PL see.INCPL engaged be.PRS.3SG

lit. ‘Jon is engaged in watching the news.’ (progressive periphrasis)

It might be tempting to conclude from this that Basque has a split alignment pattern with a Progressive tense triggering A-alignment, but this would not be correct, since in the speech of most Basque speakers there is so far no evidence that the *ari izan* construction has been reanalyzed as a single clause – Hualde and Ortiz de Urbina (2003: 284). But if this periphrasis were reanalyzed as a verb form on a par with the other non-periphrastic forms of the Basque verb, in the absence of a readjustment, this evolution would result in a split alignment pattern with A-alignment in a Present or Progressive tense.

Interestingly, there is some evidence that the grammaticalization of this periphrasis could rather to trigger a readjustment by analogy with the coding characteristics of the transitive construction in other tenses. For example (10a) and (10b) are two possible versions of a Basque sentence meaning ‘The companies are preparing the future managers’. The (a) version, with A in the Zero case and the intransitive auxiliary indexing A only, is the correct one according to normative grammar, but the (b) version, with A in the Ergative case and the transitive auxiliary indexing both A and P, was found in an official document of the Basque government (Celine Mounole, p.c.).

(10) Basque

a. *Enpresak etorkizuneko zuzendariak prestatzen ari dira.*

company.PL future manager.PL prepare.INCPL engaged be.PRS.3PL

‘The companies are preparing the future managers.’

---

<sup>17</sup> The auxiliary in the analytic conjugation of intransitive verbs is identical with the verb ‘be’, but the combination it forms with the auxiliated verb behaves differently from the combination of ‘be’ with its complement. Note that *da* as a form of the verb ‘be’ combining with nouns or adjectives in predicate function is glossed ‘be.PRS.3SG’, whereas *da* as the auxiliary of intransitive verbs is glossed ‘PRS.3SG’.

- b. *Enpresek etorkizuneko zuzendariak prestatzen ari dituzte.*  
 company.PL.ERG future manager.PL prepare.INCPL engaged be.PRS.3PL.3PL  
 same meaning as (a)

A possible explanation of this tendency to eliminate the possible alternation in the coding properties of the transitive construction resulting from the grammaticalization of the progressive periphrasis is that, originally, the development of this periphrasis was limited to some dialectal varieties of Basque (Joseba Lakarra, p.c.). It is now considered part of the standard Batua ('unified') variety, which means that it is now used by many speakers that do not have it in their native dialect, and may therefore be tempted to align it with the canonical transitive pattern.

#### 5.4.3. The 'bi-absolutive' construction of transitive verbs in Nakh-Daghestanian languages

As a rule, Nakh-Daghestanian languages have obligatory P coding and transitive constructions of the kind typically associated with obligatory P coding (A in the Ergative case, P in the Zero case, and gender-number agreement of the verb with P only), A so-called 'bi-absolutive construction' is however found with analytic verb forms expressing present tense, with both A and P are in the Zero case (hence the label 'bi-absolutive'), and the verb shows a complex agreement pattern.

For example, Avar has an analytic Present in which a participial form of the verb combines with the copula in auxiliary function. With transitive verbs, two constructions are possible. A first possibility is that A is in the Ergative case, P is in the Zero case, and the verb agrees with P only, as in the other tenses. In Ex. (11b), A is masculine and P neuter, and *b-el'ule-b b-ugo* shows neuter agreement in the prefix of the participle, in the suffix of the participle, and in the prefix of the auxiliary. A second possibility is that A and P are in the Zero case, and the verb shows a complex agreement pattern: if the auxiliated verb belongs to the class of verbs that have agreement prefixes, its prefix agrees with P, but the agreement suffix of the auxiliated verb and the auxiliary agree with A. In Ex. (11c), with the same nouns in A and P roles, *b-el'ule-w w-ugo* shows neuter agreement in the prefix of the participle only, whereas the suffix of the participle and the prefix of the auxiliary show masculine agreement (i.e., agreement with A).

(11) Avar (Alekseev and Ataev 1997)<sup>18</sup>

- a. *Emen w-ač'ule-w w-ugo.*  
 father SGM-coming-SGM SGM-COP  
 'Father is coming.'
- b. *Insu-ca χur b-el'ule-b b-ugo.*  
 father-ERG field SGN-ploughing-SGN SGN-COP  
 'Father is ploughing the field.'

<sup>18</sup> Note that the Avar noun for 'father' has two suppletive stems: *emen* in the Zero case, and *insu* in other cases.

c. *Emen* [χur b-el'ule-w] w-ugo.  
father field SGN- ploughing-SGM SGM-COP  
'Father is ploughing the field.'

The construction illustrated by Ex. (11c) can be analyzed as involving two clauses, a matrix clause headed by the copula and an embedded participial clause:

- the copula agrees with its sole argument *emen* 'father' in the Zero case;
- the agreement suffix of the participle reflects its status of head of a phrase that, taken as a whole, behaves as a predicative adjective phrase in a copular construction;
- the agreement prefix of the participle takes into account the syntactic relations within the phrase [χur b-el'ule-w].

A plausible explanation, elaborated by Harris and Campbell (1995: 187-189), is that (11c) maintains the biclausal structure of the original periphrasis, whereas in (11b), the original biclausal construction has been reinterpreted as a single clause, and the case and agreement marks have been readjusted under the pressure of the predominant pattern with A in the Ergative case, P in the Zero case, and agreement of the verb with P only. In other words, the construction illustrated by Ex. (11b) can be interpreted as resulting from the elimination of the violation of the Obligatory Coding Principle introduced by the grammaticalization of the progressive periphrases illustrated by Ex. (11c).

For a detailed presentation of the bi-absolutive construction in Nakh-Daghestanian languages, the cross-linguistic variation in its properties, and a discussion of the problems raised by its analysis, the reader is referred to Forker (2012).

## **5.5. Uncommon split alignment patterns, and the TAM periphrases of Basque**

### **5.5.1. Introductory remarks**

Dixon (1979: 95) makes the strong claim that "if a split is conditioned by tense or aspect, the ergative marking is ALWAYS found in either past tense or perfect aspect". Counterexamples have been found, in particular among Cariban languages – see in particular Gildea (1998), but Dixon (1994) discards them as insignificant, because of their "transitional" nature. However, there is nothing extraordinary in the existence of less common TAM-driven split alignment patterns, since some languages attest TAM periphrases whose grammaticalization, in the absence of a readjustment under the pressure of analogy, would automatically give rise to split alignment patterns contradicting the universal posited by Dixon.

For example, in addition to the progressive periphrasis analyzed in Section 5.4.2, Basque has several TAM periphrases whose grammaticalization could lead to the emergence of various split alignment patterns, since in the periphrases in question, the core arguments of transitive verbs are not encoded in the same way as in non-periphrastic constructions.

### 5.5.2. The future periphrasis of Basque with *joan* ‘go’

In this periphrasis, *joan* ‘go’ combines with the allative form of a verbal noun. Since *joan* is an intransitive verb, in this future periphrasis, the S/A argument of the auxiliated verb is uniformly treated as an S – Ex. (12).

(12) Basque

a. *Jonek berriak ikusiko ditu.*

Jon.ERG news.SG see.FUT PRS.3SG.3PL

‘Jon will watch the news.’ (non-periphrastic future)

b. *Jon [berriak ikustera] doa.*

Jon news.SG see.NOM.ALL go.PRS.3SG

‘Jon is going to watch the news.’ (periphrastic future)

Consequently, in the absence of a readjustment, the replacement of the non-periphrastic future by a form originating from this periphrasis (which is a very common phenomenon in the evolution of languages) could lead to a split alignment pattern with P-alignment in the future.

### 5.5.3. The modal periphrasis of Basque with *behar izan* ‘have to’

*Behar izan*, lit. ‘have need’, is a compound verb which can take a nominal complement in the Zero case, as in Ex. (13a), but is also found in a modal periphrasis in which it combines with the completive participle of the auxiliated verb – Ex. (13c) and (13e).<sup>19</sup> The behavior of this modal periphrasis is rather intricate – for a detailed discussion, see Hualde and Ortiz de Urbina (eds.) (2003: 301-308), but what is important in the perspective of this article is that, in conformity with the etymology, the person that has to do something can always be encoded as A in a transitive construction, even if the auxiliated verb is an intransitive verb assigning Zero case, as in Ex. (13c).

(13) Basque

a. *Jon-ek kotxe berri bat behar du.*

Jon-ERG car new one need have.PRS.3SG.3SG

‘Jon needs a new car.’

---

<sup>19</sup> Since in English, *need* can be either a verb or a noun, it is important to keep in mind that, in Basque, *behar* is a noun used here in the Zero case as the non-verbal element of a light verb compound whose verbal element is *izan* ‘have’. Formally speaking, *behar* can be viewed as fulfilling the P role in the construction of *izan*. However, in the construction illustrated by Ex. (14a), *Jonek* and *kotxe berri bat* behave exactly like the A and P NPs in the construction of a simplex transitive verb. Note that *du* as a form of the verb ‘have’ with the coding frame <ERG, Ø> is glossed ‘have.PRS.3SG.3SG’, whereas *du* as an auxiliary is glossed ‘PRS.3SG.3SG’.

b *Jon etxe-ra doa.*

Jon house-SG.ALL go.PRS.3SG  
'Jon is going home.'

c. *Jon-ek [etxe-ra joan] behar du.*

JON-ERG house-SG.ALL go.CPL.need have.PRS.3SG.3SG  
'Jon must go home.'

d. *Jon-ek ogi-a erosten du.*

JON-ERG bread-SG buy.INCPLPRS.3SG.3SG  
'Jon is buying bread.'

e. *Jon-ek [ogi-a erosi] behar du.*

JON-ERG bread-SG buy.CPL need have.PRS.A3SG.P3SG  
'Jon must buy bread.'

Consequently, in the absence of a readjustment, the grammaticalization of this periphrasis as the usual expression of future might lead to a split alignment pattern with A-alignment in the future, but with the atypical variety of case marking that I propose to characterize as *generalized ergative marking*.

Interestingly, the grammaticalization of the *behar izan* periphrasis with a future meaning is not attested in present-day Basque, but two or three centuries ago, the Lapurdian dialect initiated such an evolution – Mounole (2011: 191), and this semantic shift was accompanied by a possible regularization of the construction quite similar to that mentioned above for the progressive periphrasis. For example, 'he will come' occurs in the same text as *jin behar du*, with the transitive auxiliary *du*, and *jin behar den*, with the dependent form of the intransitive auxiliary *den*. The first variant is that expected from the presence of the transitive verb 'have' in this periphrasis, whereas in the second variant, the sole argument of 'come' has the coding properties expected for the sole argument of a monovalent verb.

#### 5.5.4. The resultative periphrasis of Basque

We now turn to a resultative periphrasis in which the verb *izan* 'be'<sup>20</sup> combines with the completive participle in the definite form treated as an adjectival predicate.<sup>21</sup>

Contrary to Indo-European 'past participles', this form of the Basque verb (glossed CPL) is not particularly patient-oriented, and with transitive verbs it can combine with an ergative-marked agent, like the finite forms of transitive verbs. Consequently, when transitive verbs occur in the resultative periphrasis, the raised argument can indifferently be A or P, and the non-raised argument is treated exactly as in an independent clause, which gives rise to constructions that are often designated as passive (if the raised term is P) and antipassive (if the raised term is A) – Ex. (14).

---

<sup>20</sup> The participial form conventionally used as the quotation form of Basque verbs is the same for *izan* 'be' and *izan* 'have', but the finite forms of these two verbs are distinct.

<sup>21</sup> In most dialects of Basque, nouns and adjectives in predicate function are in the definite form.

(14) Basque

a. *Jon-ek eskutitz bat idatzi du.*

Jon-ERG letter one write.CPL PRS.3SG.3SG

‘Jon wrote a letter.’ (non-periphrastic completive)

b. *Eskutitz hau [Jon-ek idatzi-a] da.*

letter DEM.SG Jon-ERG write.CPL-SG be.PRS.3SG

‘This letter has been written by Jon’ (‘passive’ variant of the resultative periphrasis, lit. ‘This letter is [Jon (having) written (it)].’)

c. *Jon [eskutitz asko idatzi-a] da.*

Jon letter many write.CPL-SG be.PRS.3SG

‘Jon has written many letters.’ (‘antipassive’ variant of the resultative periphrasis, lit. ‘Jon is [(having) written many letters].’)

In the ‘passive’ variant, there is no change in case assignment in comparison with the non-periphrastic construction of a transitive verb, and *izan* ‘be’ agrees with an S representing the P argument. Consequently, its grammaticalization would induce no modification in the alignment patterns of Basque, and the loss of verb agreement with A would even reinforce the consistency of P-alignment. By contrast, the grammaticalization of the ‘antipassive’ variant could lead to a split alignment pattern with A-alignment in the perfect. In other words, Basque attests a resultative periphrasis whose grammaticalization, in the absence of a readjustment, would automatically lead to the emergence of a split alignment pattern in clear contradiction with Dixon’s universal.

### 5.6. *Concluding remarks*

The data examined in this section shows that the grammaticalization of TAM constitutes a potential source of a wide variety of split-alignment patterns violating the Obligatory Coding Principle. But it also shows that this automatic result of the grammaticalization of TAM tends to be canceled by subsequent evolutions that align the coding properties of the verb form created by the grammaticalization process with those of the pre-existing verb forms.

Consequently, there is no need to look for direct semantic / functional explanations of the fact that just a few TAM-driven split-alignment patterns are well-attested cross-linguistically, while others are marginal or not attested at all. Given the strong tendency to eliminate the violations of the Obligatory Coding Principle resulting from the grammaticalization of TAM, it can be predicted that, among the split-alignment patterns resulting from such grammaticalization processes, the only ones that have a relatively good chance to stabilize in at least some languages are those resulting from particularly common grammaticalization paths. The explanation of the relative frequency of some particular split-alignment pattern must therefore not be sought in their intrinsic properties. The real question is why some types of evolutions leading to the grammaticalization of new TAM forms are more common than others, and this question has no direct link with alignment typology.

## 6. Conventionalization of ellipsis and the Obligatory Coding Principle

### 6.1. *Introductory remarks*

Depending on the individual languages, A and P may behave as terms of the transitive construction that must obligatorily be expressed (either by means of NPs, or through indexation), but the mere absence of any morphological material referring to a core term can also be used to signal that the missing argument must be, either anaphorically identified with a salient referent, or interpreted as indeterminate. For example, in English, the absence of P in *He is eating* implies an indeterminate interpretation of the patient; in other languages, a formally identical construction would be interpreted as ‘He is eating it’, or would be ambiguous between an indeterminate and an anaphorical reading.

The use of A or P ellipsis with an indeterminate interpretation is not limited to languages in which a particular type of alignment predominates. For example, P ellipsis with an indeterminate interpretation is perhaps particularly common among obligatory A coding languages, but some obligatory A coding languages ignore this possibility (for example, Nahuatl systematically uses ‘indeterminate object prefixes’ – see Launey (1994: 155-159)), and P ellipsis with an indeterminate interpretation is widespread among languages with other alignment patterns too (for example, in Basque, depending on the context, *ikusten dute*, with A agreement of 3rd person plural and P agreement of 3rd person singular, can equally express ‘they see it / him / her’ or ‘they can see’).

Historically, transitive constructions in which the absence of a core argument expresses indeterminacy may undergo evolutions converting them into intransitive constructions. More or less complicated scenarios can be imagined, depending on the particularities of the individual languages, but the most obvious one is that the verb in question simply loses the ability to be used in a full transitive construction, and consequently undergoes a reduction of the number of its arguments.

In obligatory A coding languages, the reanalysis of a transitive construction from which P is missing as an intransitive construction has no consequence on alignment, since a term showing A-like coding is still present. But in obligatory P coding languages, the result is the emergence of a non-canonical coding frame involving no term having the coding characteristics of P.

Symmetrically, in obligatory P coding languages, the reanalysis of a transitive construction from which A is missing as an intransitive construction has no consequence on alignment, since a term showing P-like coding is still present. But in obligatory A coding languages, the result is the emergence of a non-canonical coding frame involving no term having the coding characteristics of A.

### 6.2. *Conventionalization of P ellipsis in obligatory P coding languages: an illustration from Akhvakh*

In Akhvakh, as in most languages belonging to the Avar-Andic-Tsezic branch of the Nakh-Daghestanian family, exceptions to the rule of obligatory P coding are marginal: in the transitive construction, A in the Ergative case contrasts with P in the Zero case

and the verb agrees in gender and number with P, and with few exceptions, the coding frames of verbs include a term in the Zero case governing verb agreement in gender and number like P in the transitive construction – Ex. (15).

(15) Akhvakh

a. *Ek'wa w-oq'-ilo.*

man SGM-come-CPL.NEG.SGM

'The man did not come.'

b. *Jaše j-eq'-ile.*

girl SGF-come-CPL.NEG.SGF

'The girl did not come.'

c. *Mašina b-eq'-ile.*

car SGN-come-CPL.NEG.SGN

'The car did not come.'

d. *Ek'waš<sup>w</sup>-e jaše j-ič'-ile.*

man-ERG girl SGF-push-CPL.NEG.SGF

'The man did not push the girl.'

e. *Ek'waš<sup>w</sup>-e mašina b-ič'-ile.*

man-ERG car SGN-push-CPL.NEG.SGN

'The man did not push the car.'

f. *Jašo-de ek'wa w-uč'-ilo.*

girl-ERG man SGM-push-CPL.NEG.SGM

'The girl did not push the man.'

Depending on a complex combination of grammatical and lexical factors, verb agreement in gender and number is not always apparent.<sup>22</sup> As regards P ellipsis, depending on the context, in the absence of an NP in P role, transitive verb forms showing neuter singular agreement or devoid of any apparent agreement mark may equally have an anaphorical or indeterminate interpretation.

In addition to the canonical valency patterns characterized by the presence of a term in the Zero case governing the agreement of the verb in gender and number, Akhvakh has a limited class of verbs with non-canonical valency patterns involving an argument in the ergative case and an argument in a spatial case, but no argument

---

<sup>22</sup> In Akhvakh, gender-number agreement of verbs involves prefixes and suffixes. The presence of agreement prefixes is lexically determined (verbs divide into two morphological classes, those that have agreement prefixes in all their forms, and those devoid of agreement prefixes), whereas agreement suffixes occur in certain tenses only, irrespective of the presence or absence of lexically determined agreement prefixes. The verbs of Ex. (16) (*b-eq'urula* 'come' and *b-ič'urula* 'push', conventionally quoted in isolation with the singular neuter prefix) belong to the class of verbs with obligatory agreement prefixes, whereas the verbs of the following examples all belong to the class of verbs devoid of agreement prefixes.

that could be represented by an NP in the Zero case. With respect to agreement, the verbs in question show neuter singular default agreement – Ex. (16).

(16) Akhvakh

a. *Ek'waš<sup>w</sup>-e jašo-ga eḡ-ere godi.*  
man-ERG girl-ALL look\_at-PROG COP.SGN  
'The man is looking at the girl.'

b. *χwe-de jašo-ge ḡ'eleč'-ari.*  
dog-ERG girl-LOC bite-CPL  
'The dog bit the girl.'

c. *Mik'i-de di-ge ḡ'it'-ari.*  
child-ERG 1SG-LOC pinch-CPL  
'The child pinched me.'

There is no direct evidence that a P argument was ever present in the construction of these verbs, and several types of explanations of such exceptional valency patterns can be considered. In some cases, the most plausible explanation is the univerbation of former light verb compounds (see Section 7), but in some others, a plausible explanation is the conventionalization of P ellipsis in constructions that, originally, were perfectly canonical transitive constructions.

For example, the verb *l̄<sup>w</sup>arurula* is commonly encountered with the meaning 'hit' in a construction superficially similar to those illustrated by Ex. (16), with an argument in the Ergative case and an argument in the Locative case – Ex. (17).

(17) Akhvakh

*Ek'waš<sup>w</sup>-e jašo-ge l̄<sup>w</sup>ar-ari.*  
man-ERG girl-LOC hit-CPL  
'The man hit the girl.'

At first sight, *l̄<sup>w</sup>arurula* might appear as a bivalent verb with an exceptional valency pattern, but in fact, it is a trivalent verb, and (18) is the elliptical variant of a perfectly canonical coding frame with an oblique argument in addition to A and P.

The point is that the same verb with the same meaning 'hit' (or closely related meanings) is also found in a construction in which a term in the Zero case governing the agreement of the verb in gender and number represents the instrument used by the hitter to perform his/her action – Ex. (18).

(18) Akhvakh

a. *Ek'waš<sup>w</sup>-e jašo-ge rel'a l̄<sup>w</sup>ar-ari.*  
man-ERG girl-LOC hand hit-CPL  
'The man hit the girl with his hand.' (lit. applied the hand to the girl)

b. *Ek'was̄<sup>w</sup>-e jašo-ge ža ī<sup>w</sup>ar-ari.*  
man-ERG girl-LOC fist hit-CPL

‘The man hit the girl with his fist.’ (lit. applied the fist to the girl)

c. *Ek'was̄<sup>w</sup>-e jašo-ge č'uli ī<sup>w</sup>ar-ari.*  
man-ERG girl-LOC stick hit-CPL

‘The man hit the girl with a stick.’ (lit. applied a stick to the girl)

d. *Toχtoros̄<sup>w</sup>-e jašo-ge mežu ī<sup>w</sup>ar-ari.*  
doctor-ERG girl-LOC needle hit-CPL

‘The doctor gave an injection to the girl.’ (lit. applied a needle to the girl)

Consequently, the basic meaning of *ī<sup>w</sup>arurula* is ‘someone applies something to a surface rapidly/violently’. In Akhvakh, as in other Caucasian languages, the hittee is not conceptualized as the patient of a two participant action, but as the target at which an agent is aiming a missile, and (17) is still recognizable as a transitive construction from which P is missing: ‘The man applied [an unspecified object] to the girl’, or ‘The man aimed [an unspecified object] at the girl’. Starting from that, one can easily imagine that at least some of the bivalent verbs of Akhvakh that have non-canonical coding frames occurred initially in a construction including a P term whose elision was subsequently conventionalized.

### **6.3. Conventionalization of A ellipsis in obligatory A coding languages: illustrations from Amharic and Russian**

#### **6.3.1. Introductory remarks**

As discussed in three of the papers included in Donohue and Wichmann (2008), in languages in which A-alignment predominates, the reanalysis of P in elliptical transitive constructions (or ‘transimpersonal’<sup>23</sup> constructions) as the unique core argument of intransitive constructions can be a source of systems in which, in violation of the Obligatory Coding Principle, the set of monovalent verbs divides into two subsets differing in the alignment of their unique argument. Holton (2008) and Mithun (2008) discuss comparative evidence supporting the hypothesis that, in various languages from the Americas and Papua New Guinea, such systems developed from the reanalysis of transimpersonal constructions as intransitive constructions with P coding of the unique argument. Malchukov (2008) proposes a wider discussion of the evolutions of transimpersonal constructions, rightly pointing out that their reanalysis as intransitive constructions has no consequence on alignment patterns in obligatory P coding languages (for example, in the Iwaidjan languages discussed by Evans (2004)), whereas the same reanalysis may trigger the development of non-canonical coding frames when it occurs in obligatory A coding languages.

In this section, I illustrate this point by the comparison of Amharic and Russian impersonal constructions that can be viewed as representing two different stages in

---

<sup>23</sup> According to Malchukov (2008), this term was coined by Mary Haas – Haas (1941).

the evolution of transimpersonal constructions towards plain intransitive constructions with a non canonical alignment pattern.

### 6.3.2. *An illustration from Amharic*

Amharic is a language with obligatory A coding in which A NPs are obligatorily cross-referenced by person markers which, in the absence of a co-referent NP, normally trigger an anaphorical interpretation. Amharic also has constructions, traditionally identified as ‘impersonal’, that can be analyzed as elliptical transitive constructions in which the absence of an NP in A role exceptionally triggers an indeterminate rather than anaphorical interpretation.

For example, the state of being hungry, without any hint about a possible external cause, is rendered in Amharic by a verb showing a non-referential A index of 3rd person singular masculine, and a P index representing the person or animal being hungry – Ex. (19a). But the same verb also occurs in a canonical transitive construction in which A and P are respectively assigned the roles of stimulus and experiencer – Ex. (19b).

(19) Amharic (Leslau 2005: 43)

a. *Rabä-ñ.*

hunger.CPL. 3SGM-1SG

‘I am hungry.’, lit. ‘It hungered me.’

b. *Īnjära rabä-ñ.*

bread hunger.CPL.3SGM-1SG

‘I am hungry for bread.’, lit. ‘Bread hungered me.’

Starting from situations of this type, one can easily imagine how the loss of the construction illustrated by Ex. (19b) may result in the emergence of monovalent verbs whose exceptional construction cannot be explained as an elliptical transitive construction anymore, and must be viewed as an instance of P-alignment in a language in which A-alignment is canonical.

### 6.3.3. *An illustration from Russian*

The impersonal construction of the Russian verb *trjasti* ‘shake’ results from an evolution of this type. This verb occurs in a canonical transitive construction – Ex. (20a), but also in an impersonal construction that, synchronically, cannot be analyzed as an elliptical variant of the transitive construction, since in this construction, the participant expressed as the A term of the transitive construction can be encoded as an oblique introduced by the preposition *ot* ‘from’. In Ex. (20b), the only core term is an experiencer in the accusative case; it would be ungrammatical to add an NP in the Zero case, and an external cause can only be mentioned by means of a preposition phrase in oblique role.

(20) Russian

a. *Ja trjasu kovër.*

1SG shake.PRS.1SG carpet.ACC

‘I am shaking the carpet.’

b. *Menja trjasët (ot lixoradki).*

1SG.ACC shake.PRS.3SG (from fever.GEN)

‘I am shaking (with fever).’, lit. ‘It shakes me (from fever).’

It seems however reasonable to assume that the impersonal construction illustrated by (20b) developed as an elliptical variant of the transitive construction: ‘[An unspecified cause] shakes me’. But the fact that the cause is now encoded as an oblique introduced by the ablative preposition *ot* proves that, in the present state of Russian, this construction is no longer an elliptical variant of the transitive construction, and has been reanalyzed as a construction of its own.

## 7. Univerbation of light verb compounds and the Obligatory Coding Principle

Some languages have a particularly high proportion of predicates expressed by means of light verb compounds in which the light verb is a transitive verb, most often a verb with the meaning ‘do, make’, as in Ex.(21), and the non-verbal element is a noun encoded like the P argument of transitive verbs (Samvelian 2012: 16).

(21) Basque

a. *Haurr-ek lo egiten dute.*

child-PL.ERG sleep do.INCPL PRS.3SG.3PL

‘The children are sleeping (lit. are doing sleep).’

b. *Gizon horr-ek ez du euskar-az hitz egiten.*

man DEM.SG-ERG NEG PRS.3SG.3SG Basque-SG.INSTR word do.INCPL

‘This man does not speak Basque (lit. does not do word in Basque).’

The coding frame of such predicates can be schematized as A (X) p V, where V is the verbal element of the light verb compound, (lower-case) p represents the non-verbal element of the compound, coded as if it were the P argument of a transitive verb, A represents an argument to which A coding is assigned, and (X) represents possible additional terms whose presence depends on the argument structure of the predicate, and to which an oblique-like coding is assigned.

In obligatory A coding languages, considering the non-verbal element of the compound as a term in the construction of the light verb or considering the compound p V as a whole as the syntactic equivalent of a simplex verb does not change anything with respect to the Obligatory Coding Principle, since a term of the construction representing a participant is encoded like the A argument of a transitive verb. By contrast, in obligatory P coding languages, the principle is formally satisfied by the nominal element of the compound, insofar as it is considered a term in the

construction of the light verb, but the principle is violated if the complex predicate is taken as a whole, and only NPs representing participants are considered terms of the predicative construction.

Diachronically, there is a general tendency toward fusion of the two elements of such compounds. This univertation process converts formally transitive constructions A (X) p V (where ‘p’ symbolizes the P coding of a word that does not represent a participant) into A (X) V, i.e. constructions with a term showing A coding but no term showing P coding. In obligatory A coding languages, this results in perfectly canonical constructions, whereas in obligatory P coding languages, the same process results in a violation of the Obligatory Coding Principle. Interestingly, some obligatory P coding languages show a strong tendency toward regularization of the non-canonical coding frames resulting from this process, whereas others tend to maintain them without modification.

As already illustrated by Ex.(21) above, Basque makes wide use of light verb compounds consisting of a bare noun and the verb *egin* ‘do, make’. The argument structure of light verb compounds like *lo egin* ‘sleep’ or *hitz egin* ‘speak’ can be represented as <ERG,  $\emptyset$ >, where (uppercase) ERG symbolizes the slot for the argument of the light verb compound taken as a whole, and (lowercase)  $\emptyset$  symbolizes the slot for the non-verbal element of the compound, whose coding characteristics are similar to those of the P argument of transitive verbs. As observed by Etxepare (2003: 397), such compounds “are not instances of incorporation ... the bare nominal and the verb *egin* can be separated by a number of syntactic operations, and the bare nominal can take partitive case” – see Oyharçabal (2007) for a more detailed analysis of Basque light verb compounds.

Many of the light verb compounds of Basque correspond to simplex verbs whose root coincides with the non-verbal element of the compound, like *bultzza egin* lit. ‘do impulse’ / *bultzatu* ‘push’ – Ex. (22).<sup>24</sup>

(22) Basque

- a. *Mutil-ak ate-ari bultzza egin zion.* <ERG, DAT,  $\emptyset$ >  
 boy-SG.ERG door-SG.DAT impulse do.CPLPST.3SG.3SG.3SG  
 ‘The boy pushed the door.’
- b. *Mutil-ak ate-a bultzatu zuen.* <ERG,  $\emptyset$ >  
 boy-SG.ERG door-SG push.CPL PST.3SG.3SG  
 same meaning as (a)

In this example, a light verb compound used in the frame <ERG, DAT,  $\emptyset$ > corresponds to a simplex transitive verb, which means that the Dative argument of the light verb compound represents the same participant as the argument of the simplex verb in the Zero case. However, the arguments of the simplex verb are more commonly encoded in the same way as in the light verb construction. For example, the light verb compound *dirdir egin* ‘shine’ (lit. ‘do shining’) and the corresponding

<sup>24</sup> *Bultzatu* is the completive participle, used in Basque grammars and dictionaries as the quotation form of verbs. It can be decomposed as *bultzza-* (root) plus *-tu* (completive aspect marker).

simplex verb *dirdiratu* equally assign the Ergative case to their argument – Ex. (23). In other words, *dirdiratu* belongs to a class of verbs with no argument in the Zero case that included very few verbs in Old Basque but has grown dramatically in the history of most Basque varieties.

(23) Basque

a. *Eguzki-ak dirdir egiten du.* <ERG,  $\emptyset$ >  
sun-SG.ERG shining do.INCPL PRS.3SG.3SG  
'The sun is shining.'

b. *Eguzki-ak dirdiratzten du.* <ERG>  
sun-SG.ERG shine.INCPL PRS.3SG.3SG  
same meaning as (a)

The variation observed in the coding frames of simplex verbs cognate with the non-verbal element of an *egin*-compound can be viewed as the result of the interaction between two conflicting tendencies: a tendency to align the encoding of the arguments of the simplex verbs with the encoding of the same arguments in the light verb construction, as in Ex. (23), and a tendency to organize the coding frame of the simplex verb according to the principle of obligatory P coding, as in Ex. (22).

In (Creissels, To appear), I argue that the predominance of the tendency to align the encoding of the arguments of the simplex verbs with the encoding of the same arguments in the light verb construction observed in most Basque varieties can be explained as part of a 'conspiracy' towards extension of Ergative coding (and consequently against obligatory P coding). This phenomenon has its roots in particularities of the argument coding system that already existed in Old Basque, but it was considerably reinforced by a rule according to which, with the only exception of so-called pronominal verbs (i.e. verbs combined with the Romance reflexive clitic *se*), Spanish or French verbs borrowed into Basque assign Ergative coding to their subject, irrespective of their valency.

In the same paper, I compare the Basque situation with that of Andic languages, a group of closely related Nakh-Daghestanian languages spoken in the western part of Daghestan, which like Basque make a wide use of light verb compounds, but in which, contrary to Basque, there is a strong tendency to eliminate the violations of the Obligatory Coding Principle following from the univerbation of light verb compounds. In spite of the extensive use of light verb compounds consisting of a transitive verb and a noun in the Zero case, Andic languages have very few verbs with coding frames violating the obligatory P coding principle, even among verbs whose etymology as resulting from the univerbation of a light verb compound is obvious, and none of the verbs in question is used with a non-canonical coding frame in all Andic languages.

Therefore the question is why, in some of the obligatory P coding languages that make a wide use of light verb constructions in which the non-verbal element of the light verb compound is a noun encoded as if it represented a patient, there is a very strong tendency to eliminate the violations of the Obligatory Coding Principle that arise from the univerbation of light verb compounds, whereas in others, the

univerbation of light verb compounds contributes to an increase in the proportion of verbs with coding frames violating the principle of obligatory P coding.

In the aforementioned article I argue that, in the history of Basque, the weakening of the tendency toward regularization of coding frames contradicting the principle of obligatory P coding must be viewed as part of a general trend toward relaxation of the constraints limiting the use of Ergative coding in systems characterized by 'strict' ergative coding, according to a distinction between strict and loose ergative coding introduced by Harris (1985) and applied to Basque by Aldai (2008). And precisely, in this respect, Andic languages are strikingly different from Basque, which suggests that, in languages that stand closer to the prototype of strict ergative coding, with an ergative case relatively marked semantically, the coding frames contradicting the obligatory P coding principle that arise as the result of the univerbation of light verb compounds tend rather to be modified in order to conform to the principle of obligatory P coding.

## **8. Conclusion**

In this paper, I have examined two possible types of diachronic processes that may affect the status of argument coding systems with respect to the Obligatory Coding Principle: global changes affecting at the same time the whole set of transitive verbs and resulting in a modification of the characteristics of transitive coding, and gradual changes resulting in the creation of intransitive verbs with non-canonical alignment.

Two subtypes of global changes must be distinguished. A first possibility is that the transitive construction is replaced by another construction that was already available for transitive verbs, but with the status of a passive or antipassive variant of the basic transitive construction. A second possibility is that the grammaticalization of new TAM forms modifies the transitive type of argument coding by introducing TAM-driven alternations in the coding of A and P. As regards the possibility of gradual changes resulting in the creation of intransitive verbs with non-canonical alignment, two types have been discussed: the conventionalization of argument ellipsis, and the univerbation of light verb constructions.

The types of changes likely to automatically result in violations of the Obligatory Coding Principle (grammaticalization of TAM periphrases, conventionalization of argument ellipsis, and univerbation of light verb compounds) are very common types of diachronic processes, and one may therefore wonder why systems of transitive coding with TAM-driven alternations in the coding of A and P, or with two substantial classes of monovalent verbs differing in the coding properties of their sole argument, are not more widespread among the languages of the world. The obvious explanation is that the violations of the Obligatory Coding Principle triggered by such processes tend to be canceled by readjustment under the pressure of analogy: the alternations in transitive coding triggered by the grammaticalization of new TAM TAM tend to be eliminated by the alignment of the construction of the newly created TAM form with the pre-existing ones, and the intransitive verbs with exceptional coding frames resulting from the conventionalization of argument ellipsis or from the univerbation of light verb compounds tend to replace them more canonical coding frames. But this leads to the opposite question: if the tendency to eliminate the violations of the Obligatory Coding Principle is so strong, how is it possible that

nevertheless, systems of argument coding involving important violations of the Obligatory Coding Principle are not exceptional?

The case of Basque suggests that such situations arise as the result of the interplay of a complex set of factors. But Basque is well documented from the 16th century, and unfortunately, for most of the other languages that could improve our understanding of the tension between changes that automatically result in violations of the Obligatory Coding Principle and changes that eliminate such violations, the explanation can only be speculative, for lack of historical data.

## Abbreviations

A: agent, ABL: ablative, ACC: accusative, ALL: allative, ANTIP: antipassive, CPL: completive aspect, COP: copula, DAT: dative, DEF: definite, DEM: demonstrative, ERG: ergative, F: feminine, FUT: future, GEN: genitive, INCPL: incompletive aspect, INF: infinitive, INSTR: instrumental, LOC: locative, M: masculine, MOD: modal, N: neuter, NEG: negation, OBL: oblique, Ø: zero case, P: patient, PASS: passive, PP: past participle, PREV: preverb, PROG: progressive, PRS: present, PST: past, S: sole argument of monovalent verbs, SG: singular, V: verb, X: oblique.

## References

- Aldai, Gontzal. 2008. From ergative case marking to semantic case marking: the case of historical Basque. In Mark Donohue and Søren Wichmann (eds.) *The typology of semantic alignment*. Oxford: Oxford University Press. 197-218.
- Alekseev, Mikhail and Boris Ataev. 1997. *Avarskij jazyk* (The Avar language). Moscow: Academia.
- Benveniste, Emile. 1952. La construction passive du parfait transitif. *Bulletin de la Société de Linguistique de Paris* 48. 52-62.
- Blau, Joyce and Veysi Barak. 1999. *Manuel de kurde. Kurmanci*. Paris: L'Harmattan.
- Bobaljik, Jonathan David. 1993. On ergativity and ergative unergatives. *MIT Working Papers in Linguistics* 19. 45-88.
- Bynon, T. 2005. Evidential, raised possessor, and the historical source of the ergative construction in Indo-Iranian. *Transactions of the Philological Society* 103:1. 1-72.
- Cardona, G. 1970. The Indo-Iranian construction *mana (mama) krtam*. *Language* 46. 1-12.
- Carrier, Julien. 2012. L'expression de la transitivité en Itivimiut. Masters thesis. Université du Québec à Montréal.
- Creissels, Denis. 2008. Direct and indirect explanations of typological regularities: the case of alignment variations. *Folia Linguistica* 42-1. 1-38.
- Creissels, Denis. 2009. Uncommon patterns of core term marking and case terminology. *Lingua* 119. 445-459.
- Creissels, Denis. To appear. Univerbation of light verb compounds and the Obligatory Coding Principle. In Léa Nash and Pollet Samvelian (eds.), *Approaches to Complex Predicates*. Leiden: Brill.
- Creissels, Denis and Alain-Christian Bassène. 2013. Valency patterns for bivalent verbs in two West African languages: Mandinka (Mande) and Jóola Banjal (Atlantic). *Afrikanistik Online*.

- Creissels, Denis and Céline Mounole. 2012. Non-canonical valency patterns in Basque, variation and evolution. Paper presented at the 45th Annual Meeting of the Societas Linguistica Europaea (Stockholm, 29 August – 1 September 2012).
- DeLancey, Scott. 2004. The blue bird of ergativity. In Queixalós, Francesc (ed.), *Ergativity in Amazonia III*. Paris: Celia.
- Dixon, Robert. 1979. Ergativity. *Language* 55. 59-138.
- Dixon, Robert. 1994. *Ergativity*. Cambridge: Cambridge University Press.
- Donohue, Mark and Søren Wichmann. 2008. *The typology of semantic alignment*. Oxford: Oxford University Press.
- Etxepare, Ricardo. 2003. Valency and argument structure in the Basque verb. In José Ignacio Hualde and Jon Ortiz de Urbina (eds.) 2003. 363-426.
- Evans, Nicholas. 2004. Experiencer object in Iwaidjan languages. In Bhaskararao, P. and K.V. Subbarao (eds.), *Non-nominative subjects*. vol. 1. Amsterdam and Philadelphia: John Benjamins. 169-193.
- Forker, Diana. 2012. The bi-absolutive construction in Nakh-Daghestanian. 46(1). 75-108.
- Gildea, Spike. 1997. Evolution of grammatical relations in Cariban: How functional motivation precedes syntactic change. In T. Givón (ed.), *Grammatical relations: a functionalist perspective*. Amsterdam and Philadelphia: John Benjamins. 155-198.
- Gildea, Spike. 1998. *On reconstructing grammar: Comparative Cariban morphosyntax*. Oxford: Oxford University Press.
- Haas, Mary. 1941. Tunica. In *Handbook of American Indian languages*, vol. 4. New York: Augustin Publishers.
- Harris, Alice. 1985. *Diachronic Syntax: The Kartvelian Case (Syntax and Semantics, 18)*. New York: Academic Press.
- Harris, Alice. 1997. Review of *Ergativity* by R.M.W. Dixon. *Language* 73(2). 359-374.
- Harris, Alice and Lyle Campbell. 1995. *Historical syntax in cross-linguistic perspective*. Cambridge: Cambridge University Press.
- Hartmann, Iren, Martin Haspelmath and Andrej Malchukov. 2013. How widespread is transitive encoding? Paper presented at the ALT 10 Conference (Leipzig, August 15-18, 2013).
- Haspelmath, Martin. 2010. Comparative concepts and descriptive categories in cross-linguistic studies. *Language* 86(3). 663-687.
- Haude, Katharina and Fernando Zúñiga. To appear. Inverse systems and symmetrical voice: a unified view on languages with two transitive constructions. *Linguistics*.
- Holton, Gary. 2008. The emergence of stative-active systems in North Halmahera, Indonesia. In Donohue, Mark and Søren Wichmann (eds.), *The typology of semantic alignment*. Oxford: Oxford University Press.
- Hualde, José Ignacio and Jon Ortiz de Urbina (eds.). 2003. *A Grammar of Basque*. Berlin and New York: Mouton de Gruyter.
- Janic, Katarzyna. 2013. L'antipassif dans les langues accusatives. PhD thesis. University of Lyon.
- Laka, Itziar. 1993. Unergatives that assign ergative, Unaccusatives that assign accusative. In Jonathan Bobaljik and Colin Phillips (eds.) *Papers on case and agreement I*. volume 18. 149-172. MIT Working Papers in Linguistics.

- Laka, Itziar. 2000. Thetablind case: Burzio's generalisation and its image in the mirror. In Eric Reuland (ed.) *Arguments and case: Explaining Burzio's generalisation*. 103-129. Amsterdam: John Benjamins.
- Launey, Michel. 1994. *Une grammaire omniprédicative, essai sur la morphosyntaxe du nahuatl classique*. Paris : CNRS éditions.
- Lazard, Gilbert. 1994. *L'actance*. Paris: Presses Universitaires de France.
- Lazard, Gilbert. 1997. Review of *Ergativity* by R. M. W. Dixon. *Linguistic Typology* 1: 243–268.
- Leslau, W. 2005. *Concise Amharic dictionary* (2nd Ethiopian edition). Addis Ababa: Shama books.
- Malchukov, Andrej. 2008. Split intransitives, experiencer objects and 'transimpersonal' constructions: (re-)establishing the connection. In Donohue, Mark and Søren Wichmann (eds.), *The typology of semantic alignment systems*. Oxford: Oxford University Press.
- Mithun, Marianne. 2008. The Emergence of Agentive Patient Systems in Core Argument Marking. In Donohue, Mark and Søren Wichmann (eds.), *The typology of semantic alignment*. Oxford: Oxford University Press.
- Mounole, Céline. 2011. *Le verbe basque ancien : étude philologique et diachronique*. PhD thesis. University of Bordeaux and University of the Basque Country.
- Nedjalkov, Vladimir P. and Sergej Je. Jaxontov (eds.). 1988. *Typology of resultative constructions*. Amsterdam and Philadelphia: John Benjamins.
- Nichols, Johanna. 1992. *Linguistic diversity in space and time*. Chicago: University of Chicago Press.
- Oyharçabal, Bernard. 2007. Basque light verb constructions. In Joseba A. Lakarra and José Ignacio Hualde (eds.) *Studies in Basque and Historical Linguistics. In Memory of R. L. Trask*. Bilbao: Diputacion Foral de Gipuzkoa - Gipuzkoako Foru Aldundia and Universidad de País Vasco - Euskal Herriko Unibertsitatea, Bilbao. 787-806.
- Peterson, J. 1998. *Grammatical relations in Pali and the emergence of ergativity in Indo-Aryan*. München: LINCOM.
- Queixalós, Francesc. 2013. *L'ergativité est-elle un oiseau bleu?* München: LINCOM.
- Rezac, Milan. 2008a. The syntax of eccentric agreement: The person case constraint and Absolutive displacement in Basque. *Natural Language and Linguistic Theory* 26. 61-106.
- Rezac, Milan. 2008b. The forms of Dative displacement. From Basauri to Itelmen. In Xabier Artiagoitia and Joseba A. Lakarra (eds.) *Gramatika jaietan: Patxi Goenagaren omenez*. Universidad del País Vasco - Euskal Herriko Unibertsitatea. 709-724.
- Samvelian, Pollet. 2012. *Grammaire des prédicats complexes: Les constructions nom-verbe*. Paris: Lavoisier.
- Seržant, Ilja A. 2012. The so-called possessive perfect in North Russian and the Circum-Baltic area. A diachronic and areal account. *Lingua* 122. 356-385.
- Spreng, Bettina. 2005. Third person arguments in Inuktitut. In Solveiga Armoskaite and James J. Thompson (eds.), *Workshop on the structure and constituency of the languages of the Americas*. University of British Columbia: UBC Working Papers in Linguistics. 215-225.
- Tsunoda, Tasaku. 1985. Remarks on transitivity. *Journal of Linguistics* 21(2). 385-396.