

The Obligatory Coding Principle in diachronic perspective

Denis Creissels
University of Lyon

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 chapter examines the types of evolutions that may result either in global shifts affecting the Obligatory Coding Principle, in systematic violations of the Obligatory Coding 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 chapter is about the possible consequences of some diachronic processes (TAM grammaticalization, conventionalization of argument ellipsis, univerbation of light verb compounds, etc.) for the structure of argument coding systems, i.e. for aspects of the organization of languages commonly dealt with in terms of morphological *accusativity* vs. *ergativity*.¹

The discussion is structured around the Obligatory Coding Principle, according to which all verbal predicative constructions in a given language must include a nominal term showing a particular type of coding that can be viewed as the default type of argument coding in the language in question. This definition is intended to capture the tendency toward consistency in morphological alignment between transitive and intransitive predications.

The possible links between the diachronic processes affecting argument coding systems analyzed here and so-called ‘deep/syntactic ergativity’ are not discussed in this chapter, although a plausible connection is mentioned in section 4.2.²

The chapter 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 may result in 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

¹ Some of the questions dealt with here were already discussed in a slightly different perspective in Creissels 2008.

² A detailed discussion of this question for the types of changes dealt with in section 4 can be found in Queixalós 2013.

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 Transitivity

Verbs encoding events involving one, two, or three essential participants are designated here 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 relationship between the coding frame they select and that selected by 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*, whose definition can be formulated as follows: 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 an agent).

Break is a good example of a core transitive verb, since the lexical meaning of *break* is compatible with the highest possible degree of semantic transitivity. By contrast, *hit* or *eat* cannot be analyzed as core transitive verbs: the affected (or non-agentive) participant in a hitting event does not undergo a change of state or position, and consequently is not a typical patient; as regards eating events, the point is that the primary motivation of the action performed by the active participant in an eating event is not to change the state of the other participant or control its position, but rather to satisfy a physiological need, and consequently, the active participant in an eating event is not a typical agent.

I assume that, in the languages of the world, the set of the verbs recognizable as core transitive verbs according to this restrictive semantic definition 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.

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 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>³ is the same as that of *puskatu* ‘break’

³ On the notion of *zero case* (abbreviated as Ø in the schematization of case frames), see section 2.5.

(example (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’ (example (2)).

(1) Basque⁴

- a. *Haurr-ek ispilu-a puskatu dute.*
 child-PL.ERG mirror-SG break.CPL PRS.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) Akhvakh⁵

- a. *Mik’i-de istaka biq’wāri.*
 child-ERG glass break.CPL
 ‘The child broke the glass.’
- b. *Mik’i-la istaka harigwari.*
 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, but transitive coding is universally the default type of coding for bivalent verbs – see for example Creissels and Bassène 2013 for a detailed comparison of two languages, Say 2014 on bivalent verbs in the languages of Europe, and Haspelmath 2015 on ‘transitivity prominence’ in a worldwide sample of 36 languages.

In this chapter, 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.⁶ The verbs selecting coding frames that do not include two terms coded like the agent and the patient of core transitive verbs are designated as intransitive, regardless of the number of their (semantic) arguments.⁷

⁴ Unless otherwise stated, the Basque examples quoted in this chapter have been checked by Céline Mounole.

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

⁶ 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 chapter, 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 – 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 the framework developed in Bickel 2011 and Witzlack-Makarevich 2011, but plays no direct role here.

⁷ Note however that polysemous verbs should be characterized as transitive or intransitive in absolute terms, but rather separately for each of their possible meanings.

2.2 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*, can be illustrated by the debitive construction of Latvian analyzed by Seržant & Taperte 2016, which imposes to transitive verbs a case frame <DAT, Ø/ACC>⁸ different from the case frame <Ø, ACC> found with other TAM values (example (3)).

(3) Standard Latvian (Seržant & Taperte 2016: 200-201)

- a. *Kāpēc es šo filmu redzu?!*
 why 1SG DEM.ACC.SG film.ACC.SG see.PRS.1SG
 ‘Why do I watch this film?!’
- b. *Kāpēc man šī filma ir jā-redz?!*
 why 1SG.DAT DEM.SG.F film.SG be.PRS.1SG DEB-see
 ‘Why do I have to watch this film?!’
- c. *Kāpēc man tevi ir jā-redz?!*
 why 1SG.DAT 2SG.ACC be.PRS.1SG DEB-see
 ‘Why do I have to see you?!’

Another well-known phenomenon is the differential coding of agents or patients, i.e. variation in the coding characteristics of A or P exclusively conditioned by features inherent to the argument in question, or by its function in information structure. The differential flagging of patients (more commonly designated as ‘differential object marking’) conditioned by animacy or definiteness is particularly common. However, Iemmolo 2011 and Dalrymple & Nikolaeva 2011 argue that topicality is the crucial factor in the emergence of differential patient coding.⁹

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 provide evidence of demotion of either the agent (passivization) of the patient (antipassivization).¹⁰

⁸ In Standard Latvian, P in the debitive construction is in the Zero case with all types of NPs except for 1st and 2nd person pronouns and the reflexive pronoun, which are marked by the Accusative case.

⁹ On differential agent coding, see in particular Fauconnier & Verstraete 2014.

¹⁰ 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

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. Example (4) illustrates three verbal voice forms in Tagalog. Each clause has a *privileged argument* marked by the preposition *ang*,¹¹ and this privileged argument is the only one having access to some operations (for example, questioning). The preposition *ang* provides no indication about the semantic role of the privileged argument, but this information is given by the voice form of the verb. The other nominal terms of the clause are marked by prepositions whose choice reflects the argument structure of the verb and the semantic roles of adjuncts: *ng* (glossed CORE, used to mark adnominal possessors, but also agents and patients when they are not promoted to privileged argument), *sa* (Dative), etc. The functions of these voice alternations are quite similar to those of the alternations described in terms of passive or antipassive voices in other languages, but there is no clear asymmetry that could justify analyzing (a) as basic and (b) as passive, or (b) as basic and (a) as antipassive.

(4) Tagalog (Latrouite 2001: 123-4)

a. *Nagbigay* *ang babae* *ng liham sa kapit-babay.*

VOICE.give PRVL woman CORE letter DAT neighbour

‘The woman gave a letter to the neighbour.’

b. *Ibinigay* *ng babae* *ang liham* *sa kapit-babay.*

VOICE.give CORE woman PRVL letter DAT neighbour

‘The woman gave a letter to the neighbour.’

c. *Binigyan* *ng babae ng liham* *ang kapit-babay.*

VOICE.give CORE woman CORE letter PRVL neighbour

‘The woman gave a letter to the neighbour.’

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 should be considered as basic, and the other as a detransitivized variant. As discussed by Haude & Zúñiga 2016, this concerns in particular the direct/inverse systems that have a direct/inverse alternation for interactions between 3rd persons.

2.3 Core arguments vs. obliques

Monovalent verbs are very diverse as regards the degree of agentivity implied by the semantic role they assign to their unique argument, but in the languages of the world, the overwhelming majority of monovalent verbs divide into a small number of classes as regards the coding of their unique argument. Most of the time, there is just one major class of

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.

¹¹ This preposition is commonly designated as ‘nominative preposition’, but this term is potentially misleading, since the Tagalog system is basically different both from those for the description of which the term ‘nominative’ is traditionally used (Latin, Greek, etc.), and from those to the description of which the use of the term ‘nominative’ has been extended in more recent times (for example, Japanese, or the ‘marked-nominative’ languages of East Africa).

monovalent verbs to which almost all monovalent verbs belong, regardless of the semantic role of their unique argument. Some languages have two major classes of monovalent verbs, but languages with three or more classes of monovalent verbs including more than a handful of members each are exceptional.

Moreover, as a rule, intransitive predications (in the sense defined in section 2.1) involving non-monovalent verbs include an argument encoded like the unique argument of (a major class of) monovalent verbs.

On this basis, a notion of core argument transcending the distinction between transitive and intransitive predication can be defined as follows:

- in transitive predication, the core arguments are A (the argument encoded like the agent of core transitive verbs) and P (the argument encoded like the patient of core transitive verbs);
- in intransitive predications, the core argument is the argument whose coding coincides with that of the unique argument of (a major class of) monovalent verbs.

All the terms of verbal predicative constructions that are not core arguments will be designated as *obliques*, regardless of their status according to the *argument vs. adjunct* distinction. Terms that are analyzable semantically as arguments, but do not show the type of coding that would justify identifying them as core arguments, will be designated as *oblique arguments*.

Note that, according to this definition of core argument, (a) a language may have statistically marginal types of intransitive predication including no core argument, and (b) in a given language, the coding of the core argument in intransitive predications is not necessarily uniform, since quite a few languages have two major classes of monovalent verbs differing in the coding of their argument.

2.4 Alignment

The usual definition of ‘ergative alignment’ ($A = S \neq P$) and ‘accusative alignment’ ($P = S \neq A$) refers to properties shared by S (commonly defined as the sole argument of monovalent verbs) and one of the core terms of the basic transitive construction. This implies a more general notion of alignment whose definition can be formulated as follows:

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.

A first problem I would like to evoke here is that some uses of ‘alignment’ in the typological literature are not consistent with this definition, which results in confusion between logically independent notions. 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.

For example, Guarani (example (5)) has two sets of person markers for verbs. One of them indexes the A argument of transitive verbs and the core argument of a subclass of intransitive verbs that assign a relatively agentive role to their core argument, the other one indexes the P argument of transitive verbs and the core argument of another subclass of intransitive verbs, characterizable as assigning a relatively patientive role to their core argument. However,

transitive verbs cannot have more than one person agreement prefix, and the choice of the agreement prefix of transitive verbs is determined as follows:

- in all combinations of 1st/2nd person and 3rd person, the agreement prefix indexes the 1st/2nd person argument, whatever its semantic role (example (5a-b));
- in 2>1 combinations (2nd person A + 1st person P), the agreement prefix indexes the 1st person P, resulting in ambiguity with 3>1 combinations (example (5c));
- in 1>2 combinations (1st person A + 2nd person P), special portmanteau prefixes are used (example (5d));¹²
- if both A and P are 3rd person, the verb bears the 3rd person prefix of the agentive series.

(5) Guarani (Tonhauser 2006: 132-3)

- a. *A-hecha Juan.*
A.1SG-see Juan
'I see Juan.'
- b. *Che-hecha Juan.*
P.1SG-see Juan
'Juan sees me.'
- c. *Che-su'u-ta.*
P.1SG-bite-FUT
'You will bite me.'
or 'He/she/it/they will bite me.'
- d. *Roi-su'u-ta.*
1>2 SG -bite- FUT
'I will bite you.'

As can be seen from this example, the misnamed 'hierarchical alignment' is basically a *type of transitive coding* in which the coding characteristics of A and P are determined by their relative ranking with respect to an indexability hierarchy. It is true that this type of transitive coding raises specific problems for alignment typology, since it may be 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. This, however, is not a reason for considering this situation as a particular type of 'alignment'. In the particular case of Guarani, in spite of the hierarchical nature of indexation in the basic transitive construction, there is no difficulty in analyzing transitive-intransitive alignment as 'accusative' for one of the two classes of intransitive verbs, and 'ergative' for the other.

A second point that must be discussed before going further is the notion of S and its status in alignment typology. As observed in particular by Haspelmath 2011, S as commonly defined in alignment typology is problematic in two respects: first, the usual definition of S presupposes a uniformity in the coding of the sole argument of monovalent verbs which is not found in all languages; second, definitions of alignment types combining two primitives

¹² Interestingly, the portmanteau prefix *ro(i)-* '1>2SG' has the same form as the 1EXCL prefix of the agentive series.

referring to semantic prototypes (A and P) and a third primitive (S) whose definition does not rely on a semantic prototype is problematic in terms of logical consistency.

The solution proposed by Haspelmath 2011 is to replace S by S_U defined as the sole argument of uncontrolled change of state verbs (or ‘typical unaccusative verbs’) such as ‘die’, ‘rust’, ‘get lost’, ‘rot’, or ‘grow’. This is unquestionably an interesting solution which makes it possible to develop a typology of transitive-intransitive alignment both logically consistent and more insightful than that based on the usual definition of S.

Another possible solution relies on the observation that, from a strictly logical point of view, S is not necessary as a primitive in the definition of alignment relationships between transitive and intransitive predication. The point is that the alignment of S (either as usually defined, or as re-defined by Haspelmath) with A or P can be viewed as a particular case of *generalized transitive-intransitive alignment* defined as follows:

FOR EVERY (CODING OR BEHAVIORAL) PROPERTY THAT MAY CHARACTERIZE ARGUMENTS IN PREDICATIVE CONSTRUCTIONS, AN INTRANSITIVE CONSTRUCTION IS A-ALIGNED IF IT INCLUDES AN ARGUMENT COINCIDING WITH A FOR THE PROPERTY IN QUESTION, AND P-ALIGNED IF IT INCLUDES AN ARGUMENT COINCIDING WITH P.

This notion of generalized transitive-intransitive alignment must not be conceived as being in contradiction with the notion of restricted transitive-intransitive alignment following from Haspelmath’s replacement of S by S_U . Each of these two notions captures important aspects of alignment relationships between transitive and intransitive constructions, and combining them makes it possible to put forward interesting generalizations. For example, I am aware of no exception to the following generalization: if S_U is aligned with A in its coding properties, then intransitive constructions that do not include a term showing A-like coding are inexistent or exceptional; by contrast, languages in which S_U is aligned with P in its coding properties may have an important class of intransitive verbs assigning A-like coding to their core argument (a situation illustrated in particular by Basque). To put it somewhat differently: in split-intransitive languages (i.e. in the languages that have two non-marginal subclasses of intransitive verbs differing in their alignment properties), the monovalent change-of-state verbs may be included in a subclass of P-aligned intransitive verbs (as in Basque), or in a subclass of intransitive verbs whose core argument does not show a straightforward alignment relationship either with A or P (as in Georgian), but they cannot be included in a subclass of A-aligned intransitive verbs.

2.5 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 encompasses 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, shops signs, 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.¹³

¹³ 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 *devušk-a* ‘girl’ have

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.¹⁴ 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 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 example (8) in section 5.1.

2.6 Ergative case, ergative alignment, ergative languages

DeLancey 2004 argued 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. Although space limitations do not allow me to discuss these two points in detail here, I would like to emphasize first that the correlation between ergative alignment in argument coding as usually defined and the type of transitive coding (commonly considered ‘typically ergative’) in which unflagged P contrasts with overtly flagged A is not as strong as commonly assumed, since it does not account for the following configurations in argument coding:

- (a) the ‘marked-nominative’ type of argument flagging, with the same overt flagging for A and S contrasting with the absence of flagging for P (a type relatively rare at world level, but including the majority of the languages spoken on the African continent that have a case contrast between core arguments),
- (b) transitive constructions in which both A and P may be overtly flagged (a configuration found among others in Japanese, Tongan, Kanuri, and in some Basque varieties),
- (c) the kind of split-intransitivity found in languages like Basque (in which the same overt flagging is assigned not only to A in basic transitive coding, but also to the core argument of an important subclass of intransitive verbs),
- (d) strict ergative alignment in argument indexation combined with the absence of any flagging of core arguments at all, as in K’ichee’ and other Mayan languages,
- (e) disharmony between core argument flagging and core argument indexation (for example, ergative alignment in flagging combined with accusative alignment in indexation).

As regards disharmony between flagging and indexation, given the diachronic orientation of the present chapter, it is interesting to mention Harris and Campbell’s (1995: 257) observation that, in the changes in alignment they analyze in the Iranian languages of Pamir and in Kartvelian languages, “the pull towards consistency between subsystems [i.e. the tendency to

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.

¹⁴ See Creissels 2009 for a more detailed discussion of the shortcomings of current case terminology. Note however that in my 2009 paper, I expressed hesitation about the choice of a term for case forms of nouns that have the ability to be used in a function of pure designation, and ‘zero case’ was not among the possibilities I considered.

have the same coding of core arguments across the TAM paradigm] is stronger than that between rules [i.e. the tendency to have the same type of alignment in flagging and indexation].”

The second point I would like to emphasize in this section is that no significant correlation has been found so far between the various syntactic mechanisms that have been claimed to function on an ergative basis in some languages. Detailed studies of individual languages such as Creissels forthcoming show that the situation may be much more complex than commonly assumed even in apparently unproblematic ‘accusative’ languages,¹⁵ and the most striking thing in the literature on ‘syntactic ergativity’ is the lack of consensus between different authors analyzing the same languages (see for example Forker’s (forthcoming) criticism of Nichols’ analysis of Ingush as a syntactically ergative language).

My position is that no further progress in our understanding of the phenomena for the analysis of which the terms ‘ergative’ and ‘accusative’ have been used can be expected in so far as the terminological question is not clarified. The solution I propose is to restrict the use of ‘ergative’ and ‘accusative’ to case terminology, and to coin transparent and non-ambiguous terms for the other meanings with which these terms are used.

As regards case terminology, my proposal is to regulate the use of ‘accusative’ and ‘ergative’ as follows: if a form of nouns different from the quotation / labeling form is used to encode P (and possibly S_U) but not A, it can be labeled *accusative case*, and if a form of nouns different from the quotation / labeling form is used to encode A but neither P nor S_U, it can be labeled *ergative case*.¹⁶ Note that this definition allows using the label ‘ergative case’, not only for case forms assigned exclusively to A in the transitive construction, but also to case forms shared by A and the core argument of a substantial subclass of intransitive verbs, in split-intransitive languages such as Basque or Georgian.¹⁷

By contrast, I will avoid using ‘accusative’ and ‘ergative’ as labels for types of alignment. For example, in languages like Basque, the ergative case marks the core argument of a class of ‘unergative’ intransitive verbs 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.

This is the reason why I propose the unambiguous terms of *A alignment* and *P alignment* (defined in section 2.4 above) for the types of alignment between transitive and intransitive predications currently designated in the typological literature as accusative alignment and ergative alignment, respectively.

As already mentioned at the beginning of this section, the complexity of the relationship between types of transitive coding and types of transitive-intransitive alignment makes problematic 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,

¹⁵ To take just an example, French and Mandinka are equally unproblematic ‘accusative’ languages in core argument coding and in most of the syntactic mechanisms in which A and P behave differently, but in nominalization, Mandinka uniformly aligns A with S, whereas in French, when both A and P are expressed, P aligns with S, and A is coded differently.

¹⁶ *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 languages with TAM-driven alternations in transitive coding, for example Kurmanji Kurdish – see section 5.1, example (8).

¹⁷ As regards the so-called ‘marked-nominative’ languages, in which the same marked case form is assigned to A and to the core argument of all intransitive verbs (including S_U), contrary to the view expressed in Creissels 2009, I think now that no adaptation of any traditional term provides a satisfactory solution. The transparent term of ‘A/S case’ is a possible label for such case forms.

‘ergative vs. accusative 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.

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. A satisfactory characterization of the Basque system of core argument coding requires dissociation of the typological characteristics of the basic transitive construction (in which A in the Ergative case invariably contrasts with P in the Zero case) and those of the alignment relationship between transitive and intransitive constructions (with S_U consistently aligned with P, but also a substantial class of intransitive verbs whose core argument is aligned with A).

This is why I propose the term of *P-unmarked systems of argument coding* for systems of argument coding showing the following characteristics, commonly associated to the notion or morphological ergativity:

- (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).

Symmetrically, *A-unmarked systems of argument coding* can be defined as showing the following characteristics:

- (a) UNFLAGGED AGENTS;
- (b) FLAGGED PATIENTS;
- (c) EITHER NO INDEXATION AT ALL, OR INDEXATION OF AGENTS ONLY;
- (d) OBLIGATORY A CODING, i.e. the selection of A 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 notions of A-unmarked and P-unmarked systems of argument coding must be conceived as referring to prototypes conflating features that tend to co-occur cross-linguistically, but can nevertheless be dissociated in individual languages. In particular, ergative cases are typically found in P-unmarked systems of argument coding, but the presence of an ergative case in a language does not necessarily imply the presence of the other features of P-unmarkedness, and vice-versa.

3 The Obligatory Coding Principle

The Obligatory Coding Principle is a constraint according to which all verbal predicative constructions in a given language must include a nominal term showing a particular type of coding. Morphological accusativity / ergativity as usually defined is a particular case of this

constraint, which accounts for a cross-linguistically common type of limitation on coding frame inventories.

In coding frame inventories fully consistent with this principle, every coding frame includes a nominal term showing a given type of coding. Given the definition of A and P, this leaves two logical possibilities: in *obligatory A coding languages* (traditionally characterized as ‘morphologically accusative’), every coding frame includes a nominal term with coding properties identical to those of A in transitive coding, whereas in *obligatory P coding languages* (traditionally characterized as ‘morphologically ergative’), every coding frame includes a nominal 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 A coding to their sole argument (example (6b)), and the others assigning P coding (example (6c)).

(6) 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.3SG¹⁸
 ‘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 chapter 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 two particular types of argument coding systems: those in which the coding 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 intransitive verbs to that of A or P (as illustrated above by Guarani, cf. example (5) in section 2.4), and those with multiple transitive coding (illustrated above by Tagalog, cf. example (4) in section 2.2).

¹⁸ In Basque, the intransitive verbs that assign the Ergative case to their core argument are conjugated with the so-called transitive auxiliary, which in the basic transitive construction expresses agreement with A and P; when the transitive auxiliary combines with an intransitive verb assigning the Ergative case to its core argument, the A index expresses agreement with the unique core argument of the intransitive verb, whereas the P index takes the default value ‘3SG’.

4 Markedness reversals between the transitive construction and its variants

4.1 Introductory remarks

As already commented in section 2.2, the basic transitive construction may coexist with one or more alternative constructions of transitive verbs implying no change in the event structure. In the simple cases, in the alternative constructions, one of arguments is straightforwardly encoded like the core argument of intransitive verbs, whereas the other is either absent or encoded as an oblique. The construction is identified as passive if the participant encoded like the core argument of an intransitive 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.

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

For a proper understanding of the questions discussed in this section, the distinction between passive constructions (which encode an event structure identical to that encoded by the basic transitive construction) and anticausative or resultative constructions (which encode an event structure including no agent) is crucial.¹⁹ 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. Moreover, resultative constructions from transitive verbs are a common source not only of passives, but also of transitive perfects. 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.

It has long been observed that obligatory P coding (with the core argument of intransitive 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 typically 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

¹⁹ On resultative constructions, see Nedjalkov and Jaxontov (eds.) 1988.

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

The problem is, however, that no absolutely uncontroversial case of a global shift from obligatory A coding to obligatory P coding resulting from the reanalysis of a passive construction has been proposed so far. For example, no concrete evidence supports the widely accepted assumption of the passive origin of Basque ergativity. Passive or passive-like constructions are often mentioned in the literature as a plausible source of TAM driven alignment variations, but this is another question, which is addressed in section 5.

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 as the core argument of an intransitive verb, and the patient as 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. As already discussed by several authors (for detailed references, see Carrier 2012), 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 patient is in the Zero case, whereas the agent is in a syncretic case form used not only to flag agents in the basic transitive construction, but also in genitive function, traditionally called ‘Relative case’ (example (7a), in which the Relative case is glossed ERG according to its function in predicative constructions). In the basic transitive construction, the verb agrees with both the agent and the patient, whereas in the passive and antipassive constructions, it agrees with one argument only, which provides clear evidence of detransitivization. In the passive construction, the patient is in the Zero case, as in the basic transitive construction, but the agent is in a case form (the Ablative) distinct from that found in the basic transitive construction, and the verb agrees with the patient only (example (7b)). In the antipassive construction, the term in the Zero case is the agent; the patient is in the so-called Modal case, and the verb agrees with the agent only (example (7c)). The passive and antipassive alternations are morphologically coded on some verbs only.

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

- a. *Anguti-up arnaq kunik-taa.*
 man-ERG.SG 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 (7a) 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 agent 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 The grammaticalization of TAM and the Obligatory Coding Principle

5.1 Introductory remarks

Grammaticalization processes resulting in the emergence of new TAM forms in the inflectional paradigm of verbs are very common in the history of languages, and depending on the coding of core arguments in the source construction, they may induce TAM governed alternations in core argument coding, and sometimes also in the alignment relationship between transitive and intransitive predications.

In the grammaticalization of TAM, the source construction may be a TAM periphrasis involving nominalization of the verb and transposition of core arguments into genitival modifiers. Gildea 1992, 1998 showed that the evolution of such periphrases is a major source of TAM-driven alignment alternations in Cariban languages, and Coon 2008 argues that, in Chol (Mayan), an apparent TAM-driven alignment alternation with P-alignment in the perfective and A-alignment in the imperfective is due to the fact that the imperfective construction is a periphrasis involving a nominalized form of the verb, literally ‘[A’s V-ing P] happens’ in the transitive, and ‘[S’s V-ing] happens’ in the intransitive.

TAM periphrases analyzable as ‘raising’ constructions in which a semantic argument of the lower verb (the lexical verb) is treated syntactically as a term in the construction of the higher verb (the TAM auxiliary) are also very common cross-linguistically. For example, ‘A is engaged in V-ing P / S is engaged in V-ing’ is a common type of progressive periphrasis. Depending on the coding assigned by the higher verb to the ‘raised’ argument, the grammaticalization of such periphrases may also result in TAM-driven alternations in core argument coding, and sometimes also in TAM-driven alignment alternations.

Still another possibility is that the grammaticalization process resulting in the emergence of a new TAM form involves the reanalysis of an adjunct as a core argument.

The possible involvement of passive constructions as the source construction in evolutions leading to TAM-driven alignment alternations has often been evoked in the literature. This assumption is questionable if ‘passive’ is restricted to constructions involving no modification of the event structure and displaying the same TAM paradigm as the basic transitive construction. The ‘passive’ constructions mentioned for example by Harris & Campbell (1995: 244-5) are arguably passive-like constructions whose connection with perfectivity follows from the resultative semantics of the source construction. It is therefore debatable whether reference to passive is really necessary for a proper understanding of the evolutions in question, or perhaps the really relevant notion is rather resultativity.²⁰

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 modify 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 systems in which a TAM-driven alternation in the coding properties of A and P has no equivalent in intransitive predications, which results in TAM-driven alignment alternation. 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.

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

There is a consensus on the fact that the grammaticalization of new forms of perfects is a major source of TAM-driven alignment variations. Iranian languages provide a classical illustration of this type of change. At some stage in their history, the grammaticalization of a new form of perfect (which will be examined in more detail in section 5.3) resulted in the emergence of a TAM-driven alternation in the coding of A and P still found in some Iranian languages, for example Standard Kurmanji Kurdish (example (8)). Since no such alternation developed in the coding of the core argument of intransitive verbs (as illustrated in example (8) by the intransitive verb *hatin* ‘come’), the coding of both A and P in Standard Kurmanji Kurdish is characterized by a TAM-driven alternation between a coding identical to that of the core argument of intransitive verbs (Zero case + indexation on the verb form) and an oblique-like coding: in (8a-b), A is in the Zero case and is indexed on the verb form, like the core argument in intransitive predication, whereas P is in the Oblique case and is not indexed; in (8e-f), the term coded like the core argument in intransitive predication is P, and A shows the same oblique-like coding as P in (8a-b).

In other words, in Standard Kurmanji Kurdish, the alignment relationship between transitive and intransitive constructions is $A = S \neq P$ in the tenses that trigger the coding of A and P illustrated in (8a-b), $P = S \neq A$ in those that trigger the coding of A and P illustrated in

²⁰ 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 Tiriyó 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.

(8e-f), and this alternation in alignment originates in the grammaticalization of a new form of perfect in the history of Iranian languages.

(8) Kurmanji (Blau and Barak 1999)

- a. *Ez Sînem-ê dibîn-im.*
1SG Sinem-OBL see.ICPL-1SG
'I see Sinem.'
- b. *Sînem min dibîn-e.*
Sinem 1SG.OBL see.ICPL-3SG
'Sinem sees me.'
- c. *Ez tê-m.*
1SG come.ICPL-1SG
'I am coming.'
- d. *Sînem tê-Ø.*
Sinem come.ICPL-3SG
'Sinem is coming.'
- e. *Min Sînem dît-Ø.*
1SG.OBL Sinem see.CPL-3SG
'I saw Sinem.'
- f. *Sînem-ê ez dît-im.*
Sinem-OBL 1SG see.CPL-1SG
'Sinem saw me.'
- g. *Ez hat-im.*
1SG come.CPL-1SG
'I came.'
- h. *Sînem hat-Ø.*
Sinem come.CPL-3SG
'Sinem came.'

In such a system, the coding of the core argument of intransitive verbs cannot be identified globally with that of either A or P, which constitutes a radical violation of the Obligatory Coding Principle.

Two remarks are in order at this point.

First, it must be emphasized that the grammaticalization of TAM periphrases into new TAM forms does not necessarily trigger changes in the alignment relationship between transitive and intransitive constructions, even if they affect the coding characteristics of A and P in the basic transitive construction. For example, in Latvian, the grammaticalization of the debitive construction briefly presented in section 2.2 resulted in an alternation between $\langle \emptyset, \text{ACC} \rangle$ and $\langle \text{DAT}, \emptyset/\text{ACC} \rangle$ in transitive coding, but this change did not affect the status of Latvian as an obligatory A coding language, since transitive A's and core arguments of intransitive verbs are treated in the same way in the debitive construction.

Another important observation is that many languages in which the grammaticalization of a new TAM form resulted in a violation of the Obligatory Coding Principle similar to that illustrated by Kurmanji Kurdish 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. In some cases, 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, whereas in others, a coding alternation that initially developed in the transitive construction was subsequently extended to intransitive coding.

The first variant of this regularization process (suppression of the coding alternation that had developed in the transitive construction) occurred in many Iranian languages (for example, Persian) which at some point in their history had an argument coding system similar to that of Kurmanji Kurdish (example (8)), but subsequently aligned the coding of A and P in all tenses with the type of coding found in the present, which re-established a situation characterizable in terms of obligatory A coding.

As regards the second possible variant of the regularization process (by which a coding alternation that initially developed in the transitive construction extends to the coding of the core argument of intransitive verbs), Seržant 2012 shows that such a process occurred in the history of the North Russian Perfect: after the creation of a transitive perfect construction with adessive marking of the agent and no indexation of either A or P (example (9a)), North Russian has extended the adessive marking to core arguments of intransitive verbs in the Perfect (example (9b)).

(9) 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 across 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 overt flagging is used for A and for the core argument of intransitive verbs.

Among the languages with TAM-driven alternations between A alignment and P alignment, the configuration found in Iranian languages and illustrated above by Kurmanji Kurdish, with P alignment in past tense or perfective aspect, is particularly widespread cross-linguistically, 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.

5.3 Split-alignment resulting from the grammaticalization of a new perfect form: problems in reconstructing the scenario

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 if 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, and if the notion of passive is broadened to include various types of passive-like constructions, then the question

is: what is the exact nature of the passive-like constructions whose reanalysis may lead to the emergence of new perfect forms, and what are the possible scenarios. This is a tricky question, since for the languages in which this change is historically attested (the Indo-Iranian languages), the interpretation of the historical data is far from obvious, hence the long-standing controversy about the emergence of split-alignment patterns in the languages in question.

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 this evolution 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 (example (10)).

(10) 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.1SG letter.ACC written.ACC
‘I am concerned by the fact that a letter is written.’
(lit. ‘I have (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 article, 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 decisive 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 resultative situation is encoded as an adjunct added to the resultative clause, as in example (11b). This construction cannot be described as a possessive clause with an embedded resultative clause, since it does not involve the verb *haben* ‘have’ standardly used to express 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 of the type illustrated in (10c).

(11) 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.’)

The plausibility of the reanalysis of a dative-marked adjunct referring to an indirectly affected participant as an agent showing argumental properties is confirmed by the observation of the possible interpretations of the Spanish construction in which anticausative *se* combines with a dative NP. The original meaning of this construction is clearly that the referent of the dative NP is indirectly affected by an event occurring more or less spontaneously, but in (12), the dative NP can only be interpreted as representing an involuntary agent, and the infinitive can only be interpreted as controlled by the dative NP, which points to the reanalysis of a former adjunct as an argument.

(12) Spanish (Rivero 2008: 221)

- A Ana se le quemaron las niñas al bañar-las*
 Ann.DAT 3REFL 3SG.DAT burned.3PL the girls at.the bathe.INF-3PL.ACC
 ‘Ann (accidentally) burned the girls when bathing them.’

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.²¹ 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

²¹ 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 core argument of intransitive verbs in the same tense.

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 (11b) is reanalyzed as a transitive construction with a new tense form expressing perfect semantics. If the same resultative construction is available with intransitive verbs (as in English *The man is gone* / *The mirror is broken*), and if the resultative construction of intransitive verbs undergoes the same reanalysis as a perfect without any change in its form, the construction of the perfect form resulting from this reanalysis will be something like (intr.) *The man is gone* / (tr.) *By the child the mirror is broken*, in contradiction with the rule of obligatory A coding.

Another scenario likely to lead to the same result is the reanalysis of periphrases in which a resultative form of the verb is used nominally with a genitival modifier representing the agent, i.e. something like literally ‘P is A’s V-ed’. Similarities between the coding of agents and that of possessors can be viewed as evidence supporting the reconstruction of this type of scenario. Creissels (1979: 523-529) discusses this hypothesis for Hungarian and K’ichee’ (Maya). Such an evolution does not necessarily lead to the emergence of a perfect with alignment properties different from those of the pre-existing TAM, but depending on the details of argument and possessor marking in the language, this is unquestionably a possibility.

5.4 Progressive periphrases and split alignment

Cross-linguistically, 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 be compared with *Mary is [in the garden]*), or Spanish *María está [comprando regalos para los niños]* (to be compared 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, and consequently, to a systematic violation of the Obligatory Coding Principle.

Non-verbal predications generally involve an argument encoded like the core argument of intransitive verbs. Consequently, in languages in which A alignment is canonical, the A argument of a transitive verb treated as the core argument of an intransitive predication 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 then treated as the core argument of an intransitive predication.

This can be illustrated by the Basque progressive periphrasis in which the intransitive compound verb *ari izan* ‘be engaged in’ combines with nominal complements marked typically locative (example (13a)), or with clausal complements headed by the so-called ‘Incomplete Participle’, used also to form the non-periphrastic Present of the verbs that do

not have synthetic finite forms (example (13b) and (13d)). The construction with a clausal complement is a raising construction in which the A/S argument of the auxiliated verb is uniformly treated as the core argument of *ari izan*. Since *ari izan* is an intransitive verb showing P alignment, with transitive verbs (and also with the intransitive verbs that assign Ergative coding to their core argument), this results in coding characteristics different from those of the same argument in non-periphrastic constructions (example (9d-e)).

(13) Basque

- a. *Jon lanean ari da.*
 Jon work.SG.LOC engaged be.PRS.3SG
 ‘Jon is working.’ (lit. ‘Jon is engaged in work.’)
- b. *Jon paseatzen da.*
 Jon walk.ICPL PRS.3SG²²
 ‘Jon is walking.’ (non-periphrastic Present)
- c. *Jon [[paseatzen] ari] da.*
 Jon walk.ICPL engaged be.PRS.3SG
 ‘Jon is walking.’ (progressive periphrasis)
- d. *Jonek berriak ikusten ditu.*
 Jon.ERG news.PL see.ICPL PRS.3SG.3PL
 ‘Jon is watching the news.’ (non-periphrastic Present)
- e. *Jon [[berriak ikusten] ari] da.*
 Jon news.PL see.ICPL engaged be.PRS.3SG
 ‘Jon is 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 TAM-governed variation in the coding properties of the transitive construction and in alignment, with uniform A alignment in a Present or Progressive tense contrasting with split alignment in other tenses.

A similar process whereby a periphrasis ‘be occupied with’ has given rise to A alignment in the progressive is discussed by Gildea 1998 (chapter 12).

As regards the progressive periphrasis of Basque, it is nevertheless interesting to observe that 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, (14a) and (14b) are two possible versions of a Basque sentence meaning ‘The companies are preparing the future managers’. The (a) version, with the A argument of the auxiliated verb in the Zero case and the P argument not indexed, is the correct

²² 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’.

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, pers.com.).

(14) Basque

- a. *Enpresak etorkizuneko zuzendariak prestatzen ari dira.*
 company.PL future manager.PL prepare.ICPL engaged be.PRS.3PL
 ‘The companies are preparing the future managers.’
- b. *Enpresek etorkizuneko zuzendariak prestatzen ari dituzte.*
 company.PL.ERG future manager.PL prepare.ICPL engaged be.PRS.3PL.3PL
 same meaning as (a)

The tendency to eliminate the alternation in the coding properties of the transitive construction resulting from the grammaticalization of the progressive periphrasis may be reinforced by the fact that, originally, the development of this periphrasis was limited to some dialectal varieties of Basque (Joseba Lakarra, pers.com.). 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 be particularly prone to aligning it with the canonical transitive pattern.

The so-called ‘bi-absolutive construction’ of transitive verbs in Nakh-Daghestanian languages is another example of a progressive periphrasis whose grammaticalization may result in a systematic violation of the Obligatory Coding Principle in obligatory P coding 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). The bi-absolutive construction is a periphrasis expressing progressive aspect or present tense in which 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. In intransitive predication (example (15a)), the core argument is invariably in the Zero case and invariably controls the agreement of both the copula and the auxiliated verb. 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 example (15b), 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 example (15c), 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).

(15) Avar (Alekseev and Ataev 1997)²³

- a. *Emen w-ač'ule-w w-ugo.*
 father SG.M-coming-SG.M SG.M-COP
 'Father is coming.'
- b. *Insu-ca χur b-el'ule-b b-ugo.*
 father-ERG field SG.N-plowing-SG.N SG.N-COP
 'Father is plowing the field.'
- c. *Emen [χur b-el'ule-w] w-ugo.*
 father field SG.N-plowing-SG.M SG.M-COP
 'Father is plowing the field.'

The construction illustrated by example (15c) 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 (15c) maintains the biclausal structure of the original periphrasis, whereas in (15b), 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 example (15b) can be interpreted as resulting from the elimination of the violation of the Obligatory Coding Principle introduced by the grammaticalization of the progressive periphrasis illustrated by example (15c).

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

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 to the connection between perfective and ergativity assumed by Dixon have been found, in particular among Cariban languages (see in particular Gildea 1992, 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 it is not difficult to find languages having TAM periphrases whose grammaticalization, in the absence of a readjustment under the pressure of analogy, would automatically give rise to alignment variations contradicting the universal posited by Dixon.

²³ Note that the Avar noun for 'father' has two suppletive stems: *emen* in the Zero case, and *insu* in other cases.

For example, in addition to the progressive periphrasis analyzed in section 5.4, Basque has several TAM periphrases whose grammaticalization could lead to the emergence of various alternations in the coding properties of the transitive construction and in alignment patterns.

A first example is the future periphrasis in which *joan* ‘go’ combines with the allative form of a verbal noun. Since *joan* is an intransitive verb showing P alignment, in this future periphrasis, the A/S argument of the auxiliated verb encoded as the core argument of *joan* is uniformly in the Zero case; if the auxiliated verb is transitive, P is also in the Zero case, and it is not indexed, since the non-finite verb forms of Basque do not express agreement with any of their arguments (example (16b)).

(16) Basque

- a. *Jonek berriak ikusiko ditu.*
 Jon.ERG news.PL see.FUT PRS.3SG.3PL
 ‘Jon will watch the news.’ (non-periphrastic future)
- b. *Jon [berriak ikustera] doa.*
 Jon news.PL see.NMLZ.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 the emergence of a new future form imposing the same coding (with in particular zero flagging) to A and the core argument of all intransitive verbs, and consequently uniform A alignment contrasting with the split alignment pattern found with the other TAM forms.

Another case in point is the debitive periphrasis with *behar izan* ‘need’, lit. ‘have need’. This transitive compound verb can take a nominal complement in the Zero case, as in example (17a), but it is also found in a debitive periphrasis in which it combines with the completive participle of the auxiliated verb (example (17c) and (17e)).²⁴ The behavior of this 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 chapter 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 to its core argument, as in example (17c).

(17) Basque

- a. *Jonek kotxe berri bat behar du.*
 Jon.ERG car new one need have.PRS.3SG.3SG
 ‘Jon needs a new car.’

²⁴ 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 example (17a), *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 etxera doa.*
 Jon car new
 ‘Jon is going home.’
- c. *Jonek [etxe-ra joan] behar du.*
 Jon.ERG house.SG.ALL go.CPL need have.PRS.3SG.3SG
 ‘Jon must go home.’
- d. *Jonek ogia erosten du.*
 Jon.ERG bread.SG buy.ICPL PRS.3SG.3SG
 ‘Jon is buying bread.’
- e. *Jonek [ogia erosi] behar du.*
 Jon.ERG bread.SG buy.CPL need have.PRS.3SG.3SG
 ‘Jon must go home.’

Cross-linguistically, debitive constructions are a common source of future tenses. In Basque, in the absence of a readjustment, the grammaticalization of this periphrasis as the usual expression of future might lead again to the emergence of a tense form imposing uniform A alignment, but with the atypical variety of flagging in which the same marked case form is assigned to the A argument of transitive verbs and to the core argument of all intransitive verbs.

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* in an independent clause (18a), and as *jin behar den* in a relative clause (18b).

(18) Old Lapurdian (Mounole 2011: 191)

- a. *Jin behar du.*
 come. CPL need have.PRS.3SG.3SG
 ‘He will come.’
- b. *jin behar den*
 come. CPL need be.PRS.3SG.REL
 ‘who will come’ (relative clause)

In (18a), in conformity with the etymology of this periphrasis, the auxiliary *du* is a form of the transitive auxiliary ‘have’, in spite of the intransitive nature of the auxiliated verb, whereas in (18b), the auxiliary *den* is a dependent form of the intransitive auxiliary ‘be’,²⁵ which points to a possible readjustment in the coding properties of this periphrasis when the auxiliated verb was intransitive.

²⁵ The corresponding independent form would be *da*; in the same context, the transitive auxiliary would occur as *duen*.

We now turn to a resultative periphrasis of Basque in which the verb *izan* ‘be’²⁶ combines with the completive participle in the definite form treated as an adjectival predicate.²⁷

Contrary to Indo-European ‘past participles’, this non-finite 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, as in (19b)) and antipassive (if the raised term is A, as in (19c)).

(19) 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’
 (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.’
 (lit. ‘Jon is [(having) written many letters].’)

In the ‘passive’ variant (19b), there is no change in case assignment in comparison with the non-periphrastic construction of a transitive verb, since the core argument of *izan* ‘be’ has the same coding characteristics as the P argument of transitive verbs. 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 (19c) could lead to a TAM-driven alignment alternation with uniform 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 an alternation in alignment in clear contradiction with Dixon’s universal.

5.6 Concluding remarks

The data examined in this section show that the grammaticalization of TAM constitutes a potential source of a wide variety of alternations in the coding characteristics of the transitive construction, which may mechanically induce violations of the Obligatory Coding Principle, depending on the treatment of intransitive verbs in the source construction. But we also have seen that this automatic result of the grammaticalization of some TAM periphrases tends to be canceled by subsequent evolutions that do not necessarily align the coding frame imposed by

²⁶ 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.

²⁷ In most dialects of Basque, nouns and adjectives in predicate function are in the definite form.

the new TAM form with that of the pre-existing TAM forms, but at least modify it in such a way as to eliminate the alignment variation.

Consequently, there is no need to look for direct semantic / functional explanations of the fact that just a few TAM-driven alignment variations are well-attested cross-linguistically, while others are marginal or not attested at all. The fact that a particular TAM grammaticalization process at some point in the history of a given language has a side effect on alignment or not depends entirely on the coding of core arguments in the source construction, and cross-linguistically, verb forms expressing the same type of TAM value can emerge from the grammaticalization of a variety of source constructions with various types of configurations in terms of core argument coding. Given the strong tendency to eliminate the violations of the Obligatory Coding Principle resulting from the grammaticalization of TAM, the only TAM-driven alignment variations that have a relatively good chance to surface again and again in different languages are those likely to result from particularly common grammaticalization paths. The explanation of the relative frequency of some particular TAM-driven alignment variations must therefore not be sought in the semantics of TAM forms. The real question is why some types of evolutions leading to the emergence of new TAM forms are more common than others, and this question has no direct link with alignment typology.

6 Conventionalization of argument 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 anaphoric 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 do not have this possibility (for example, Nahuatl systematically uses ‘indeterminate object prefixes’ (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 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 intransitive verbs include a term in the Zero case governing verb agreement in gender and number like P in the transitive construction (example (19)).

(20) Akhvakh

- a. *Ek'wa w-oq'-ilo.*
man SG.M-come-CPL.NEG.SG.M
'The man did not come.'
- b. *Jaše j-eq'-ile.*
girl SG.F-come-CPL.NEG.SG.F
'The girl did not come.'
- c. *Ek'waš^w-e jaše j-ič'-ile.*
man-ERG girl SG.F-push-CPL.NEG.SG.F
'The man did not push the girl.'

Depending on a complex combination of grammatical and lexical factors, verb agreement in gender and number is not always apparent.²⁸ 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 anaphoric 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 that could be represented by an NP in the Zero case. With respect to agreement, the verbs in question show neuter singular default agreement (example (20)).

²⁸ 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 example (19) (*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.

(21) Akhvakh

- a. *Ek'waš^w-e jašo-ga eḡ-ere godi.*
 man-ERG girl-ALL look_at-PROG COP.SG.N
 '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'warurula* is commonly encountered with the meaning 'hit' in a construction superficially similar to those illustrated by example (21), with an argument in the Ergative case and an argument in the Locative case (example (22)).

(22) Akhvakh

- Ek'waš^w-e jašo-ge l̄'war-ari.*
 man-ERG girl-LOC hit-CPL
 'The man hit the girl.'

At first sight, *l̄'warurula* might appear as a bivalent verb with an exceptional valency pattern, but in fact, it is a trivalent verb, and (22) 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 (example (23)).

(23) Akhvakh

- a. *Ek'waš^w-e jašo-ge rel'a l̄'war-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'waš^w-e jašo-ge č'uli l̄'war-ari.*
 man-ERG girl-LOC stick hit-CPL
 'The man hit the girl with a stick.'
 (lit. 'applied a stick to the girl')

Consequently, the basic meaning of *Ēʷarurula* 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 (22) 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

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’²⁹ constructions) as the unique core argument of intransitive constructions can be a source of systems in which, in violation of the Obligatory Coding Principle, intransitive verbs divide into two subsets differing in the alignment of their core 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 core 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 the evolution of transimpersonal constructions towards plain intransitive constructions with a non-canonical alignment pattern.

Amharic is a language with obligatory A coding in which A is obligatorily indexed by person markers which, in the absence of a co-referent NP, normally trigger an anaphoric 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 anaphoric 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 (example (24a)). 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 (example (24b)).

(24) Amharic (Leslau 2005: 43)

- a. *Rabä-ñ.*
 hunger.CPL.3SG.M-1SG
 ‘I am hungry.’ (lit. ‘It hungered me.’)

²⁹ According to Malchukov 2008, this term was coined by Mary Haas (Haas 1941).

- b. *Īnjära rabä-ñ.*
 bread hunger.CPL.3SG.M-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 example (24b) 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.

The impersonal construction of the Russian verb *trjasti* ‘shake’ results from an evolution of this type. This verb occurs in a canonical transitive construction (example (25a)), but also in an impersonal construction that, synchronically, cannot be analyzed as an elliptical variant of the transitive construction, since the participant expressed as the A term of the transitive construction can be encoded as an oblique introduced by the preposition *ot* ‘from’. In example (25b), 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.

(25) 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 (25b) 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 example (26), and the non-verbal element is a noun encoded like the P argument of transitive verbs (Samvelian 2012: 16).

(26) Basque

- a. *Haurr-ek lo egiten dute.*
 child-PL.ERG sleep do.ICPL 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.ICPL
 ‘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 univerbation 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 example (26) 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 *bultza egin* lit. ‘do impulse’ / *bultzatu* ‘push’ (example (27)).³⁰

(27) Basque

- a. *Mutil-ak ate-ari bultza egin zion.* <ERG, DAT, \emptyset >
 boy-SG.ERG door-SG.DAT impulse do.CPL PST.3SG.3SG.3SG
 ‘The boy pushed the door.’

³⁰ *Bultzatu* is the completive participle, used in Basque grammars and dictionaries as the quotation form of verbs. It can be decomposed as *bultza-* (root) plus *-tu* (completive aspect marker).

- b. *Mutil-ak ate-a bultzatu zuen.* <ERG, Ø>
 boy-SG.ERG door-SG push.CPL PST.3SG.3SG
 same meaning as (a)

In this example, a light verb compound with the coding frame <ERG, DAT, Ø> corresponds to a simplex transitive verb with the coding frame <ERG, Ø>, and the Dative argument of the light verb compound represents the same participant as the argument of the simplex verb in the Zero case. However, in such situations, the prevailing trend in Basque is rather to encode the arguments of the simplex verb 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 simplex verb *dirdiratu* equally assign the Ergative case to their argument (example (28)). In other words, *dirdiratu* belongs to the class of verbs with no argument in the Zero case and an argument in the Ergative case, which included very few verbs in Old Basque but has grown dramatically in the history of most Basque varieties.

(28) Basque

- a. *Eguzki-ak dirdir egiten du.* <ERG, Ø>
 sun-SG.ERG shining do.ICPL PRS.3SG.3SG
 ‘The sun is shining.’
- b. *Eguzki-ak dirdirutzen du.* <ERG>
 sun-SG.ERG shine.ICPL PRS.3SG.3SG
 same meaning as (a)

The variation observed in the coding frames of Basque 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 example (28), and a tendency to organize the coding frame of the simplex verb according to the principle of obligatory P coding, as in example (27).

In Creissels 2015, I argued that the prevalence 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 intransitive verbs borrowed into Basque assign Ergative coding to their core argument, irrespective of their valency.

In the same article, I compared 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 clearly resulting from the univerbation of a light verb compound, and none of the verbs in question is used with a non-canonical coding frame in all Andic languages.

For example, in the case of ‘listen’, three situations are found among Andic languages:

- Some Andic languages express ‘listen’ by means of a construction involving the noun ‘ear’ in addition to the noun phrases encoding the two participants, as in Godoberi *hāt’uk’ja rikī*, literally ‘fix the ear (on someone/something)’ (example (29)). Formally, this construction is an instance of the regular coding frame <ERG, Ø, ALL> with *hāt’uk’ja* ‘ear’ in the Zero case.
- Others have a verb ‘hear’ with the exceptional coding frame <ERG, ALL>: Tindi *aniḡiḡa* (example (30)), Chamalal *woḡuk’la*, Bagvalal *aḡtila*;
- A verb ‘hear’ with the regular coding frame <Ø, ALL> is found in two Andic languages: Akhvakh *hādaḡurula* (example (31)), Karata *ādukala*.

(29) Godoberi (Saidova 2006)

Waḡu-di imu-qī hāt’uk’ja rikki rukkida.
 son-ERG father-LOC/ALL ear hold.INF must.ICPL
 ‘The son must listen to his father.’

(30) Tindi (Magomedova 2003)

Di-qā aniḡiḡa hik’i oḡ^w-ī
 1SG-LOC/ALL listen.ICPL NEG DEM.M-ERG
 ‘He does not listen to me.’

(31) Akhvakh

Waḡa imo-ga hādaḡari.
 boy father-ALL listen.CPL
 ‘The boy listened to his father.’

Interestingly, ‘fix the ear on’ is the obvious etymology of Akhvakh *hādaḡurula* (compare with *hāde* ‘ear’, *biḡurula* ‘fix’ – the root of this verb is *-iḡ-*), in spite of the fact that the NP representing the listener is assigned the Zero case instead of the Ergative case that should be expected from this etymology.

The variation in the expression of ‘listen’ in Andic languages provides therefore evidence supporting the reconstruction of the following evolution:

- at a first stage, the coalescence of a trivalent verb selecting the regular frame <ERG, Ø, ALL> with a noun in the Zero case creates a bivalent verb with the exceptional coding frame <ERG, ALL>;
- at a second stage, attested by Akhvakh and Karata, the exceptional coding frame resulting from this evolution may be regularized into <Ø, ALL>.

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 argued 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 2008.³¹ 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 consequently 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 chapter, 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 results in TAM-driven alternations in core argument coding that affect the uniformity of alignment relationships across the TAM paradigm. 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 new TAM forms, 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 alignment, or with two substantial classes of intransitive verbs differing in the coding properties of their core 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 alignment alternations triggered by the grammaticalization of new TAM forms with a special coding of A and P tend to be eliminated by the alignment of the construction of newly created TAM forms with the pre-existing ones, or by a change in the coding of the core argument of intransitive verbs restoring the alignment relationship found in the remainder of the TAM paradigm,

³¹ In languages with strict ergative coding, ergative coding tends to be limited to agents of core transitive verbs in contexts implying a high degree of semantic transitivity, whereas in languages with loose ergative coding, ergative coding is widely used to encode the most agent-like argument of bivalent verbs that are not core transitive verbs, irrespective of the precise semantic roles they assign and of the contexts in which they are used.

- 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 by 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 with very few exceptions (mainly Kartvelian and Indo-Iranian languages), 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.

As rightly pointed to me by Sonia Cristofaro, this is in fact a general problem in diachronic typology: there are some relatively well-described cases, and plausible scenarios for the development of particular synchronic patterns can be put forward, but we don't have the quantitative evidence about the actual frequency of the relevant processes that would allow the assessment of hypotheses about how frequent these scenarios are cross-linguistically.

Acknowledgments

Sincere thanks are due to the editors of this volume as well as to three reviewers for suggestions that helped me to improve the article considerably. I am also very grateful to Martin Haspelmath and Iija Seržant for their valuable comments. The usual disclaimer applies.

Abbreviations

A	agent	M	masculine
ABL	ablative	MOD	modal
ACC	accusative	N	neuter
ALL	allative	NEG	negation
ANTIP	antipassive	NMLZ	nominalization
COP	copula	OBL	oblique case
CPL	completive aspect	Ø	zero case
DAT	dative	P	patient
DEB	debitive	PASS	passive
DEF	definite	PL	pluriel
DEM	demonstrative	PP	past participle
ERG	ergative	PREV	preverb
F	feminine	PRF	perfect
FUT	future	PROG	progressive
GEN	genitive	PRS	present
ICPL	incompletive aspect	PRVL	privileged argument
INF	infinitive	PST	past
INSTR	instrumental	REL	relativizer
LOC	locative	S	sole argument of (a non-marginal

	subclass of) monovalent verbs	TAM	tense-aspect-modality
SG	singular	V	verb
S _U	sole argument of uncontrolled change of state verbs	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*, 197-218. Oxford: Oxford University Press.
- 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.
- Bickel, Balthasar. 2011. Grammatical relations typology. In Jae Jung Song, ed. *The Oxford handbook of linguistic typology*, 399-444. Oxford: Oxford University Press.
- 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.
- Coon, Jessica. 2008. When ergative = genitive: Nominals and split ergativity. In Natasha Abner and Jason Bishop (eds.), *WCCFL XXVII: Proceedings of the 27th West Coast Conference on Formal Linguistics*, 99-107. Somerville, MA: Cascadilla Proceedings Project.
- Creissels, Denis. 1979. *Les constructions dites possessives, étude de linguistique générale et de typologie linguistique*. Habilitation thesis (thèse d'état). University of Paris IV.
- 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. 2015. Univerbation of light verb compounds and the Obligatory Coding Principle. In Léa Nash and Pollet Samvélian (eds.), *Approaches to complex predicates*, 46-69. Brill.
- Creissels, Denis. Forthcoming. Grammatical relations in Mandinka. In Balthasar Bickel & Alena Witzlack-Makarevich (eds.) *Handbook of grammatical relations*. John Benjamins.
- Creissels, Denis and Alain-Christian Bassène. 2013. Valency patterns for bivalent verbs in two West African languages: Mandinka (Mande) and Jóola Banjál (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).
- Dalrymple, Mary and Irina Nikolaeva. 2011. *Objects and information structure*. Cambridge University Press.
- DeLancey, Scott. 2004. The blue bird of ergativity. In Queixalós, Francesc (ed.), *Ergativity in Amazonia III*, 1-15. 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.
- Ettxepare, Ricardo. 2003. Valency and argument structure in the Basque verb. In José Ignacio Hualde and Jon Ortiz de Urbina (eds.), *A grammar of Basque*, 363-426. Mouton de Gruyter.
- Evans, Nicholas. 2004. Experiencer object in Iwaidjan languages. In Bhaskararao, P. and K.V. Subbarao (eds.), *Non-nominative subjects*. vol. 1, 169-193. Amsterdam and Philadelphia: John Benjamins.
- Fauconnier, Stefanie and Jean-Christophe Verstraete. 2014. A and O as each other's mirror image? Problems with markedness reversal. *Linguistic typology* 18(1), 3-49.
- Forker, Diana. 2012. The bi-absolutive construction in Nakh-Daghestanian. *Folia Linguistica* 46(1), 75-108.
- Forker, Diana. Forthcoming. Ergativity in Nakh-Daghestanian languages. In Jessica Coon, Diane Massam, & Lisa Travis (eds.), *The Oxford Handbook of Ergativity*. Oxford University Press.
- Garrett, Andrew. 1990. The origin of NP split ergativity. *Language* 66, 261-296.
- Gildea, Spike. 1992. *Comparative Cariban morphosyntax: on the genesis of ergativity in independent clauses*. Oxford: Oxford University Press.
- Gildea, Spike. 1997. Evolution of grammatical relations in Cariban: How functional motivation precedes syntactic change. In T. Givón (ed.), *Grammatical relations: a functionalist perspective*, 155-198. Amsterdam and Philadelphia: John Benjamins.
- 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.
- Haspelmath, Martin. 2011. On S, A, P, T, and R as comparative concepts for alignment typology. *Linguistic Typology* 15(3), 535-567.
- Haspelmath, Martin. 2015. Transitivity prominence. In Andrej Malchukov and Bernard Comrie (eds.), *Valency Classes in the World's Languages*, Vol. 1, 131-147. Berlin: De Gruyter Mouton.
- Haude, Katharina and Fernando Zúñiga. 2016. Inverse systems and symmetrical voice: a unified view on languages with two transitive constructions. *Linguistics* 54(3), 443-481.
- 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*, 252-276. Oxford: Oxford University Press.
- Hualde, José Ignacio and Jon Ortiz de Urbina (eds.). 2003. *A grammar of Basque*. Mouton de Gruyter.

- Iemmolo, Giorgio. 2011. *Towards a typological study of differential object marking and differential object indexation*. Doctoral dissertation. Pavia: Università degli studi di Pavia.
- Janic, Katarzyna. 2013. *L'antipassif dans les langues accusatives*. Doctoral dissertation. 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.
- Latrouite, Anja. 2011. *Voice and case in Tagalog: the coding of prominence and orientation*. Doctoral dissertation. Düsseldorf: Heinrich-Heine Universität.
- 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*, 76-100. Oxford: Oxford University Press.
- Magomedova, Patimat. 2003. *Tindinsko-russkij slovar'* [Tindi-Russian dictionary]. Maxačkala: Dagestanskij Naučnyj Centr Rossiskoj Akademii Nauk.
- 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*, 297-333 Oxford: Oxford University Press.
- Mounole, Céline. 2011. *Le verbe basque ancien : étude philologique et diachronique*. Doctoral dissertation. 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*, 787-806. Bilbao: Diputacion Foral de Gipuzkoa - Gipuzkoako Foru Aldundia and Universidad de Pais Vasco - Euskal Herriko Unibertsitatea, Bilbao.
- 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*, 709-724. Universidad del País Vasco - Euskal Herriko Unibertsitatea.
- Rivero, María Luisa. 2008. Oblique subjects and person restrictions in Spanish. In Roberta D'Alessandro, Susann Fischer, and Gunnar Hrafn Hrafnbjargarson (eds.), *Agreement restrictions*. 215-250. Mouton de Gruyter.

- Saidova, Patimat. 2006. *Godoberinsko-russkij slovar'* [Godoberi-Russian dictionary]. Maxačkala: Dagestanskij Naučnyj Centr Rossiskoj Akademii Nauk.
- Samvelian, Pollet. 2012. *Grammaire des prédicats complexes: Les constructions nom-verbe*. Paris: Lavoisier.
- Say, Sergey. 2014. Bivalent verb classes in the languages of Europe. A quantitative typological study. *Language Dynamics and Change* 4(1), 116-166.
- 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.
- Seržant, Ilja A. & Jana Taperte. 2016. Differential argument marking with the Latvian debitive. In Axel Holvoet and Nicole Nau (eds.), *Argument realization in Baltic*, 199-258. John Benjamins.
- 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*, 215-225. University of British Columbia: UBC Working Papers in Linguistics.
- Tonhauser, Judith. 2006. *The temporal interpretation of noun phrases: evidence from Guaraní*. Doctoral dissertation. Stanford University.
- Tsunoda, Tasaku. 1985. Remarks on transitivity. *Journal of Linguistics* 21(2), 385-396.
- Witzlack-Makarevich, Alena. 2011. *Typological Variation in Grammatical Relations*. Doctoral dissertation. University of Leipzig.