

**Abstract title: The Obligatory Coding Principle (alias Obligatory Case Parameter) in diachronic perspective**

**Abstract category: oral**

**Theme session: typological hierarchies in synchrony and diachrony**

In the recent typological literature (Dixon 1994 and others), accusativity / ergativity is defined in terms of  $S=A \neq P$  vs.  $S=P \neq A$  alignment, but morphological accusativity / ergativity can be viewed as a particular case of a more general principle underlying the organization of verbal valency in languages that have consistent  $S=A \neq P$  or  $S=P \neq A$  alignment, the Obligatory Coding Principle. According to this principle, regardless of the number of arguments, the only available coding frames are, either (in 'accusative' languages) those including a term with coding properties identical to those of A, or (in 'ergative' languages) those including a term with coding properties identical to those of P. A formal elaboration of this principle can be found in the generative literature under the name of Obligatory Case Parameter.

I would like to present a paper discussing a particular type of diachronic process that may be responsible for the development of coding frames contradicting the Obligatory Coding Principle in languages with (quasi-)obligatory P-like coding: the univerbation of light verb compounds.

Some languages have a very high proportion of predicates expressed by means of light verb compounds whose non-verbal element is a noun encoded like the P argument of typical transitive verbs, and diachronically, there is a general tendency toward univerbation of light verb compounds. When the nominal element of the compound is coded like a patient, this process converts a formally transitive construction  $A(X)pV$  (where lower case 'p' symbolizes the P-like coding of a word that does not represent a participant, and (X) refers to possible oblique terms representing additional participants) into  $A(X)V$ , i.e. a construction with a participant coded like A and no participant coded like P. In languages with obligatory A-like coding, this results in perfectly canonical constructions, whereas in languages with obligatory P-like coding, the same process automatically results in the emergence of constructions violating the Obligatory Coding Principle.

Two opposite tendencies can be observed among languages with obligatory P-like coding: either the verbs resulting from the univerbation of  $pV$  compounds tend to maintain the exceptional coding frame  $A(X)V$ , or they tend to regularize it.

The first tendency is predominant in Basque. Basque makes a wide use of light verb compounds whose verbal element is *egin* 'do', and in many cases, the light verb compound is synonymous with a simplex verb whose root coincides with the non-verbal element of a do-compound, as in *bultza(tu) / bultza egin* 'push', or *dirdira(tu) / dirdir egin* 'shine'. In most present-day Basque varieties, the predominant tendency is that the simplex verb assigns to its arguments a coding identical to that observed in the light verb construction.

The tendency toward regularization can be illustrated by Andic languages (Nakh-Daghestanian). For example, in Andic languages, the translational equivalent of 'listen to' is either a light verb construction whose etymological meaning is 'fix ear at', with the coding frame  $\langle \text{ERG, ALL, abs} \rangle$ , or a simplex verb resulting from the univerbation of this compound. In some of the languages that have a simplex verb resulting from the univerbation of the compound 'fix ear at', this verb maintains the non-canonical coding frame  $\langle \text{ERG, ALL} \rangle$ , but in some others, its coding frame has been regularized as  $\langle \text{ABS, ALL} \rangle$ .

In my presentation, I would like to discuss a possible correlation with the distinction between strict and loose ergative coding (Harris 1985).