

Romance and Balkan factive islands in a nanosyntactic light

Lena Baunaz (University of Zurich, lena.baunaz@gmail.com)

Eric Lander (Stockholm University, eric.lander@su.se)

Claim. Factive verbs (*regret*, *remember*) select CPs that are presupposed to be true; non-factive verbs (*believe*, *want*) don't (Kiparsky & Kiparsky (K&K) 1971). Long-distance extraction out of factive CPs create **Weak Islands (WI)**, with only argument extraction being possible (Rizzi 1990, a.o). In addition to WIs, Romance and Balkan factive constructions may also involve **Strong Islands (SI)**, when both arguments and adjuncts are banned for extraction). The traditional approach in the literature is that the embedded CP is actually different in the case of factive constructions (K&K 1971, Rizzi 1990, de Cuba & Ürögdi 2009, Haegeman & Ürögdi 2010, a.o). On the basis of Romance (French (Fr.), Italian (It.)) and Balkan (Torlakian Serbian (TS), Croatian (Cr.), Bulgarian (Bg.), Modern Greek (MG)), we claim that this approach is basically correct, but with the added detail that the matrix predicate's structure is crucial too. More precisely, we claim that three features are responsible for the island effects observed; normally these features are spelled out as a comp(lementizer), but sometimes they can be split between the matrix verb and the comp, with consequences for lexicalization.

Data. In (1) it is seen that 'remember'+*deto* in Bg. yields SI effects (incompatible with both argument and adjunct extraction), whereas 'remember'+*če* yields WI effects (compatible with argument extraction but not adjunct extraction). In Fr. (2) the same effect is seen with 'understand'+IND (SI effects) and 'understand'+SUBJ (WI effects). These alternations also correlate with a change of meaning in the main predicate. In TS (3), *što* triggers a strong emotive reading of *žao mi je* 'regret' (with an experiencer subject, like 'feel bad' or 'feel sorrow'), which is less prominent with *da* (which yields something like 'regret to inform'). In Fr. (4), *comprendre*+SUBJ implies a more empathic subject (\approx 'understand') than with *comprendre*+IND (\approx 'realize'). We refer to these types as *emotive factives* vs. *cognitive factives*.

The verbal typology is even more complex than this. Verbs like 'dream', for ex., can be cognitive (triggering embedded IND in Romance) or emotive (triggering SUBJ). We view these facts in terms of syncretism, with an underlying functional sequence (fseq) $F4 > F3 > F2 > F1$, as in (5). The packaging of verbs may vary (and evolve) within one and the same language, as well as cross-linguistically. In this talk, our verb typology is extended to about 10-15 verbs in Fr., It., Bg., Cr., MG and TS.

Background. We adopt the nanosyntactic idea that morphemes are internally complex and composed of syntactico-semantic features which are hierarchically ordered according to a fseq. Hence comps and verbs are complex morphemes, lexicalizing structures of different sizes, as in (5), (6). In (6), Part(itive) comps range over (a given set of) propositional variables ('true' or 'false') and Spec(ific) comps have the property of locating the complement proposition with respect to a given point of reference, binding a single propositional variable, which corresponds to a single truth value ('true') (cf. Roussou 2010).

Analysis. Spec and Part comps contain FACT and are selected only by factive predicates. Spec comps are the largest and trigger SI effects; Part comps are medium-sized and trigger WI effects. Non-factive matrix predicates, on the other hand, must select the bare (*c*) comp. Romance and Balkan lexicalize these structures as in (7), with Fr. (and It.) showing a complete syncretism.

Now, combining the two sequences in (5) and (6) yields (8). Typically, the features in (8) will behave as two separate fseqs, with the matrix predicate spelling out the F features and the complementizer spelling out the Spec, Part, FACT, and *c* features. Fr. (and Bg.) illustrate such clear-cut cases, (9). In other languages, the fseqs are not divided so neatly by lexical item. In Cr., *da* is the bare comp selected by non-factives and not expected to cause any island effects, but it sometimes surfaces in a WI context

with a factive matrix predicate. Our proposal is that Part, which leads to the WI, is packaged on the matrix verb, leaving only FACT and *c* to be lexicalized by *da* (10). Thus the syntactic features necessary for WI effects are present but lexicalization obscures this fact.

One of the few contexts in Cr. in which SI effects are observed is similar to Eng. *Which article did you regret **it that** I had selected? As discussed, SI effects are due to the full structure Spec > Part > FACT > *c*. Since the main verb does not contain Spec in its lexical entry, Spec must be lexicalized by something else, namely *to* ‘it’, which leaves Part, FACT, and *c* to be spelled out by *što*, as shown in (11).

The lexical entry for *žao mi je* is at least [F3 F2 F1 Part] (10); with *što*, a subset of this feature set is spelled out (11). The lexical entry for *to* is [Spec D]; interestingly, in comp *š-to*, *to* seems to spell out D only (shrinking), i.e. *š-* = [Part FACT *c*] + *to* = [D].

Conclusion. Our analysis accounts for factive islands in Romance and Balkan. In both cases, comp is responsible for the blocking effect, but features of the comp-fseq can also be absorbed on the matrix verb, with lexicalization effects on comp.

- (1) a. Koji pomniš, **če/??deto** sreštna na pazara? (Bg)
 who remember.2SG that met.2SG at the market
 b. *Kogai pomniš, **če/deto** sreštna Maria na pazara?
 when remember.2SG that met.2SG M. at the market
- (2) a. Quelle voiture il comprend que Marie **ait/*a** acheté rapidement?
 which car he understands/realizes that M. has.SUBJ/IND bought quickly
 b. *Comment est-ce qu’il comprend que Marie **ait/a** acheté la voiture ?
 How Q he understands that M. has.SUBJ/IND bought the car
- (3) Žao mi je **što/da** si povrijedio Ivana. (TS)
 sorry. 1sg.dat aux that AUX.past.2SG hurt.PAST.PART John
 ‘I feel bad/regret to inform you that you hurt John.’
- (4) a. *Because he is such an understanding guy,*
 Georges comprend que Léon ait besoin de temps. [= emotive factive]
 G understands that L. have.SUBJ need of time
 b. *After hours of explanation (he is not very bright),*
 Georges comprend que la terre est/*soit ronde. [=cognitive factive]
 G. understands that the earth is.IND/be.SUBJ round
- (5) Verb fseq : Volitional (F4) > Emotive (F3) > Cognitive (F2) > Sentient (F1)
- (6) Comp fseq : Specific > Partitive > FACT > *c*
- (7)
- | | | | | | | |
|-------------------------------|------------------------|---------------|---------|-----------------|---------------|----------------------------|
| | <i>c</i> | No Island | Fr. que | Bg. <i>da</i> | ST <i>da</i> | Cr. <i>da</i> |
| | Part > FACT > <i>c</i> | Weak Island | Fr. que | Bg. <i>če</i> | ST --- | Cr. <i>što</i> |
| Spec > Part > FACT > <i>c</i> | | Strong Island | Fr. que | Bg. <i>deto</i> | ST <i>što</i> | Cr. <i>to</i> + <i>što</i> |
- (8) F4 > F3 > F2 > F1 > Spec > Part > FACT > *c*
- (9) [*dire* = F1] + [*que*, Bg. *da* = *c*] No island
 [*comprendre* = F3 F2 F1] + [*que*, Bg. *če* = Part FACT *c*] Weak island
 [*comprendre* = F2 F1] + [*que*, Bg. *deto* = Spec Part FACT T *c*] Strong island
- (10) [*žao mi je* = F3 F2 F1 Part] + [*da* = FACT *c*] Weak island
- (11) [*žao mi je* = F3 F2 F1] + [*to* = Spec D] + [*što* = Part FACT *c*] Strong island