A nanosyntactic analysis of mood selection in French and Balkan languages

The *ABA theorem states that syncretism can target only adjacent regions in a paradigm (Bobaljik 2007, 2012; Caha 2009). We argue that *ABA in combination with a 'peeling' approach (Caha 2009) as implemented in Nanosyntax (NS) can account for puzzling facts about IND(icative) and SUBJ(unctive) mood selection in Romance vs. Balkan.

Facts. Romance marks SUBJ mood on the verb (*verbal mood*), while Balkan uses a special SUBJ COMP(lementizer) (*clausal mood*, see Sočanać 2017, a.o). In Balkan (and Modern Greek (MG)), SUBJ morphology on V has been lost and replaced by *perfective non-past* (PNP) morphology (see Giannakidou 2009, Sočanać 2017 a.o), encoding aspect (and tense) but not mood.

Puzzle. SUBJ is a *dependent mood* selected under verbs associated with special semantic features. Based on (MG), Giannakidou 1998 (and subseq.) argue that veridicality – the stance taken by some individual toward the truth of a proposition – licenses mood choice: veridicality licenses IND, non-veridicality licenses SUBJ. While the correlation mood/veridicality applies very well in Balkan/Slavic (Todorovic 2012), it fails in Romance: despite being veridical, Romance emotive factives (*regret*) trigger the SUBJ mood (Quer 2009 a.o). As a solution, A1 2018 (a.o) refines the notion of veridicality and argues for a tripartition of embedding verbs in French/Balkan (directives/desire/emotive factives/epistemic/ saying), relative to the truth of the embedded proposition w.r.t both subject and speaker (= strong veridicality, SV), to either subject or to speaker (= relative veridicality, RV), or to neither (= non-veridicality, NV). Crucially, verbs can be syncretic in being either SV or RV (T.1). The distinction is tracked by different morphemes in C in Balkan, or by SUBJ vs. IND marking on V in Romance (T. 2-3).

T.1	**English translation**	**MG**	**SC**	**Bg**	**Fr**
SV	'remember' 'regret' 'understand'	*thimame*	*sjetiti se* *%žaliti*	*pomnja* *sǎžaljavam*	*se rappeler* *comprendre*
RV	'remember' 'regret' 'understand'	*thimame* *lipame*	*% sjetiti se* *žaliti*	*pomnja* *sǎžaljavam*	*regretter* *comprendre*
NV	'say' ; 'want'	*leo; thelo*	*reći* ; *željeti*	*kazvam; iskam*	*dire ; vouloir*

T.2	**Mood on C**		**Mood on V**	
	MG	**Fr**	**MG**	**Fr**
SV	pu	que	IND	IND
RV	pu /oti	que	IND	SUBJ
NV$_1$	oti	que	IND	IND
NV$_2$	na$_{SUBJ}$	que	(PNP)	SUBJ

T.3	**MG**	**Bulg.**	**Serbian**	**Croatian**	**Fr**
SV	pu	deto	što	da	que
RV	pu oti	če	što	da	que
NV$_1$	oti	če	da	da	que
NV$_2$	na$_{SUBJ}$	da$_{SUBJ}$	da	da	que

Yet this threeway distinction does not provide an orderly way of accounting for mood selection. In MG and Bg, veridicality and mood-marking appear to be unrelated: NV verbs can take either IND or SUBJ complements (T.3), splitting the NV group in two. Even more problematic is French: SV and NV$_1$ verbs select for IND complements, while RV and NV$_2$ verbs for SUBJ complements, resulting in an *ABA violation (T.2). On the other hand, there are no problematic *ABA violations in T.3, i.e. veridicality cleanly tracks COMP selection in French and Balkan.

Claim. We reconcile the way in which mood is realized with the theory of COMP selection. The ABA in French is thus the result of unduly mixing up two independent processes: (i) the internal structure of COMPS in terms of veridicality, and (ii) the selection of embedded mood by certain matrix verbs. Taken on its own, neither process violates the *ABA theorem. We show that mood is triggered by the internal structure of the selecting predicate. T.4 shows that IND and SUBJ obey the adjacency requirement on syncretism, i.e. there are no illicit ABAs. The problem arises in the observation that a COMP$_{RV}$ 'selects' a SUBJ embedded verb. However, the correct way of framing this fact is not that the COMP selects the embedded mood; rather, it is the higher matrix verb that is responsible for this. It is important to recognize

that veridicality is an (at least partially) independent variable, with different COMP structures being compatible with different matrix verbs (T.4): COMP_{RV} and COMP_{NV2} are compatible with matrix verbs taking the SUBJ mood; COMP_{SV} and COMP_{NV1} are compatible with verbs taking the IND mood.

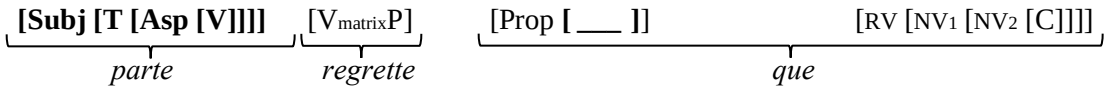
T.4	**Predicate class**	**Its internal structure** (functional sequence (fseq))	**Selects**	**COMP compatibility**
Directives		CAUSE > VOLITIONAL > EMOTIVE > SENTIENT	SUBJ	NV_2
Desire		VOLITIONAL> EMOTIVE >SENTIENT	SUBJ	NV_2
Emotive factives		EMOTIVE > SENTIENT	SUBJ	RV
…of saying/epistemic		SENTIENT	IND	SV, NV_1

Assumptions. The fseq of the COMP morpheme is [SV [RV [NV_1 [NV_2 [C]]]]], different subsets of which may be lexicalized (A1 2016/2018). The other main complex constituents of the sentence are the embedded V with inflectional layers ([$\text{V}_{\text{embedded}}$P]) and the matrix V ([$\text{V}_{\text{matrix}}$P], see T.4). Each of these three constituents is built in a separate cognitive 'workspace' (Pantcheva 2011, Starke 2013, a.o.). These constituents are then brought together into the shared workspace of the sentence (1).

Analysis. (1) is the derivation of a RV structure like *Jean regrette que Marie parte* 'John $\text{regrets}_{\text{RV}}$ that_{RV} M. $\text{leave}_{\text{SUBJ}}$'. First, the complex morpheme COMP_{RV} is built in its own workspace, and then merged in the clausal fseq, as indicated by the subscript #1. Next, the complex [$\text{V}_{\text{embedded}}$P] is built with its inflectional (TAM) layers (Subj = SUBJ morphology; Prop is for *proposition*, which is the feature yielding IND morphology) and merged with COMP_{RV}. (see #2 in (1)). Then V_{matrix} is built and merged (#3) at the top of the clausal fseq (note that we remain agnostic about the exact labels of the projections in the clausal fseq). As seen in T.4, it is the class of V_{matrix} that determines whether the embedded mood is IND or SUBJ. Here *regret* is an emotive-factive matrix verb (i.e. [V_{matrix}P] = [EMOTIVE [SENTIENT [VP]]]), which at this stage of the derivation is in a local relation to [$\text{V}_{\text{embedded}}$P]. We posit that this kind of relation makes it possible for V_{matrix} to select some subset of the full verbal structure (in this case a verb with SUBJ inflection) and move it to the left of V_{matrix}P. Importantly, removing the SUBJ structure leaves behind a complement-less Prop (since SubjP has been moved out); this Prop 'peel' can be spelled out as part of COMP itself, Fr. *que* (i.e. the structure of *que* as stored in the lexicon includes the Prop feature, as seen in (2)). We assume that [V-matrix *regrette* [C *que*]] moves up later to give the correct word order. However, it is not *only* in such cases that *que* can be spelled out. *Que* can also be spelled out if an indicative verb has been selected/moved to the left of V_{matrix}P, as seen in (3). The lexical entry in (2) can still, by the Superset Principle, spell out the leftover structure in (3) as *que*. That is to say, the lexical structure with the peel in (2) is still a superset of [SV [RV [NV_1 [NV_2 [C]]]]] (or any subset of this structure). In other words, *que* is spelled out whether indicative or subjunctive has been selected by the matrix verb. Thus V_{matrix}P does not directly determine the kind (size) of COMP. In other words, COMP grows and shrinks independently of the behavior of the matrix verb and the process of mood-selection happening above it. The lexical entry in (2) is able to accommodate this fact if we adopt the Revised Superset Principle, given in (4).

This analysis has interesting consequences for Balkan and its PNP inflection. Assuming that PNP involves subextracting [Asp [V]] out of the verbal complex, the peel left behind will be [Subj [T [__]]], as in (5) (Prop is not involved because MG *na*-complements are non-propositional, see Roussou 2010). If *na* spells out this peel and thus contains Subj and T in its lexical entry, we explain why *na* is used as a marker of subjunctive but also as an infinitival marker in T (a syncretism widespread in Balkan; Sočanać 2017 a.o).

(1) [V_{matrix}P] #3 [Prop [Subj [T [Asp [V]]]]] #2 [RV [NV_1 [NV_2 [C]]]] #1



(2) < que ⇔ [Prop **[___]**] [RV [NV_1 [NV_2 [C]]]] >

(3) **[Prop [Subj [T [Asp [V]]]]]** [V_{matrix}P] [___] [$\text{COMP}_{\text{SV/NV1}}$]

(4) A lexical entry L may spell out a syntactic node SN iff the features of L are a superset of the features dominated by SN. (Vanden Wyngaerd 2018: 289, his (6))

(5) < *na* ⇔ [Subj [T __]] [NV_2 [C]] >