The Copula in VSO languages: Projection, labeling and X(P)-movement.

1. Introduction. One of the most attractive features of the VP remnant analysis of many VSO languages is that in copular sentences what is fronted is the entire predicate and not a head. Yet, in what follows, I will share the caveats w.r.t. this approach pointed out by McCloskey (2005), and develop an analysis of copular clauses in Irish and Māori which relies on an independently motivated modification of the Labeling Algorithm (=LA) proposed by Chomsky (2013). The emerging picture is the attested word orders are due to the ‘blind’ application of this modified LA, modulo the presence or not of an overtly realized BE head and φ-features, addressing thus quite directly the issue of whether the Predicate Phrase has a head that projects, and whether it can move up or not, and which specific circumstances this is possible.

2. VSO and Tense-doubling. In Irish, evidence from embedded VSO and subject-ADV orders (McCloskey (1991), (1997)), indicates that V raises to a position (=F) higher than T, whereas subjects leave vP: [CP C [TP V+F [TP S T [φ ADV]... In addition, sentential modifiers before complementizers suggest that F is Fin: [ForceP [Top/FocP PP [FinP Comp.+V [TP S... (Hendrick (2000)). However, it’s unclear is why adjacency must hold within the C+F+(V+)T complex.

A crucial property of Irish/Sc. Gaelic is the existence of multiple exponence of specific tense values in T and F, e.g. Ir. go gcuireann é (gl. that Pres put-Pres he) ‘that he puts’ vs. gur chuir é (gl. that Past put-Past he) ‘that he put’. My proposal then is that this is the key factor involved:

(1) If two heads in the same phase, X and its complement Y, share specific values (=‘i’) for a certain interpretable feature, no ZP can intervene: [X [ (*ZP) [Y [Z]]]].

This approach can be extended (pace Medeiros (2014), which defends a VP remnant analysis for its close relative Hawai’ian) to Māori, a Polynesian language which displays the following sequence: TMA1-V-(TMA2)-S...; where TMA1 realizes tense in C, and TMA2 is located in T; since embedded clauses (truly so, given indexical shift) are directly introduced by TMA1 and there are no lexical complementizers, locating TMA1 in C seems well motivated.

It must be noticed that (1) predicts that there are will be no SVO lgs. displaying tense doubling:*did S V-ed,*that-ed S V-ed, etc., and it excludes some apparent counterexamples: i) complementizer agreement, as in West Flemish, etc., involves the sharing of uninterpretable agreement features, and hence, doesn’t falsify (1), which deals with the sharing of interpretable feature values.ii) E. for DP to V clauses involve the sharing of [+finite], not specific tense values, iii) E. conditional inversion, as in had Mary said this... involves movement of AUX in T to C, and hence is not subject to 1), if it is a condition on external merge. In any case, the status of (1) within minimalism is unclear, and a principled account is required.

3. Labeling. Let’s assume that α in F [α DP [TP]] can be labeled either via criterial φ-feature agreement, leading to (2)a, as in Chomsky (2013), or via X-movement iff (1) holds, as in (2)b:

(2)a F [α=agr. DP [agr.]]
(2)b F+T1 [α=TP DP [t1]]

Hence, (1) is derived from the LA if we maintain endocentricity and syntactic head movement just under its limited conditions, which will not be ‘conditions’, but just the overt realization of an UG option. If so, α=TP in (2)b doesn’t entail true subject agreement, a desirable result because subject agreement in Irish must be analyzed as D-incorporation (Baker&Hale (1990), etc.), due to the complementarity between lexical subjects and V+agr. (e.g. beireann fir lit.: carry-pres-3sg men), and in Māori, φ-agreement is completely absent.

4.1. Irish. Considering (2)a and b, both possibilities are seen in Irish nonverbal predication (Carnie(2000), etc): the subject-first order (3)a is found with definite or referential attributed properties, and corresponds to the Labeling option in (2)a, whereas the predicate-first order (3)b is found with indefinite, non referential attributes, and corresponds to (2)b:

(3)a Is é Seamus an captaen b Is teangeolaí (í) Máire
COP AGR S. the captain COP linguist (AGR) M.
‘Seamus is the captain’ ‘Maire is a linguist’

(3)b is an instance of Labeling via Predicate X-movement, as in (2)b, including apparently phrasal Predicates (Carnie, loc. cit): its structure is [\text{Pred}_Máire [\text{Pred} (í) \text{teangeolaí}]], headed by the optional post-predicate pronominal to which the predicate moves before the complex head moves to the coindexed COP. This is the only option because subject raising would be string-vacuous, and it takes place under endocentricity plus ‘i-sharing’, as in (1)/(2)b. In (3)a, the attribute is definite, hence, it is phrasal and X-movement is excluded; the structure is then a true Small Clause, and Labeling via 𝜓-agreement as in (2)a is the only option, phonologically realized as an obligatory pronominal element, which disregards the complementarity effect.

4.2. Māori. First, equative sentences, with a definite predicate, are true SCs \text{[sc \text{XP}_{\text{subj}} \text{YP}_{\text{pred.}}]} and as such, given the lack of 𝜓-agreement, the option of criterial labeling as in (2)a, (3)a for Irish, is unavailable; hence, fronting of the subject or the predicate, to a left peripheral TOP position headed by the particle ko are the only two options for the otherwise unlabelled object:

(4)a Ko ia te rooia b Ko te rooia ia
Part. he the lawyer Part. the lawyer he
‘s/he is the lawyer’ ‘the lawyer is s/he’

With indefinite attributes, instead of ko-fronting, the entire Pred. headed by the element he is fronted. The categorial status of he is a very debatable issue: Chung & Ladusaw(2004) argue that it is an indefinite composed via the non saturating mode Restrict, but the problems posed by its polyfunctional character remain (e.g., Cook(1999), and Du Feu(1996) which show that he, w can head verbal predicates in Hawai’ian and Rapanui). Here, I suggest that Māori he can the head of the PredP, in a way similar to Irish pronouns in (3)b, and therefore, the label for PredP can be assigned via Pred⁰ raising to C, where tense is realized, hence an instance of (2)b above:

\[ C_{\text{Tense}}[^{\text{Pred}^0\text{he...}}],[^{\text{lo-PredP}} \text{XP}_{\text{subj}}][^{\text{Pred}^0\text{t}} \text{YP}_{\text{pred.}}]. \]

This approach is supported by the fact that he-predicates are sensitive to tense, he'i being the form used in future statements, as in (5)a vs b:

(5)a He rooia ia b Hei rooia ia
Pred lawyer s/he Pred-FUT lawyer s/he
‘s/he is/was a lawyer’ ‘s/he will be a lawyer’

Given that equative sentences are insensitive to tense, one could say that the difference between (4) and (5) reproduces in language particular terms, the long-standing debate (see Moro(2013, ch.5) about the nature of copular clauses, either as device to link the rhêma with Tense, as in the Aristotle’s De Interpretatione, whih would correspond to (5), or as a true copula that merely links subject and predicate, as in Abelard’s work, which would be instantiated by (4). Finally, existential sentences and Irish tá, roughly compatible with stage level Preds, would fit into the picture, as instances of AUX-Subject-Pred⁰, an option not precluded by (1) or the options in (2).
