

# TAG and Raising in VSO Languages\*

Heidi Harley and Seth Kulick

Institute for Research in Cognitive Science

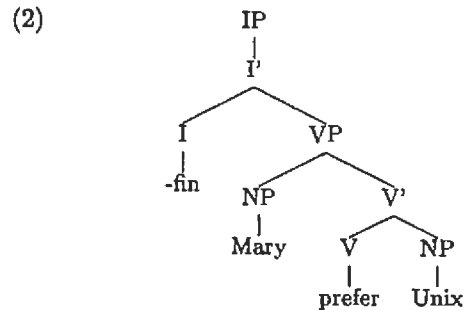
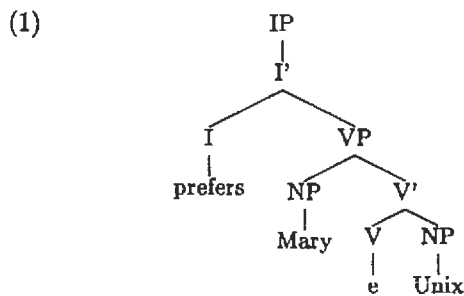
University of Pennsylvania

Suite 400A, 3401 Walnut Street, Philadelphia, PA 19104-6228

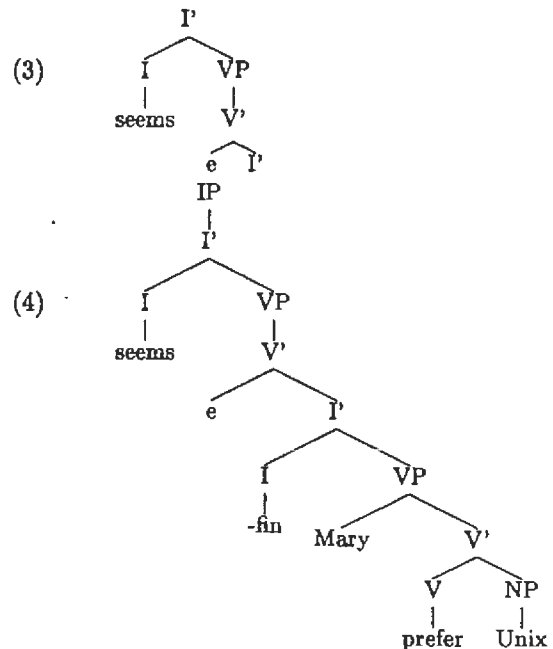
hharley@babel.ling.upenn.edu, skulick@linc.cis.upenn.edu

The derivation of unbounded Subject-to-Subject Raising in languages like English is a problem particularly elegantly treated by Tree Adjoining Grammar. The adjoining operation inserts auxiliary trees headed by raising verbs between the subject in Spec-IP and the root verb, distancing the subject from its original local relationship with the root verb and producing a final multi-clausal structure with the subject in the final subject position in the matrix clause.

Verb-initial languages could pose a challenge to unadorned TAG in this central paradigm if it can be shown that they exhibit true raising structures. Consider the possible structures of the pseudo-English VSO finite and non-finite clauses in (1) and (2). In (1), the tensed verb appears to the left of its subject, and in (2) a structure with a non-finite verb to the right of the subject is shown. (This reflects the fact that in general, VSO clauses are only VSO in the finite case).



If an auxiliary *seems* tree like (3) were to adjoin to (2) above, the result would be (4) below, not a true Raising structure at all, as the subject remains in its original position in the embedded clause. In the formal system of basic TAG, it is generally true that no VSO language is predicted to exhibit a true raising structure, since the finite raising verb must appear in initial position.



\*We would like to thank Randall Hendrick, Aravind Joshi, Tony Kroch, Maggie Tallerman, and two anonymous reviewers for valuable comments. This work is supported by grant NSFSTC89-20230.

The linguistic question, then, is whether it can be shown that a VSO language does exhibit a true

raising structure in which the subject is in the matrix clause. This is a non-trivial question for two reasons. First, the string will exhibit identical word order whether or not the subject is in the matrix or embedded clause, since subjects in finite clauses follow the verb. Secondly, even if it is possible to show that the subject is in the matrix clause, it must be shown that the verb in question is a true Raising verb, and not a Control verb, controlling an in situ null argument in the embedded clause. Only when both these conditions are met can we show that basic TAG is insufficient to treat VSO raising.

In Welsh, a Celtic VSO language, there are two verbs which are potential raising verbs, *digwydd* ('happen'), and *dechrau* ('begin'). We can immediately test whether or not the subject of these verbs appears in the matrix clause by using a participial form of the verb, with a finite auxiliary in initial position. If the subject is in the embedded clause, as in (4), it should make no difference whether or not the raising verb is finite or participial; it should continue to precede the embedded subject; the counterpart to *Mary has seemed to prefer Unix in the past* should be *has seemed Mary to prefer Unix in the past*. On the other hand, if the subject is in the matrix clause, the raising participle should appear to the right of the subject, since it is non-finite. We can immediately see the latter is the case:

- (5) Mae Siôn yn digwydd bod yn gweld  
 Is John prt.happen be.inf prt.see  
 Mair  
 Mary  
 'John happens to be seeing Mary' (Hendrick 1988)

We must then show that *digwydd* is a raising verb, not a control verb. Following Hendrick (1988), we make this argument from the behavior of expletives. Expletives are possible as the subject of raising verbs, but not of control verbs: *There seems/\*tries to be a spider on the wall*.

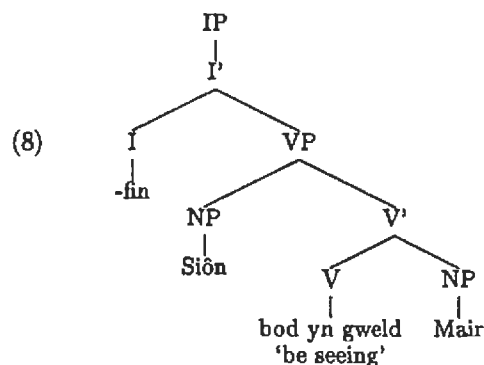
Welsh has an expletive subject *yna* that behaves essentially identically to English *there*, appearing in locative, existential and possessive constructions, as in (6) below.

- (6) Mae yna oriad gyda John  
 Is there a key with John  
 'There is a key with John/ John has a key.'

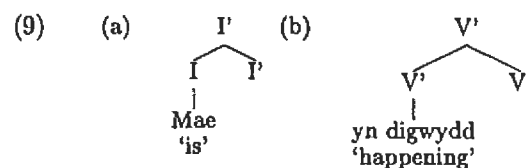
Crucially, then, this expletive may appear as the subject of *digwydd*, but not of control verbs like *mynd* ('go' in future sense). The latter also differ from the former in that they require an overt complementizer *i* to appear between the matrix and embedded clauses:

- (7) a. Mae yna yn digwydd bod oriad  
 Is there prt.happen be.inf a key  
 gyda Siôn  
 with John  
 'There happens to be a key with  
 John/ John happens to have a key.'  
 b. \*Mae yna yn mynd i bod  
 Is there prt.go Comp be.inf  
 oriad gyda Siôn  
 a key with John  
 'There is going to be a key with  
 John/John is going to have a key.'

Welsh, then, is a VSO language with a true raising structure. The word order in finite clauses entails that the basic TAG adjoining mechanism will not be able to generate the structures necessary, and recourse to a multicomponent derivation must be made. Consider the non-finite tree in (2) above, repeated as (8) to represent the structure of the embedded Welsh non-finite clause in (5)<sup>1</sup>



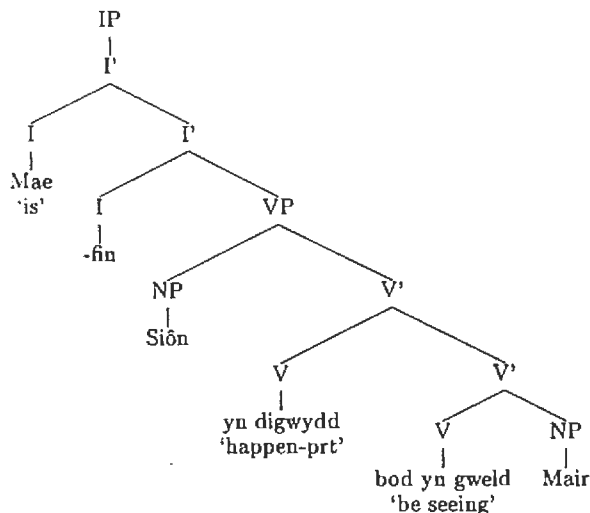
In order to get *Siôn* into subject position of the matrix clause, in a sentence like (5) above given this structure, two auxiliary trees must adjoin into the elementary tree, as shown in (9ab). One tree, headed by *Mae*, the finite copula, must substitute/adjoin in to the elementary tree in front of *Siôn*, and another, headed by the participle form of the raising verb, *yn digwydd*, must adjoin in below *Siôn*, creating the raising structure. Let us consider what such auxiliary trees must look like:



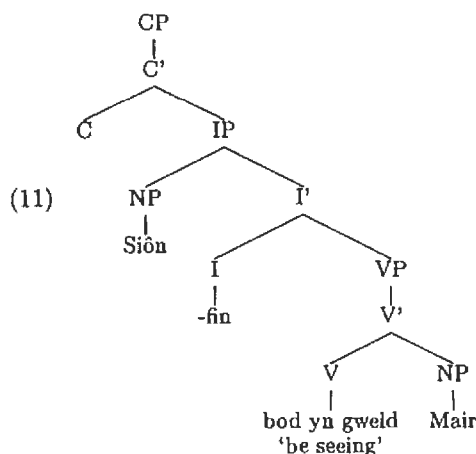
If we adjoin these trees into the elementary tree in (8), we arrive at the final structure in (10):

<sup>1</sup>We represent here *bod yn gweld* as a complex NP for convenience. The use of a VP-shell might be more desirable, although that issue is irrelevant for this discussion.

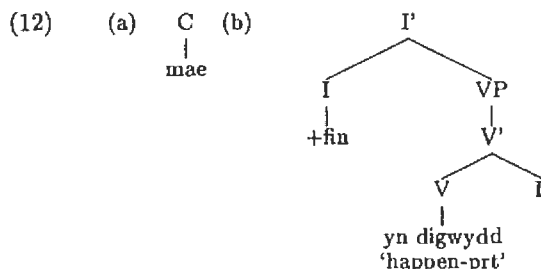
(10)



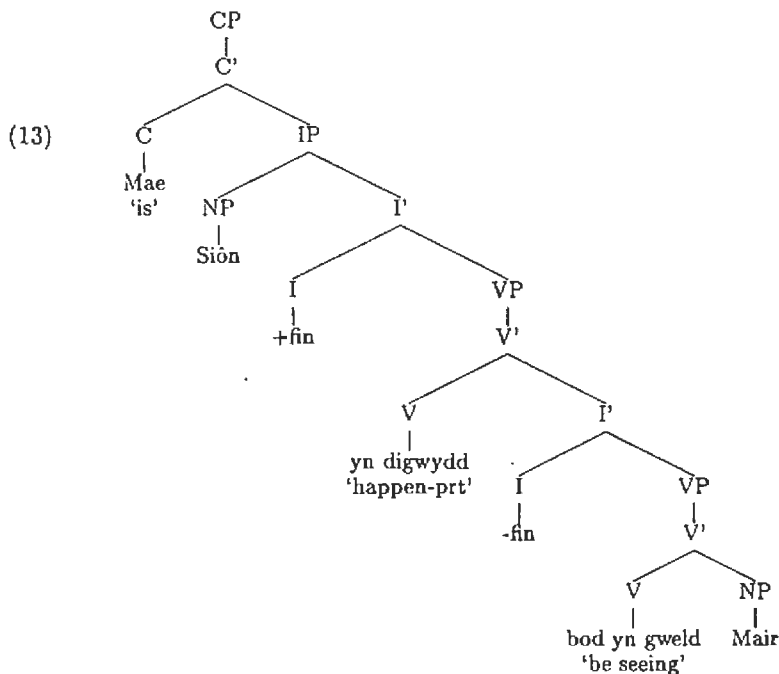
While this provides us with the correct final word order, it is linguistically unsatisfactory for two reasons. Firstly, we have destroyed the relationship between the [-fin] I head and the non-finite form of the verb *bod* by interpolating the participle *yn digwydd* (which itself needs to be related to the finite form *Mae*, now separated from it by the [-fin] head). Secondly, in a purely theory-internal problem, if Spec-VP is universally a theta-position, which is widely assumed, the subject *Siôn* is in a theta position in what is now the matrix clause. That is, "raising" has been to a theta-position, a theoretically incoherent result. Both these problems are avoided if we assume a different final clause structure for Welsh VSO sentences than that presented in the finite VSO structure in (1). The problem here is that the finite verb in (1) has raised only as far as I. This creates the dual problem above: if finite verbs are in the I head, the multicomponent auxiliary tree will always interfere with non-finite I head of the elementary tree in a raising structure, and the subject must appear in the specifier of VP, as there is no higher non-theta position available.



In an infinitive clause, the subject will still appear in Spec-IP, rather than Spec-VP, giving the correct SVO order for the infinitive. (Note that since in TAG there is no "movement" of the subject, it is not impossible to place the subject in [Spec, VP] in the lower clause, while ending up in [Spec, IP] in the higher clause.) Our revised elementary tree for the nonfinite clause is shown in (11), and the auxiliary trees which will adjoin into this structure are shown in (12):



Consider, on the other hand, the possibilities which arise if finite verbs in Welsh raise as far as C. In this case the subject appears in [Spec, IP], a non-theta position, and as we shall see, no problem for the insertion of the topmost auxiliary tree will arise for the MC-adjunction necessary to derive the raising structure.



This adjunction gives us the final structure for the raising construction, (13), which makes much more linguistic sense than the IP tree above:

The result seems to suggest that in a TAG framework, the only VSO languages which are predicted to exhibit raising structures will feature positioning the finite verb in C.<sup>2</sup>

The derivation we end up with is essentially identical to that proposed by Frank (1992) for an analogous problem in English: the formation of the question "Does John seem to like Mary?". This supports the view that the problem raised by that particular derivation (the requirement of a multi-component set) was not just a weird quirk, but rather just one example of the widespread need for such a derivation.

## References

Frank, R. (1992) *Syntactic Locality and Tree Adjoining Grammar: Grammatical, Acquisition and*

<sup>2</sup>Irish, which has been argued to not move its finite verb to C (McCloskey 1992), is an interesting case. In the "Exceptional Raising Construction" (McCloskey 1984b, Stowell 1989), an apparently-raised NP appears with a preposition (or perhaps a Case marker). The proper analysis of this construction is very unclear, and we leave it aside. Whether or not Irish has the more traditional subject raising structure is unclear (e.g., McCloskey (1984a) claims that there is, while Stenson (1981) states that there is none.) Clearly more study is required.

*Processing Perspectives*, Doctoral dissertation, University of Pennsylvania.

Hendrick, R. (1988) *Anaphora in Celtic and Universal Grammar*, Dordrecht:Kluwer Academic Publishers.

McCloskey, J. (1984a) "Case, Movement, and Raising in Modern Irish," in *West Coast Conference on Formal Linguistics*, ?

McCloskey, J. (1984b) "Raising, Subcategorization, and Selection in Modern Irish," *Natural Language and Linguistic Theory* ?, ?

McCloskey, J. (1992) "On the Scope of Verb Movement in Modern Irish," Technical Report LRC-92-10, Linguistic Research Center, Cowell College, University of California at Santa Cruz.

Stenson, N. (1981) *Studies in Irish Syntax (ars Linguistica 8)*, Gunter Narr Verlag, Tübingen.

Stowell, T. (1989) "Raising in Irish and the Projection Principle," *Natural Language and Linguistic Theory* 7, 317-359.