

Anne ABEILLE

Département de sciences du langage,  
 Université de Paris 8,  
 2 rue de la liberté, 93526 Saint Denis  
 abeille@frp8v11.bitnet

**Introduction**

Lexicalized Feature-based TAGs have been used for describing various syntactic phenomena across several languages, and for writing sizable grammars for French and English (Abeillé 1988 and 1991a, Abeillé et al. 1990) which have been incorporated into a fully implemented parsing system (Schabes 1990)<sup>1</sup>. The main linguistic advantages of such a formalism are that:

- it provides a larger domain of locality than CFG-based formalisms such as HPSG or LFG,
- it allows factoring recursion from the domain of local dependencies,

- as a consequence of both above properties, it allows the grammar to be totally lexicalized (Schabes et al. 1988).

However, a certain number of syntactic phenomena are difficult to represent in this framework. We focus here on French pronominal clitics which are cases of non canonical argument realization. We show how they can be naturally handled using the "Synchronous TAG" extension of the formalism as a further set of wellformedness conditions on the language and we extend this treatment to other cases of mismatch between syntactic attachment and semantic role.

**1. Motivation**

French pronominal clitics have been subject to numerous studies in both theoretical and computational linguistics. We restrict here our discussion to syntactic properties of clitics and thus ignore most of their morphological, phonological or binding properties<sup>2</sup>. We show that they cannot be handled by existing TAG variants such as FOTAGs or MCTAGs (unless major changes are made to the existing French TAG grammar) but that Synchronous TAGs provide an elegant and inexpensive solution.

**1.1. Some linguistic properties of French clitics** French pronominal clitics fall into 6 groups: the nominative ones (*je, on, il...*), the 1st, 2d person and reflexive ones (*se, me...*), the accusative ones (*le, la...*), the dative ones (*lui, leur...*) plus *en* and *y* which each forms a class on its own. They follow each other in a rigid order: nom < refl < acc < dat < y < en, and are subject to numerous corestrictions<sup>3</sup>: no identical

ambiguous forms can cooccur even if they play different (and compatible) syntactic functions:

(1a) *Paul songe à Marie dans le jardin.* (Paul is dreaming about Marie in the garden)

(1b) \* *Paul y y songe* (Paul dreams of her there)

(2a) *Paul remplit un verre de vin* (Paul fills one glass with wine)

(2b) \* *Paul en en remplit un* (Paul fills one with it)

No 1st or 2nd person direct object may cooccur with a dative clitic: \* *Paul me leur présente* (Paul introduces me to them).

Clitics do not have the same syntactic properties as NPs: they do not coordinate with NPs, nor take NP modifiers, and are usually assigned a specific category (Clit). It is commonly agreed upon that French pronominal clitics (plus the negative marker *ne*) form a syntactic verbal cluster with the main verb (or with the inflected auxiliary if there is one) of the sentence<sup>4</sup>.

In standard French, pronominal clitics are in complementary distribution with syntactic phrases as complements of various lexical heads (V, N, A). They may appear on a verbal head of which they are not an argument:

(3) *Jean en est aimé* (cpl of the past participle)

(4) *Jean lui est fidèle* (cpl of the adjective)

They are also subject to numerous lexical constraints. Object cliticization may be ruled out by certain verbs which impose a non clitic pronominal form:

(5a) *Jean (ressemble + pense) à Paul* (Jean resembles/thinks about Paul)

(5b) *Jean lui (ressemble + \* pense).*

(5c) *Jean (?\* ressemble + pense) à lui*

(6a) *Jean (pense + aime) que Marie lui raconte des histoires.* (Jean thinks/likes it that Marie tells him stories)

(6b) *Jean (le pense + \* l'aime).* (ruled out with intended meaning)

Clitics *en* and *y* may also behave as adverbial modifiers. Several proposals have been made in computational linguistics for handling these phenomena. Bès 1988 gave a GPSG treatment of French clitic objects, which essentially considers them as preverbal NPs (with a special <Clit> feature), except for inverted subject clitics which he considers verbal suffixes. He does not consider adverbial uses nor cases of non local dependencies or of auxiliary-verb combinations. Lexical entries for clitics in

1. It is also the basis of an interactive NLP environment (Paroubek et al. 1992). See Joshi 1987a and Schabes et al. 1988 for an introduction to TAGs.

2. French pronominal clitics are subject to obligatory sandhi phenomena such as liaison and elision.

3. These properties can easily be described with a finite automaton; Cf Gross 1989.

4. See for example Gross 1968, Kayne 1975 for various arguments for a Clitics-Verb constituent in French. See Abeillé 1992 for arguments against a VP constituent.

dislocated constructions are duplicated. Baschung et al. 1987 present a UCG treatment with a more restricted coverage which considers object clitics as functors taking (on their right) a verb expecting an NP (or PP) argument on its right and returning a more saturated verb. They do not give a unified treatment of subject clitics which they consider as standard NPs nor do they handle non local dependencies. Bès et al 1989 give a slightly modified version of this treatment allowing for what they call French "semi-free" word order. Miller 1991, using a HPSG-style framework, considers clitics as "phrasal affixes" on VP and uses optional lexical rules to update the subcategorization frame of the corresponding verbal entries and foot features to keep track of the presence of a clitic in the tree. He accounts for a lot of non local dependencies (including causative constructions) but needs extra constraints to handle locality constraints. He does not talk about inverted clitics nor dislocated constructions. This treatment is not, to our knowledge, implemented in a computational application.

### 1.2. Difficulties with existing TAG variants

We can first put aside "frozen" clitics which are easily handled in Lexicalized TAGs: they do not play any semantic role and "co-head" the elementary trees of their predicate ("impersonal" *il, se* in "middle" constructions and various idioms). Clitics with a semantic role (adverbial modifiers or arguments of a verb, an adjective or a noun) are more difficult to handle.

One could think of adding to the existing tree families (associated with the predicates) an elementary tree (with a substitution node for the clitic and an empty argument node) for each possible clitic (or clitic combination) realization. This would be somewhat unprincipled and lead to numerous extra trees<sup>5</sup>, unless one generates the new trees by metarules or lexical rules (Vijay-Shanker and Schabes 1992, Becker 1990). It would also separate the syntax of adverbial clitics from that of argumental ones and disallow many non local dependencies.

One might also consider extensions of TAGs, such as FOTAG or MCTAG variants. In Multicomponent TAGs, used for example by Kroch and Joshi 1986, the elementary units of the grammar are tree sets (noted here with braces) instead on single trees. Adjunction (or substitution) thus becomes simultaneous adjunction (or substitution) of the different members of a tree set into a tree (or into the members of another tree set). The different members of such a set need not be of the same type, and we could use a set comprising an auxiliary tree headed by the clitic and an initial tree headed by the empty string for the corresponding argument position:

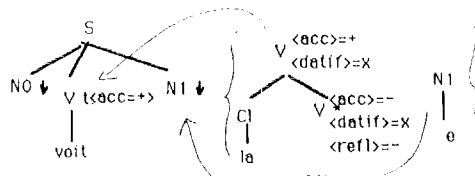


Figure 1. *Jean la voit* with MCTAGs

The substitution node corresponding to the NP realization of the complement is thus filled with the empty string when the clitic adjoins. However, this will not work for PP complements, since in this case, the whole PP subtree with the preposition "co-head" would have to be "zeroed" when the clitic adjoins, an operation not provided by the formalism<sup>6</sup>.

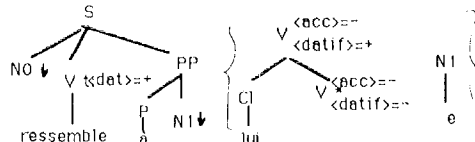


Figure 2. Blocking *Jean lui ressemble* with MCTAGs.

"Free Order" TAGs is a variant analogous to the ID/LP format for GPSG which was first introduced by Joshi 1987b and developed by Becker et al. 1991. Argument clitics would thus be treated as moved NPs (substituted at the same NP nodes) according to LP statements such as the following: N1(+Clit) < V. This representation might be attractive for handling cases of "clitic climbing" or non local dependencies but faces unescapable problems:

- It is similarly unable to account for the argument PP/clitic alternation, since the node to be substituted is an NP, not a PP in the non clitic case.
  - It prevents from having an unified syntactic representation of the different clitics (it does not make any syntactic distinction between NP and clitic subjects)
  - It does not regard clitics-verb as a constituent in the syntax, and it is difficult to see how corestrictions between clitics could be handled (the same warning holds for prohibiting adverb insertion between clitics and verb).
- Current TAGs versions thus do not provide a satisfactory account of the non frozen pronominal clitics. We now turn to an alternate representation which will not be strictly syntactic but involve the syntactico-semantic interface defined in the Synchronous TAGs framework.

## 2. A Synchronous TAG representation

**2.1. Synchronous TAG** Synchronous TAGs have been introduced by Shieber and Schabes 1990 to characterize correspondences between Tree-Adjoining languages. They can be used for relating two TAGs for two different languages for the purpose of machine translation (Abeillé et al. 1990), or for relating a

5. Two different trees would be needed for *Jean le lui donne* (Jean gives it to him) where first clitic=N1, second clitic =N2, and for *Jean me le donne* (Jean gives it to me) where first clitic=N2, second clitic =N1.

6. We recall that, due to their lack of semantic autonomy, we consider argument-marking prepositions as co-heads of the elementary tree of their predicate, contrary to prepositions heading adjuncts which are autonomous heads of their auxiliary tree.

syntactic TAG and a semantic one for the same language, for the purpose of generation or semantic analysis<sup>7</sup>. We consider here the latter case and assume that both syntactic and semantic TAGs are lexicalized and feature-based<sup>8</sup>.

In Synchronous TAGs, TAG elementary trees are paired with semantic units expressed in a logical form language which is also a lexicalized TAG. The correspondences are stated as tree pairings with some of the nodes of the paired trees possibly linked. The following are examples of such pairings<sup>9</sup>:

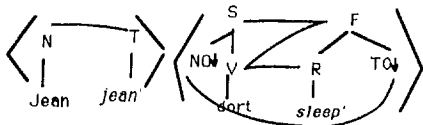


Figure 3. Pairing with semantic trees for *Jean* and *NO dormir*

The links between syntactic and semantic nodes mean that an operation at one node in the syntactic tree has an equivalent combination at the linked node in the semantic tree (and vice versa). More precisely, the semantic interpretation of a sentence is built "synchronously" with its syntactic derivation, by choosing a pair of elementary trees (a syntactic one: T1, a semantic one: L1) from the grammar and repeatedly updating it as follows (Shieber and Schabes 1990):

- choose a link between two nodes n1 and n2 in the pair (T1, L1),
- choose another pair of trees (T2 and L2), the root nodes of which match the respective categories of the two nodes chosen above,
- form a new pair by combining T2 at node n1 and L2 at node n2 and remove the link (if T2 and L2 have links, these are preserved in the result).

The definition allows for the operations performed at each node to differ, one being an adjunction and the other a substitution for example. It also allows for a node to be linked to several other nodes: in this case, only the "consumed" link is removed (the others are preserved and

we adopt here the convention that they are "pushed" upwards in the case of an adjunction)<sup>10</sup>.

Since multiple links are allowed, one syntactic node may be linked to several semantic nodes, and conversely, one semantic node to several syntactic nodes. This allows for the ability of a given syntactic element to play different semantic roles and for different syntactic realizations of the same semantic role. For example, as explained by Abeillé and Schabes 1990, this formalism naturally accounts for the adverbial status of adjectives inserted into idiomatic expressions: *Jean a pris une vraie veste* = *Jean a vraiment pris une veste* (Jean has really come a cropper). We want to show here that, provided it is taken as part of the well-formedness conditions of the grammar, it also accounts for the properties of French pronominal clitics.

**2.2. Augmenting TAG wellformedness conditions with synchronicity** So far, well-formed sentences of a TAL have been defined in the syntax only. In this respect, an input sentence is accepted by the TAG parser iff it obeys the following conditions:

- it can be derived from an initial tree rooted in S,
- all features unify at each node in the corresponding derived tree.

It is however possible to allow for the definition of the well-formed sentences of the language to be given jointly by the syntax and by the semantics, so that among the ill-formed sentences certain will be accepted by the syntax but rejected by the semantic rules if they cannot assign any interpretation to them. Such semantic filtering is not without history (Cf Pullum 1984) but it seems especially fruitful with the Synchronous TAG formalism because:

- the syntax and the semantics use exactly the same formalism,
- the syntactic and semantic derivations are necessarily built up in parallel.

The following well-formedness constraint is thus added to the parser: a sentence is acceptable iff it has at least one valid semantic interpretation built "synchronously" with it. By valid semantic interpretation, we mean that:

- it can be derived from an initial semantic tree,
- all features unify at each node in the corresponding derived semantic tree.

Several linguists have also suggested such semantic filtering for cases usually thought of as more syntactic (e.g. Sag and Klein 1982). The purpose of this paper is to advocate this device in various cases which all involve the syntax-semantic interface among which French pronominal clitics.

**2.3. French pronominal clitics with Synchronous TAGs** We rely on the existing elementary trees in the grammar to which we add substitution nodes for all possible clitics. Both clitics and corresponding NP, AP or PP nodes are optional in the syntax, their alternate realization is triggered by the

7. See Shieber and Schabes 1991 for some arguments for the use of trees instead of flat structures in Logical Form, and for the use of adjunction as an alternative to substitution in LF.

8. We refer the reader to Schabes et al. 1988 for more details on LexTAGs. Suffice here to say that the TAG elementary trees are systematically associated with lexical "heads" anchoring them in the lexicon (and required not to be empty) and are combined together either by substitution or adjunction. Feature structures are also associated at the nodes of the elementary trees and constraints on combining trees can be expressed in terms of success or failure of unification (Vijay-Shanker and Joshi 1988). Not all features are shown here.

9. The non terminal names in the semantic TAG are mnemonic for Formula, Term and Relation. Only the relevant links are shown.

10. We refer the reader to Shieber and Schabes 1990 for formal aspects of Synchronous TAGs (which are more powerful than TAGs).

associated semantic representation. We show how we get the following distribution:

- (7a) *Jean intéresse Marie* (Jean is interesting for Marie)
- (7b) *Jean l'intéresse*
- (7c) \**Jean intéresse*
- (7d) \**Jean l'intéresse Marie*

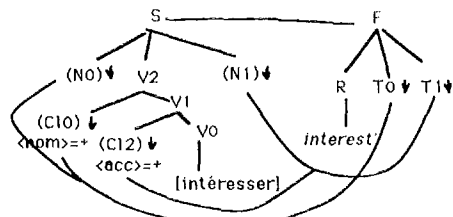


Figure 4. Elementary tree pairing for *NO intéresseur NI*. Both NP arguments of *intéresser* are linked with the corresponding Terms in the semantic tree. But the Clitic nodes are also linked to those Terms. They are thus prevented cooccur since only one substitution is allowed for each Term on the semantic side. Sentence (7b) is thus derived as follows:

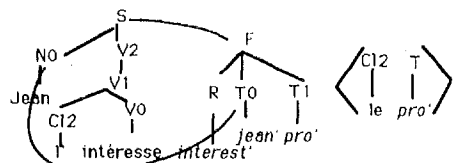


Figure 5. Synchronous derived trees for *Jean l'intéresse*. The four sentences (7a)-(7d) are allowed in the syntax but only sentences (7a) and (7b) are associated a synchronous interpretation. No interpretation is possible for sentence (7c) because its derived semantic tree is incomplete: the T1 is obligatory in the semantics. No interpretation is possible for sentence (7d) because whenever the clitic or the NP tree is substituted, it synchronously fills the T1 term and prevents the other to be substituted.

A motivation for treating cases (7c) and (7d) (i.e. cases of argument missing or cooccurrence between clitics and full phrases) as "semantically" deviant is that it seems to be always possible to construct a context in which they could be accepted<sup>11</sup>. We thus consider all the argument nodes to be optional and compatible in the syntax, their realization will be incompatible in the semantics (and might be obligatory if the semantic representation specifies so).

Handling clitic corresponding to PPs is now straightforward, as shown in the following pairing:

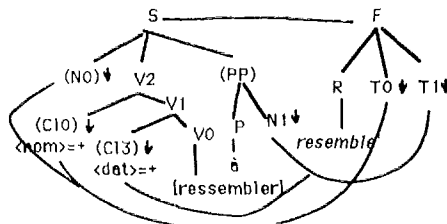


Figure 6. Elementary tree pairing for *NO ressembler à NI*. Notice that although N1 is the argument of the verb, it is the PP as a whole which is marked as optional (and will be prevented to cooccur with a dative clitic). The same result could be achieved if one considers the clitics to be adjoined (instead of substituted) on the syntactic side but this will necessitate a richer feature system to check clitic ordering and compatibility (see Abcillé 1991a for a previous account along these lines). In order to keep the feature system as simple as possible, we provide in fact nodes for all possible clitic realization (argumental or adverbial ones) in the corresponding elementary trees. The complete tree for a transitive verb like *voir* is thus the following (with clitic numbering as in section 1.1)<sup>12</sup>:

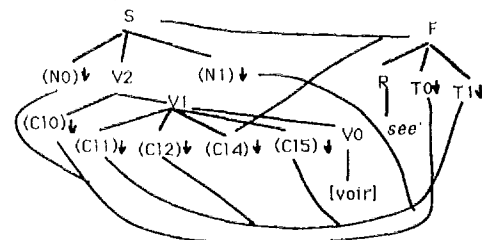


Figure 7. Elementary trees for *NO voir NI*. We will thus get *Jean y en voit, Il se voit* etc...

**2.4. Ambiguity and haplogy** Ambiguities are provided by the multiple links coming out of a clitic node: *en* for example can correspond to an accusative or a genitive complement, *y* can be a dative complement or a locative adjunct... If one takes a verb with an optional complement (such as *songer*) and the ambiguous clitic *y*, sentence (8a) below is parsed as ambiguous whereas only the adverbial reading is selected for *y* in sentence (8b)<sup>13</sup>:

(8a) *Jean y songe.* (1. Jean is dreaming there/ 2. Jean is dreaming about this)

(8b) *Jean y songe à ses soucis.* (Jean is dreaming there about his problems)

11. In fact (7d) is OK in spoken French, which can be shown to exhibit "clitic doubling", and (7c) may be improved as in: ? *Si Jean n'est pas intéressant c'est parce qu'il ne sait pas intéresseur.*

12. See Miller 1991 for arguments for having the subject clitics separated from the complements ones.  
13. Contrary to Miller 1991, we do not consider that (8a) has a third "haplogy" interpretation (Jean is dreaming there about it), nor do sentences such as *Jean en remplit.* (Jean is filling some with it)

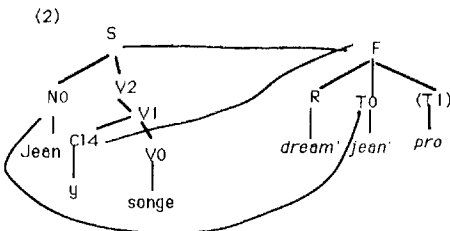
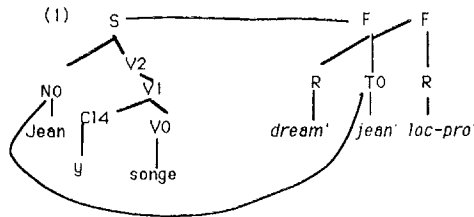


Figure 8. Two possible derivations for *Jean y songe*  
 Notice that sentences (1b) and (2b) above are not generated since there is only one position for *y* and one for *en* in the syntax.

**2.5. Further constraints** Clitic insertion is ruled out in extraction contexts:

- (9) *Je sais qui Jean regarde* (I know who Jean is looking at)
- (9a) \**Je sais qui Jean le regarde*
- (9b) *C'est Marie que Jean regarde* (It is Marie that Jean is looking at)
- (9c) \**C'est Marie que Jean la regarde*

In these constructions, which correspond to distinct elementary syntactic trees (in the tree family of their head) the obligatory syntactic realization of the extracted element naturally prevents the substitution of the corresponding clitic.

This representation is also directly suitable for marking various constraints, e.g. structural ones (ruling out *en* direct object for PP complements) or lexical ones (verbs which forbid cliticization of their complement such as *penser à Nhum* or *aimer que P*). As for feature equations, certain links are structurally defined as part of the elementary trees regardless of their lexical head (and there will be no link between *en* and the T complement node in the tree family for verbs with a PP complement), other links are brought by the lexical head and only certain verbs with a PP complement (*ressembler* but not *penser*) will define a link between the dative Clitic and their T complement node.

We now show how the representation sketched above naturally extends to some cases of so-called "non local" dependencies and to cases of cooccurrence between the clitic and the corresponding argument.

**2.6. Locality constraints and non local dependencies.** As noted by Shieber and Schabes 1990, locality constraints are inherent to the formalism of Synchronous TAGs. Contrary to Miller 1991, who runs the risk of allowing too many non local dependencies

with the FFP, we thus do not need to add specific locality constraints.  
 Notice first that some "non local" dependencies in a PSG are treated locally in a TAG framework. Examples of these are sentences with raising verbs (adjoining to V1) or auxiliaries (adjoining to V0) following the word order: *Jean peut le voir, Jean l'a vu*. Adjoining a raising or an auxiliary verb only updates the links coming out of the V1 or V0 node and does not interfere with the links of the clitics. We straightforwardly get: *Jean aurait été aimé de Marie* = *Jean en aurait été aimé* (Jean would have been loved by Marie). The agreement equations are the following (with *o-agr* for object agreement):

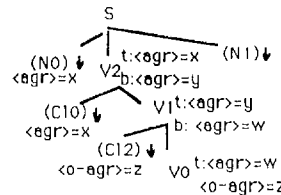


Figure 9. Agreement pattern for transitive verbs  
 Cliticization of the complement of an adjective is directly allowed in copular constructions: *Jean est fidèle à Marie* = *Jean lui est fidèle* (Jean is faithful to Marie). Copular constructions are (independently) treated as extended elementary trees in the TAG grammar, with the adjective co-heading the elementary tree with the verb and the clitic and PP(de) complement nodes belonging to the same tree<sup>14</sup>.

However, such cliticization is ruled out for modifying adjectives. Sentences such as: *J'ai rencontré une fille fidèle à Jean* = \**Je lui ai rencontré une fille fidèle* (I met a girl faithful to Jean) are not generated since the T node corresponding to the complement of the adjective does not belong to the same semantic tree as that of *N0 rencontrer N1* with which the clitic tree must be combined.

The same "local" treatment holds for cliticization of complements of nouns. It is allowed in light-verb constructions such as: *Jean a fait le portrait de Marie* = *Il en a fait le portrait* (Jean made a picture of Marie), which are represented as extended elementary trees with the light verb and the predicate nominal co-heading the structure. It is ruled out by our treatment when the NP is in a productive argument position<sup>15</sup>.

14. There are differences in acceptability for cliticization with verbs taking adjectival arguments: *Je lui (sais + ?\*trouve) Jean fidèle*. (I know/find Jean faithful to her). The difference is the same for extraction out of the AP: *A qui (sais+ ?\*trouve)-tu Jean fidèle ?* and is accounted for by different syntactic elementary trees (an extended one for *savoir*, one with the AP to be substituted for *trouver*).

15. This is obviously too strong a constraint since there are cases where the clitic corresponds to a complement of a noun at an arbitrary level of embedding, such as: *Le bord de la couverture de ce livre est déchiré* = *Le bord de la couverture en est déchiré* (The corner of the cover of the book is torn out).

en, which allows a direct object with a null head-noun: *Jean achète deux pommes* = *Jean en achète deux* (Jean buys two apples). In such cases, the determiner heads a syntactic N initial tree but its semantic tree is an auxiliary T tree which adjoins to the T node filled by the clitic. We also account for cases where the dative (human) clitic is semantically equivalent to a possessive, a construction typical of nouns of inalienable possession (such as body parts) combined with certain action verbs: *Ce docteur soigne les dents de Jean* = *Ce docteur lui soigne les dents* (This doctor treats Jean's teeth). Here, the clitic *lui* will be paired with an auxiliary T tree (as that for possessive determiners) and its node will be linked with the T node of the direct complement (for the verbs allowing this); it may thus cooccur with the NP complement.

For lack of space, we do not develop here "clitic climbing" in causative constructions which require either multicomponent trees on the semantic side or lexical rules adding causative double-headed elementary trees to the existing tree families.

**2.7. Cooccurrence between clitics and full phrases** Such cooccurrences are exhibited by inverted contexts such as: *Qui Jean voit-il ?* Such inverted clitics are represented in the syntax as auxiliary trees which trigger an inversion feature (Abeillé 1991a) and adjoin to the right of the inflected verbal form. On the semantic side, they are represented as ambiguous: they are associated with Terms and may thus alternate with non inverted clitics or NP subjects (*Qui voit-il ?*), provided the verbal nodes are linked with the subject T node. But they are also associated with auxiliary trees adjoining a question marking at the top F node (and thus allowed to cooccur with NP subjects).

Other cooccurrences are exhibited by dislocated constructions such as: *Jean, Marie l'a vu* or *Marie l'a vu, Jean* (Jean, Marie saw) which tend to generalize in spoken language. Right dislocation for complement clitics can be accounted for straightforwardly with the existing elementary trees if one allows for an alternative semantic representation of the clitics, namely an auxiliary emphatic semantic tree (adjoining to the Term node already filled by the NP) instead of a full Term<sup>16</sup>:

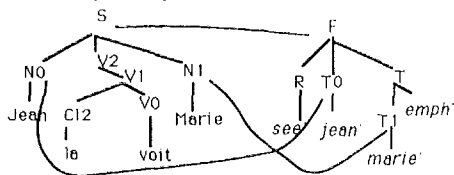


Figure 10. Derivation of a right dislocated construction. Multiple dislocations are thus allowed: *Je le lui ai confié, l'argent, à Jean* (I gave Jean the money) as well as "median" ones: *Jean l'a confié, l'argent, à Marie*. We do not consider here left dislocations which have different

16. This alternative representation of clitic pronouns as semantically vacuous is similarly used by Grimshaw 1982 for Spanish clitic doubling (optional feature Pred in the clitic entry).

syntactic properties and for which a purely semantic principle seems necessary in order to bind the preposed NP (Cf Hirschbuhler 1975, Fradin 1988).

**3. Extensions**

We show how the synchronous TAG framework naturally handles other cases of discrepancies between syntactic attachment and semantic role, for various non canonical configurations, while keeping the semantic component very simple and strictly compositional.

**3.1 "Argument-adjuncts"** It has often been noted that syntactic adjuncts may behave semantically as arguments<sup>17</sup>. These adjuncts may be in complementary distribution with arguments, such as possessive determiners and PP(de) complements: *la venue de Jean/ sa venue* (Jean's coming, his coming) vs \**la venue/ \*sa venue de Jean*. Synchronous TAGs naturally allow to represent such possessives as auxiliary trees in the syntax and as initial trees in the semantics. For example, the above distribution with the noun *venue* is accounted for by linking the attachment point of the determiner (the top N) and the substitution node of the PP complement to one and the same substitution node in the corresponding semantic tree<sup>18</sup>:

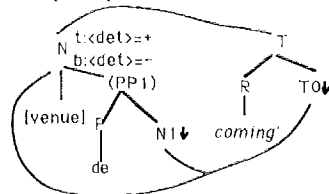


Figure 11. Elementary trees for *venue*. The phrase \**la venue* is disallowed because substitution is obligatory at T0 and the definite article *la* is not paired with a Term initial tree. Other such alternations involve "relational" adjectives such as *voyage présidentiel/voyage du président* and are handled similarly (with the relational adjective paired with a semantic term).

**3.2. "Extended" adjuncts** It has also been noticed that adjuncts may have a semantic scope well beyond their syntactic attachment point. For example, raising verbs, negative adverbials, quantifiers all have a semantic sentential scope although they attach to verbal or nominal items. These discrepancies are easily handled with Synchronous TAGs provided corresponding links are added to the elementary tree pairs (e.g. between V1 and the F node for raising verbs; Cf Abeillé 1991b).

Conversely, we can handle cases of "narrow" scope such as extraposed relative clauses which attach to S although they semantically modify an NP: the syntactic S tree of the relative clause is paired with an auxiliary T rooted tree in the semantics which adjoins to the T node

17. We recall that in TAGs, there is a structural distinction between modifiers (which are adjoined) and complements (which are substituted).

18. For some linguistic arguments for representing determiners as syntactic adjuncts, see for example Abeillé 1991a.

corresponding to the modified NP (provided S nodes are linked with the argument T nodes)<sup>19</sup>.

### Conclusion

We have shown how some non canonical arguments can be naturally handled using the formalism of Synchronous TAGs, provided the syntax-semantic synchronicity is incorporated as a well-formedness condition in the grammar. We have applied this treatment to French pronominal clitics and handle their complementary distribution with complements, without increasing the number of elementary trees in the grammar. Thanks to the extended domain of locality of TAGs, their locality constraints are handled (similarly as subadjacency) without specific stipulations. We also handle cases of non local dependencies, provided one adds alternative semantic representation for ambiguous clitic complements. The same idea can be extended to other cases of mismatches between syntactic attachment and semantic role, such as "extended" adverb scoping or extraposition.

### References

- A. Abeillé 1988. A lexicalized Tree Adjoining grammar for French, *Tech. report*, Univ. of Pennsylvania, Philadelphia.
- A. Abeillé, K. Bishop, S. Cote, Y. Schabes, 1990. "A Lexicalized Tree Adjoining Grammar for English", *Tech. Report*, Univ. of Pennsylvania, Philadelphia.
- A. Abeillé, Y. Schabes, 1990. "Non compositional discontinuous constituents in TAG", *Symposium on discontinuous constituency*, Tilburg (forthcoming van Horck, Wiesjma (eds), Mouton de Gruyter).
- A. Abeillé, Y. Schabes, A. Joshi, 1990. "Using Lexicalized TAGs for machine translation", *13th Coling*, Helsinki.
- A. Abeillé, 1991a. *Une grammaire lexicalisée d'arbres adjoints pour le français: application à l'analyse automatique*, Thèse de Doctorat, Université Paris 7. (forthcoming Editions du CNRS, Paris)
- A. Abeillé 1991b. "Raising verbs in Tree Adjoining grammar", *26th Linguistik Kolloquium*, Poznan (Vetulani, Darski (eds), Springer Verlag).
- A. Abeillé, 1992. "Sur la notion de syntagme verbal en français", *Linguisticae Investigationes*, XVI:2.
- K. Baschung, G. Bès, A. Corluy, T. Guillotin 1987. "Auxiliaries and clitics in French UCG", *3rd European ACL*, Copenhagen.
- T. Becker, 1990. "Metarules for TAGs", *1st TAG Workshop*, Sarrebrück.
- T. Becker, A. Joshi, O. Rambow 1991. "Long-distance scrambling and Tree Adjoining Grammars", *5th European ACL*, Berlin.
- G. Bès 1988. "Clitiques et constructions topicalisées dans une grammaire GPSG du français", *Lexique*, n°6, Lille.
- G. Bès, C. Gardent, 1989. "French Order without order", *4th European ACL*, Manchester.
- J. Grimshaw, 1982. "On the lexical representation of romance reflexive clitics", in Bresnan, *The mental representation of grammatical relations*, MIT Press.
- B. Fradin, 1988. "Approches des constructions à détachement : la reprise interne", *Langue française*, 78.
- M. Gross, 1968. *Syntaxe du verbe*, Cantilène, Paris.
- M. Gross 1989. "The use of finite automata in the lexical representation of natural language" in Gross & Perrin (eds), *Electronic dictionaries and automata in Computational Linguistics*, Springer Verlag.
- P. Hirschbuhler, 1975. "A non transformational source for left hand NPs", *Linguistic Inquiry*, 4.
- A. Joshi, 1987a. "An introduction to TAGs" in Manaster Ramer, *The Mathematics of language*, Benjamins.
- A. Joshi, 1987b. "The relevance of TAG to generation", in Kempen (ed), *Natural Language Generation*, Dordrecht.
- R. Kayne, 1975. *French syntax*, MIT Press.
- A. Kroch, A. Joshi, 1986. "Analysing Extraposition in a TAG", in Huck and Ojeda (eds): *Discontinuous Constituents*, Syntax and Semantics vol. 20, Academic Press.
- P. Miller, 1991. *Clitics and constituents in Phrase structure grammar*, PhD Thesis, Univ. Utrecht. (forthcoming Garland, New York)
- P. Paroubek, Y. Schabes, A. Joshi, 1992. "XTAG: a graphical Workbench for developing TAGs", *3rd Conf. on Applied NL Processing*, Trento, Italy.
- G. Pullum, 1984. "Syntactic and Semantic parsability", *10th COLING*, Stanford.
- I. Sag, E. Klein, 1982. "The syntax and semantics of English expletive pronoun constructions", in Barlow et al. (eds) *Developments in GPSG*, vol. 2, Indiana Ling. Club, pp. 92-136.
- Y. Schabes, A. Abeillé, A. Joshi, 1988. "Parsing strategies with lexicalized grammars: Tree adjoining grammars", *12th COLING*, Budapest.
- Y. Schabes, 1990. *Computational and Mathematical Properties of Lexicalized Grammars*, PhD Dissertation, Univ. of Pennsylvania, Philadelphia.
- S. Shieber, Y. Schabes, 1991. "Generation and Synchronous TAGs", *Computational Intelligence 7:4*.
- S. Shieber, Y. Schabes, 1990. "Synchronous Tree Adjoining Grammars", *13th COLING*, Helsinki.
- K. Vijay-Shanker, A. Joshi, 1988. "A Feature-based Tree Adjoining Grammar", *12th COLING*, Budapest.
- K. Vijay-Shanker, Y. Schabes, 1992. "Structure sharing in TAGs", *14th COLING*, Nantes.

<sup>19</sup>. This case has been handled by "local" MCTAGs (Kroch and Joshi 1986) with "empty" trees for coindexing the NP and the extraposed relative. Due to the inherent locality of Synchronous TAGs, the same effect of clause boundedness as "local" MC-TAGs is achieved.