Principles for a Computational Account of Adverb Syntax

Christopher Laenzlinger
Laboratoire d'analyse et de technologie du langage
Université de Genève

Septembre 1994

Adverbials are a rich and as yet relatively unexplored system, and therefore anything we say about them must be regarded as quite tentative

N. Chomsky. *Aspects* (p. 219)

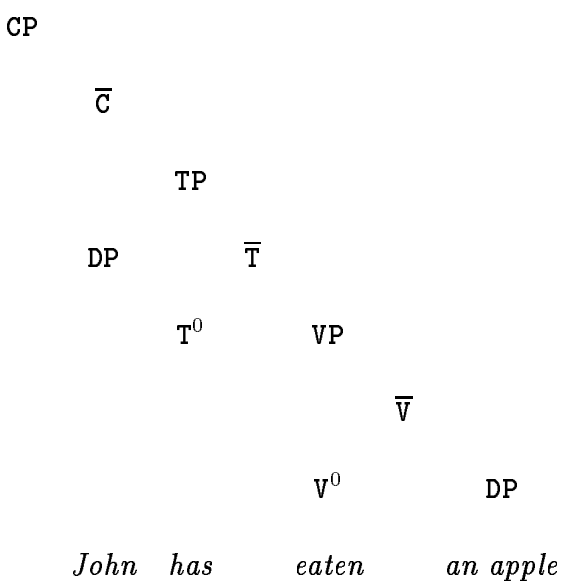
The incorporation of a treatment of adverbs in a computational model is rather difficult, because of the gaps still present in the theoretical specifications. For instance, the way in which adjunction interacts with phrase structure component is not clearly defined. The model of $\overline{\mathtt{X}}$-structure presented in Laenzlinger (1993) makes a step forward in delineating precisely the structural configuration in which adverbs can enter and their scopal properties. We will discuss the integration of this $\overline{\mathtt{X}}$ model into the GB–oriented parsing system IPS (Interactive Parsing System) developed at the LATL at the University of Geneva. There exists an English and French version of the parser (*cf.* Wehrli 1992; Laenzlinger & Wehrli 1991). We will also present parsing strategies for a treatment of adverbs based on the 'double–Spec' model.

The IPS system contains a module that links the lexical component to the syntactic component. It is the $\overline{\mathtt{X}}$ generator of phrase structures. The $\overline{\mathtt{X}}$ mechanism generates maximal projection of level 2 (XP) from head categories of level 0 ($\mathtt{X}^0$). Lexical categories correspond to N(oun), V(erb), A(djective), Adv(erb) and P(reposition), and functional categories to D(eterminer)[1], T(ense) and C(omplementizer). Three types of action operate on the $\overline{\mathtt{X}}$ module: phrasal projection, attachment to the left (attachment of specifiers), and attachment to the right (attachment of complements). Specifiers and complements are implemented as lists that can be empty. Attachments of complement are of two kinds: formal attachments, which are ruled by selection and require adjacency (selection of auxiliaries, of determiners), and non-formal attachments, which concern mainly subcategorized elements.

The clausal structure corresponding to the input string *John has eaten an apple* is parsed as (1), on the basis of the $\overline{\mathtt{X}}$ generator and lexical information:

[1]The $\overline{\mathtt{X}}$ format follows Abney's (1986) DP hypothesis, according to which the functional head $\mathtt{D}^0$ selects a lexical NP.

(1)



Adverbs project their own maximal category, AdvP. The attachment of an adverb is unproblematic when it immediately precedes the participle, as in (2a), or the adjective, as in (2b). In the former case, it is attached in Spec of VP, and, in the latter case, in Spec of AdjP. These Spec positions are typically adverb positions.

(2)a. John has **completely** eaten the apple.

b. a **very** good apple.

We know that the distributional range of adverbs is not restricted to the above Spec positions. Consider the case of modal adverbs in French. The adverb *probablement* ('probably') in its sentence reading may appear in clause initial, clause final and pre-participial positions, and also between the subject and the auxiliary, and between the verb and its object. The overall distributional range is illustrated in (3a-e).

(3)a. Probablement, Jean a lu le livre de Chomsky.

b. Jean a lu le livre de Chomsky, probablement.

c. Jean a probablement lu le livre de Chomsky.

d. Jean, probablement, a lu le livre de Chomsky.

e. Jean a lu probablement le livre de Chomsky.

The parsing of the adverb in (3a) and (3c) does not induce problems of attachment. In the former case, the adverb is attached as Spec of VP; the VP node is a legitimate attachment site for adverbs. In the latter case, the sentence initial position of the adverb corresponds to Spec of CP.[2] Given the clause structure in (1), all the other cases in (3) involve problems of attachment. No specifier position is directly accessible to the attachment of the adverb. Let us proceed case by case. The modal adverb in (3b) is sentence final. The standard model of the $\overline{X}$ schema adopted for French and English does not allow the right–branching

[2]Such an assumption concords with our analysis of modal adverbial distribution, according to which the adverb must be in Spec–head relation to C^0, at the latest at LF.

of Spec positions. The adverb can be inserted in a complement position. Specifiers and complements correspond to lists of constituents. The presence of an argument and an adjunct in the same list of Complements is not prohibited by the $\overline{X}$ generator (no binary branching constraint). However, the lack of structural distinction between an argument and an adjunct is inadequate for the interpretative modules (θ-module, chain formation module, binding module). An argument behaves differently from an adjunct with respect to these modules (an object argument contrary to an adverb is assigned a θ-role, a Case and is lexically governed). Moreover, the scope of adverbials must not be totally dissociated from their distributional properties. In other words, the syntax and the semantics of adverbials ought not to be assigned fully independent analyses/parses.

On the assumption that the syntactic component – the pre-semantic component – associates the distribution of adverbials with their scope properties, an additional structural mechanism is needed to treat adjunct attachments. Adjunction is a possible phrasal tool, but its implementation mirrors the formal problems we discussed in Laenzlinger (1993). Either the $\overline{X}$ system is expanded so as to compose with adjunction positions (*cf.* Dorr 1993 for a similar approach) or adjunction is a structural mechanism independent of the $\overline{X}$ system which applies to well-formed completed phrasal structures.

On computational grounds, both approaches lead to practical inadequacies. The iteration of the adjunction process is problematic: two potential maximal adjunction positions are accessible to adjunct attachment at the level of any maximal node. Such an iteration entails unwanted attachment alternatives in the course of parsing. Consider the sentence final position of an adverb as in (4):

(4) [$_{CP}$ [$_{TP}$ Jean lira [$_{VP}$ [$_{DP}$ ce livre]]]]... certainement.
'John will certainly read this book'

The set of potential attachment sites for the adverb in (4) contains any maximal node on the right edge of the structure built until then, namely CP, TP, VP, DP. The range of possible attachment sites can be, however, filtered out on the basis of the lexico–semantic content of the adverb.

Identical technical difficulties in parsing arise whether we adopt an $\overline{X}$ integrated theory of adjunction or a theory of adjunction independent of the $\overline{X}$ system. Attachment alternatives should be reduced to the minimum, for their computation is particularly costly. We will show that even heuristics for licensing attachment adjunction sites cannot help in resolving the problematic cases of adverbial attachment.

Before entering into the details of adverb parsing, we briefly sketch out the specificities of the IPS system. The parsing algorithm is data–driven with a left–to–right reading of the input string. It combines a bottom up approach (phrase structure building) with a top down filtering (thematic checking, chain formation). Alternatives are treated in parallel. Structure building is incremental: incoming material is integrated into maximally well–formed structures. To avoid some problems inherent in bottom up algorithms, notably concerning the possible generation of multiple locally well-formed structures which prove subsequently to be incompatible with the overall structure, Wehrli (1992) develops a 'right corner' strategy based on the following principles (p. 873):

- constituents can be attached as soon as they have been hypothesized;
- attach incoming items to maximally expanded constituents in the left contexts;
- constituents specify a list of active attachment sites on their right edge;

- all attachments are considered in parallel;

Suppose that adjunction is the appropriate device for adverb attachment. The adjunction process (whether $\overline{\mathtt{X}}$ internal or $\overline{\mathtt{X}}$ external) reduplicates a maximal node that serves as a possible attachment site for adjuncts. The 'right corner' strategy is relevant to adverbial attachments to the right. Adverbial attachments to the left follow the usual bottom up algorithm. Once an adverb has been read and its category projected, the parser attempts to attach it by means of the standard $\overline{\mathtt{X}}$ procedure as a specifier, typically as `Spec` of `VP` or `Spec` of `AP`. If no `Spec` attachment works out, the system calls for the adjunction procedure. If adjunction is an $\overline{\mathtt{X}}$ integrated procedure, the system takes into account the set of possible adjunction sites in the right context. If adjunction is implemented as a post $\overline{\mathtt{X}}$ procedure, the system creates a specific adjunction site in the right context.

The parsing of a sentence-initial temporal adverbial, as in (5), resorts to adjunction sites. As the adverb is the first item of the input string to parse, its right context will constitute its domain of attachment. Once the clausal structure has been built from the verb *mange* with *Jean* attached as a specifier of `TP` (subject position), the parser attempts to attach the adverb as a specifier. The range of possible `Spec` attachment sites is `CP`, `TP` and `DP` (subject). `Spec` of `CP` is usually the position of *wh*–phrases; `Spec` of `TP` is typically the position of the subject; `Spec` of `DP` is the position of phrasal adjuncts. As none of these `Spec` positions is suitable for the adverb attachment, the parser considers the set of active adjunction sites in the right context, namely the expanded segmental nodes `CP`, `TP` and `DP` (subject). These are filtered out on the basis of the lexico–semantic feature of the adverb. Only `TP` is a licensed attachment site for a temporal adjunct.

(5) Maintenant Jean mange.
'Now John is eating'

Left-attachment of adjuncts follows the 'right corner' strategy and is restricted to adjunction sites. Indeed, attachment to $\overline{\mathtt{X}}$ nodes (*i.e.* in `Compl` of $\mathtt{X}^0$) applies only to selected and subcategorized complements. Consider the sentence final adverb in (6):

(6) John has read the book entirely.

The set of active nodes in the left context which function as possible attachment sites for the adverbial constituent corresponds to `CP`, `TP`, `VP`, `DP` (object), `NP`. The set is filtered out on the basis of the scope properties of the adverb (encoded in terms of lexical features). As *entirely* is a `VP` adverb, it must adjoin to `VP`.

The parsing of intrasentential adverbials is less straightfoward than that of peripheral adverbials. First, we know that some adverbs may occupy a position at S–structure which is embedded in the domain of their scope. The case of `VP` internal occurrence of modal adverbs is instanced below:

(7) John has probably eaten an apple.

If the scope of an adverb is calculated as a function of its structural position, then a modal adverb must be adjoined to the clausal node, at the latest at LF. Following the proposal of Q-raising of such adverbial operators, we assume that the modal adverb occupies its scope

position at the semantic level. The semantic parsing of the sentence in (7) produces a LF–structure that depends on its well–formed S–structure (output of syntactic parsing).[3] [4] In the S–structure of the sentence, the modal adverb is attached as Spec of VP. In its LF–structure the modal adverb is adjoined to CP. Thus, Logical Form is the ultimate level for expressing the scope of (quantificational) adverbials. Those adjuncts that cannot Q-raise (qualificational adverbs) must express their scope at S–structure.

Adjunction is not always a sufficient tool to treat adverbial attachment. To parse the sentences below, we must stipulate some additional transformational operation.

(8)a. Jean, probablement, a lu le livre de Chomsky.

b. Jean a lu probablement le livre de Chomsky. (wide scope)

Given the clause structure in (1), the application of the adjunction procedure to the modal adverb is blocked by the subject argument in (8a) and by the object argument in (8b). The former must be attached as a specifier of TP and the latter as a complement of $\overline{\text{V}}$. The formal analysis of (8a) that we have proposed is that the order Subj Adv Aux derives from subject topicalization (in Spec of TopP). The sequence V_{part} Adv Obj in (8b) is obtained by participial verb movement. None of these analyses is compatible with a phrasal system that limits the functional categories of the clause to CP and TP. The limitation of phrasal projections understandably obeys effeciency constraints on parsing ('avoid unbound or non selected empty categories').

Extraposition of arguments, as in (8), requires some strategies of parsing. Let us start with the extraposition of the subject in (8a). The parser proceeds as follows: it reads the first element of the input string, the word *Jean*, and projects a DP constituent. Then, it reads the adverb and projects an AdvP constituent. As no attachment can be performed so far, the two constituents are put in a stack. The reading of the auxiliary *a* gives rise to the projection of a TP and of a CP by extension. The top element on the stack to be considered is the adverbial constituent. Its adjunction site is licensed as CP (propositional adverb). The second element of the stack is a DP. The only hypothesis of analysis to pursue is also to adjoin the DP constituent to CP. Provided that the procedure of adjunction relying on the rule in (9) may apply recursively, more than one adjunction site is possibly created on the same maximal node, here CP.

(9) XP $\longrightarrow$ (YP) XP (YP)

The interpretation of the CP–adjoined DP is successfully achieved by the formation of a chain linking the adjoined DP to the empty Spec–TP position.

The case of object extraposition in (8b) requires rather subtle strategies. At the point of reading the adverb, the parser has built up the left context corresponding to the structure (10):

(10) [$_{\text{CP}}$ [$_{\text{TP}}$ Jean a] [$_{\text{VP}}$ lu]]

[3] Such a modularization of grammatical analyses mirrors the T–model of Chomsky (1981): the syntactic component feeds the semantic component.

[4] The IPS system does not apply semantic parsing for the time being. The parsing of semantic structures is still at the planning stage.

The attachment site for *probablement* is licensed as CP (lexical selection). The obligatory occurrence of (clausal) complement extraposition in French and in English, as in (11), motivates a strategy of parsing according to which the attachment of postverbal adjuncts ought to be delayed until the end of the sentence has been reached, *i.e* until the last element of the sentence has been read and parsed. This is the only situation in which an incremental analysis cannot be maintained.

(11)a. Jean a dit hier que Marie viendra.

b. Mary told yesterday that Mary will come.

The parsing of (8b) follows two hypotheses of analysis at the VP level. The verb *lu* projects either an intransitive VP or a transitive one. Only the second hypothesis turns out to be compatible with the subsequent input data. The last constituent of the sentence is a DP constituent, *le livre de Chomsky*. The categorial matching between the extraposed constituent and the possible subcategorized complements of the verb (here a DP) predicts the presence of an empty category in Compl of $\overline{V}$: a chain is formed with the extraposed DP. This is a relatively simple way of dealing with chains whose head follows their foot on surface. Thus, the delay strategy allows a treatment of traces preceding their antecedent (extraposition) in a system endowed with a left-to-right reading of the input string.

According to the adjunction phrasal rule in (9), more than one constituent may be successively attached to segments of the same maximal category (iterative application of the rule). The rule recursively applies to multiple adverbial attachments, as in (12):

(12)a. Yesterday, at ten, the President died.

b. Jean a lu courageusement entièrement le livre de Chomsky.

The attachment of the temporal adverbials in (12a) is performed at the TP level, and the attachment of the manner adverbs in (12b) at the VP level. A possible way of implementing the potential iteration of adjoined positions is to use a list. The order in the list must reflect the hierarchy of positions in the structure. In the case of multiple left adjunction, the first element of the head of the list is hierarchically higher within the projection than the other elements of the list. This implies that the first element precedes the others in the surface string (linearly). In the case of multiple right adjunction, the tail of the list is hierarchically higher than the other elements of the list. This implies that the last element of the list follows the others in the surface string. With respect to the occurrence of multiple adverbs, the list is ordered so as to express the hierarchical constraints on adverbs, while preserving the scope relations of each adverb.

A precompilation of possible adjunction sites may be executed in order to limit the number of potential attachments by adjunction. For instance, no adjunction is allowed to the node NP:

(13) *[$_{DP}$ the [$_{NP}$ probably [$_{NP}$ men]]]

The number of adjunctions per maximal category is theoretically unlimited. But semantic constraints and psycholinguistic consideration (understanding human faculties) impose reasonable restrictions on the number of adjoined elements (adjuncts). Note that, if no semantic incoherence and no problem of understanding arise, the number of adverbials may be relatively high:

(14) In France, near Paris, in a garden, yesterday, after dinner, at ten, I smoked a cigar.

Recall that the 'double–Spec' model of $\overline{\mathtt{X}}$ theory we propose in the present research is an attempt to remedy the lack of formalization inherent in the Theory of Adjunction (notably with licensing conditions, recursivity, $\overline{\mathtt{X}}$ relatedness, segmentation, barrierhood, *etc*). The 'double–Spec' model leads to a more adequate treatment of adverb syntax than the Theory of Adjunction does. We will show here that the integration of this model into the $\overline{\mathtt{X}}$ generator of the IPS system is a feasible task, and that it fits a treatment of adverb attachments. As specifiers correspond to lists (possibly empty), we propose that the Spec list has a binary structure ordered as [Spec–$\overline{\mathtt{A}}$, Spec]. The $\overline{\mathtt{X}}$ generator which relies on the two rules in (15) imposes the linear precedence of an $\overline{\mathtt{A}}$–element over an A–element contained in the same projection.

(15) $\mathtt{XP} \longrightarrow \mathtt{Spec\text{-}}\overline{\mathtt{A}} \quad \mathtt{Spec\text{-}A} \quad \overline{\mathtt{X}}$
$\overline{\mathtt{X}} \longrightarrow \mathtt{X}^0 \quad \mathtt{Compl}$

For instance, the linear precedence of the modal adverb over the non-pronominal subject is expected in French complex inversion (*cf.* (16a)), and, similarly, the precedence of the temporal adverb over the subject in (16b)).

(16)a. [Peut-être Jean] viendra-t-il.
'Maybe John will come'

b. [Yesterday John] came.

The level of attachment of adverbs is (pre-)determined on the basis of lexical information. Their modification (*i.e.* scope) properties are encoded in the lexicon. A structural constraint guarantees their attachment to the projection of the head they modify (*i.e.* in the Spec-$\overline{\mathtt{A}}$ of that projection). This is exactly what the *Adv*–Criterion stated by Laenzlinger (1993) expresses. A manner adverb is predictably attached to VP, either on the left or on the right:

(17)a. Jean a [$_{\mathtt{VP}}$ péniblement [$_{\overline{\mathtt{V}}}$ travaillé.]]

b. Jean a [$_{\mathtt{VP}}$ [$_{\overline{\mathtt{V}}}$ travaillé] péniblement.]
'John worked painstakingly'

A temporal adverb undergoes either left–attachment or right–attachment to TP:

(18)a. [$_{\mathtt{TP}}$ Hier, Jean [$_{\overline{\mathtt{T}}}$ a travaillé]]

b. [$_{\mathtt{TP}}$ Jean [$_{\overline{\mathtt{T}}}$ a travaillé] hier]
'Yesterday John worked'

A modal adverb in its sentence final or sentence initial position is attached to CP. It may also be attached to the VP level at S–structure and Q-raises at LF, *i.e.* attached to CP semantically:

(19)a. [$_{\mathtt{CP}}$ Probablement, [$_{\mathtt{TP}}$ Jean a travaillé]]

b. [$_{\mathtt{CP}}$ [$_{\mathtt{TP}}$ Jean a travaillé], probablement]

c. [$_{CP}$ [$_{TP}$ Jean a [$_{VP}$ probablement [$_{\overline{V}}$ travaillé]]]]
'Probably John worked'

The treatment of aspectual adverbs is rather problematic, given the restricted set of clausal categories (CP, TP, VP) at disposal in the IPS parser. Nevertheless, the distribution of aspectual adverbs may be integrated into the TP/VP phrasal system. The adverb *souvent*, for instance, may be assigned two readings. If topicalized, as in (22a), it modifies the Event specifications. Otherwise, it is VP–related in that it modifies the aspectual specifications of the verb, as in (22b). In the former case, the parser should attach the adverb to TP, the head of which contains the Event specifications of the verb, and in the latter case to VP.[5]

(22)a. [$_{TP}$ Souvent, Jean [$_{\overline{T}}$ pleure]]
'Often John cries'

b. [$_{TP}$ Jean a [$_{VP}$ souvent [$_{\overline{V}}$ pleuré]]]
'John often cried'

The cases of multiple adjacent adverbs given in (23) deserve particular attention. The co-occurrence of an aspectual adverb with a manner adverb in Spec–$\overline{A}$ of VP, as in (23a), is not expected, nor is the co-occurrence of an aspectual adverb with a time adverb in Spec–$\overline{A}$ of TP, as in (23b) (they are not adverbs of the same semantic type for absorption).

(23)a. Jean a souvent péniblement ouvert la porte.
'John often painstakingly opened the door'

b. Hier, souvent, Jean a pris la parole.
'Yesterday often John opened the door'

Recall that the $\overline{X}$ model allows one Spec-A and one Spec–$\overline{A}$ per projection. Implemented as such, the model fails to treat cases like (23), especially because clausal projections are limited in number (for procedural efficiency). Assume that the number of Spec–$\overline{A}$ per projection is not limited to one. Then, the list of specifiers contains a unique Spec-A and possibly more than one Spec–$\overline{A}$. The order among Spec–$\overline{A}$ elements of the list is determined according to their respective scope properties. For instance, the parsing of (24) should fail, because it does not fit the hierarchical constraints on multiple adverbs.

(24)a. *Jean a péniblement souvent ouvert la porte.
'John painstakingly often opened the door'

[5]The difference in VP–adverb distribution between French and English, as illustrated in (20) and (21), relies on the surface position of the verb in these two languages. The Pollockian analysis stipulates on the basis of the intrasentential position of VP–adverbs (and of negation) that the verb raises to inflection (I^0) in French, while it remains in V^0 in English. The position of the verb determines the order V Adv in (20) and, inversely, Adv V in (21):

(20)a. Jean rencontre **souvent** Marie.

b. Jean quitta **soudainement/courageusement** la salle.

(21)a. John **often** meets Mary.

b. John **suddently/courageously** left the room.

b. *Jean a péniblement malheureusement ouvert la porte.
'John painstakingly unfortunately opened the door'

To increase the number of Spec–$\overline{\text{A}}$ per list may have unwanted effects on the range of parsing hypotheses. We have to impose specific constraints on the types of Spec–$\overline{\text{A}}$ elements that are allowed to co-occur in the same projection and on their respective order of occurrence.

Importantly, the phrasal rules stated in (15) do not account for the case of right-attachment of adverbs, as in (17b), (18b) and (19b), nor do they for the cases of complement extraposition, as those illustrated below:

(25)a. John told me yesterday that Mary cried.

b. Jean a vu hier le film dont tu m'as parlé.

On the assumption that clause final adverbs and extraposed constituents occupy a right–branching Spec–$\overline{\text{A}}$ position, the $\overline{\text{X}}$ generator must rely on the slightly modified rules in (26):

(26) XP $\longrightarrow$ Spec-$\overline{\text{A}}$ Spec-A $\overline{\text{X}}$
XP $\longrightarrow$ Spec-A $\overline{\text{X}}$ Spec-$\overline{\text{A}}$
$\overline{\text{X}}$ $\longrightarrow$ X^0 Compl

To compute the potential left-attachment/right-attachment of Spec–$\overline{\text{A}}$ is obviously a difficult task. Recall that specifiers are implemented as lists. If the two Specs are left–branching, the list is calculated on the basis of the order [Spec–$\overline{\text{A}}$, Spec]. But, if the Spec-$\overline{\text{A}}$ is right–branching (right–attachment), some strategy must be used so as to recover the content of the Spec list. We propose the strategy of delayed XP closure, which means that all maximal nodes remain active for attachment, until the end of the sentence has been reached. Consider an input string containing a time adverb between the participle and the nominal complement:

(27) Jean a mangé hier une pomme.
'John ate an apple yesterday'

Once the participle has been read, two hypotheses of analysis are pursued in parallel. The verb *mangé* projects either an intransitive structure (no DP expected) or a transitive one (a DP expected). The parsing of the time adverb leads to its attachment in the right–branching Spec–$\overline{\text{A}}$ of TP. The adverb is inserted into the list of Spec-TP elements. The subject is already part of this list, since it is attached to the left of TP. As the two specifiers are not linearly adjacent, the order within the list is irrelevant. Following the strategy of delayed XP closure, the parsing of the adverb does not induce the closure of the VP constituent. The structure that has been parsed so far is given in (28).

(28) [$_{\text{CP}}$ [$_{\text{TP}}$ Jean a [$_{\text{VP}}$ mangé]...] [$_{\text{AdvP}}$ hier]]

If no other element follows the adverb (*Jean a mangé hier*), the intransitive VP analysis works out. If a DP is parsed after the adverb, the intransitive option fails. The transitive analysis is pursued: an empty category in Compl of $\overline{\text{V}}$ is postulated in accordance with the subcategorization of the verb. The lexical DP *une pomme* is attached in the right–branching Spec–$\overline{\text{A}}$ of CP and forms a chain with the empty category in Compl of $\overline{\text{V}}$. The fully parsed structure is given below:

(29) [$_{\text{CP}}$ [$_{\text{TP}}$ Jean a [$_{\text{VP}}$ [$_{\overline{\text{V}}}$ mangé [$_{\text{DP}}$ e]$_i$]]] [$_{\text{AdvP}}$ hier] [$_{\text{DP}}$ une pomme]$_i$]