



HAL
open science

French missing object constructions

Gabrielle Aguila-Multner, Berthold Crysmann

► **To cite this version:**

Gabrielle Aguila-Multner, Berthold Crysmann. French missing object constructions. *Glossa: a journal of general linguistics* (2021-..), 2022, pp.1-35. 10.16995/glossa.6478 . halshs-03505490

HAL Id: halshs-03505490

<https://shs.hal.science/halshs-03505490>

Submitted on 31 Dec 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

French missing object constructions

Gabriel Aguila-Multner
Université de Paris, LLF, CNRS
gabriel.aguilamultner@gmail.com

Berthold Crysmann
Université de Paris, LLF, CNRS
berthold.crysmann@gmail.com

Abstract This paper discusses a class of French *à*-infinitival constructions, where the missing direct object corresponds to an external argument, either being the antecedent noun in an attributive use, or else a raised argument in a subject or object predication or in the *tough* construction. We investigate the internal and external properties of these constructions and show that (i) the construction displays passive-like properties and (ii) control and raising verbs may intervene between the marker *à* and the missing object verb, as shown on the basis of a corpus study. We observe that while the construction as a whole behaves like a passive where the erstwhile logical object ends up being promoted to external argument, the logical subject is still accessible for control, both from within the *à*-infinitive and from outside. Building on Grover (1995), we analyse these double subjects in a way similar to passives, involving lexical promotion of the direct object valency to subject, yet without concomitant subject demotion. Raising of the missing object as a secondary subject will make it available at the top of the construction, where it can serve as an external argument for modification or complementation. The present analysis thus captures the full set of *à*-infinitival missing object constructions in a unified fashion, capturing its passive-like properties and the extended domain of locality.

Keywords: French; missing object; *tough* construction; passive; infinitive; raising; control

French missing object constructions (henceforth MOCs) are infinitival phrases introduced by marker *à* that involve a dependency between a direct object (DO) that is locally missing, but that corresponds to an external argument that the construction modifies or predicates on. An example is given in (1), in which *à lire* is an instance of a MOC, its antecedent noun (*livre*) corresponding to the missing object of the verb *lire*. As can be seen from the translation, such constructions can (at least in some uses) find an English equivalent in a similar phrase headed by *to*.

- (1) un livre *à lire*
a book to read
'a book to read'

The MOC has been treated variously as a long distance extraction, as a bounded reduced relative and as a kind of passive. In this paper we borrow insights from the last two treatments and propose an analysis as an infinitival passive with an extended but bounded locality.

In a first section we will describe the external properties of the MOC, before turning to its internal properties in Section 2. In Section 3, we present new data relating to the internal properties in the form of a corpus study on the locality of the construction. Sections 4–5 present and discuss previous analyses, and our proposal is exposed in detail in Section 6.

1 Distribution

We find four distinct contexts that license the MOC in French: attributive, subject predication, object predication, and *tough* contexts.

1.1 Attribute

Probably the most wide-spread use of the MOC is the attributive one, illustrated by example (1), where the MOC modifies an antecedent noun that corresponds to the missing object. This property displays some superficial similarity to infinitival relatives (Abeillé et al. 1998), as in (3). However, the two constructions differ from each other in that the subject in the MOC can be realised internal to the infinitival clause by an optional *by*-phrase (see (2) and Section 2.1), an option not found with infinitival relatives.

- (2) un livre à lire par chaque parent à mon avis
 a book to read by every parent at my opinion
 ‘a book to read by every parent in my opinion’
- (3) un endroit où aller pour se détendre
 a place where go to relax
 ‘a place to go to relax’

1.2 Subject predication

Similar to other attributes in French, the construction can also be used predicatively:

- (4) Ce livre est à lire.
 this book is to read
 ‘This book is to be read.’

This use extends to other copular verbs, e.g. *sembler*, and does not find an equivalent in English, as the change to a passive in the translation demonstrates.

1.3 Object predication

Predicative use of the MOC is not limited to subject predication, but it extends to object predication as well. For example, it can be used with verbs like *avoir* as an object predicate:

- (5) J’ ai un livre à lire.
 I have a book to read
 ‘I have a book to read’ or ‘I have to read a book.’
- (6) Je l’ ai à lire.
 I 3SG.M have to read
 ‘I have to read it.’

At first glance, the sentence in (5) could simply be interpreted as an attributive use of the MOC attaching to the direct object’s head noun, with the meaning ‘I have a book to read’. This is in fact one of the possible analyses of (5), which is structurally ambiguous; the alternative structure can be revealed by pronominalisation of this DO in (6)

independently of the MOC, left in situ. Since such pronominalisation normally targets saturated phrases and not partial NPs, this is evidence of an object predication structure. The semantics of this object predication are slightly but noticeably different: the missing object, expressed as a direct object of *avoir*, is not interpreted as an argument of it – that is, no owning relation need exist between the subject and the book here, and the sentence could felicitously be uttered of a book that has yet to be bought or borrowed.

Most importantly, this object predication structure comes with an additional control relation where the subject of *avoir* ‘have’ is interpreted as the subject of the infinitive. With other verbs displaying this object predication pattern (*donner*, *laisser*, ...), however, the controller is not the matrix subject but an indirect object, as is the case in example (7).

- (7) Je le donnerai à lire aux étudiant-es.
 I 3SG.M will give to read to the students
 ‘I will give it for the students to read.’

This control phenomenon poses a challenge for existing accounts of the missing object construction, as any analysis will therefore need to expose pointers to both core arguments of a MOC to be accessed by the upstairs verb. The downstairs logical object will be realised as an upstairs NP complement, while the downstairs logical subject can be controlled by some argument of the upstairs verb: *avoir* exerts subject control, whereas *donner* exerts control by means of its indirect object.

1.4 Tough constructions

Finally, the MOC appears as the complement of so-called *tough* adjectives, its missing object being understood to correspond to the adjective’s external argument: this will be the head noun in the case of attributive use (8), or the subject in the case of predicative use (9).

- (8) un livre facile à lire
 a book easy to read
 ‘a book that is easy to read’
- (9) Ce livre est facile à lire.
 this book is easy to read
 ‘This book is easy to read.’

Similarly to English, *tough* predicates are not strictly limited to adjectives, but may include (idiomatic) nominals like *de la tarte* ‘piece of cake’ (10a) or verbal expressions like *prendre (du temps)* ‘to take (time)’, *coûter (une somme)* ‘to cost (something)’ (10b-10c).

- (10) FrWaC (2010) examples (Baroni et al. 2009)
- a. Example from france5.fr
 un avortement c’ est pas de la tarte à assumer
 a abortion it is not PART.F.SG pie to take responsibility of
 ‘having had an abortion is not a piece of cake to live with’

- b. Example from *lesverts.fr*
 c' est le genre de renseignement qui prend du temps à obtenir
 it is the kind of information that takes PART.M.SG time to obtain
 'it's the kind of information that takes time to obtain'
- c. Example from *hardware.fr*
 Ce GPU coûte très cher à fabriquer à Nvidia
 this GPU costs very expensive to make to Nvidia
 'This GPU costs Nvidia a lot to produce'
- (11) a. Ce GPU coûte très cher à Nvidia à fabriquer
 b. Ce GPU coûte des milliers d'euros à fabriquer à Nvidia
 this GPU costs INDEF.PL thousands of euros to make to Nvidia
 'This GPU costs Nvidia thousands of euros to produce'
 c. * Ce GPU est des milliers d'euros à fabriquer
 this GPU is INDEF.PL thousands of euros to make
 ('This GPU is worth thousands of euros to produce')

This last example displays an additional case of control of the subject of the MOC: the *tough* use of verbs like *coûter* not only raises the missing object to subject (as do all *tough* predicates), but additionally licenses an indirect object controller (*à Nvidia*) interpreted as the subject of the infinitive. Importantly, this indirect object is a dependent of the matrix verb and not internal to the MOC, as evidenced by its possibility to permute with the MOC (11a); neither could it be a dependent of *cher* (otherwise known as a *tough* adjective on its own), as it can be replaced by a noun phrase expressing a cost such as *des milliers d'euros* ('thousands of euros') in (11b), which cannot function as a *tough* predicate (11c). *Tough* predicates therefore further the case for an analysis that keeps the subject available for control in addition to extracting the missing object.

Notice that French lacks several other cases of missing object constructions found in English, namely purpose infinitives, *too* and *enough* complements, and *need* predicates (Grover 1995), which simply do not find any parallel in French.

Finally, we leave aside substantival uses of the MOC that do not express the missing object as in *à manger* 'food' ('eatable material', literally 'to eat'), *à boire* 'drink' ('drinkable material', literally 'to drink'), an unproductive pattern (e.g. *à lire* cannot form an NP to mean 'reading material') which we consider to be lexicalised phrases.

2 Internal properties

2.1 Passive properties

The examples given so far leave the subject of the infinitive of the MOC unexpressed (although in the object predication and *tough* uses, it may be controlled). Yet this subject can be realised internally as an oblique marked by *par* ('by'), a realisation typically found with participial passives:

- (12) un livre à lire par toute la classe
 a book to read by all the class
 'a book to be read by the whole class'

Together with the promotion of the same verb's DO to be used as a general external argument, this property has led to analyses of the MOC as a modal (see next section) infinitival passive (Giurgea & Soare 2007).

2.2 Modality

As in English, the MOC carries some modality: *un livre à lire* denotes a book that one either must or can read. However, the status of modality in the *tough* construction is less clear. Although an analysis of *tough* adjectives as covert (bouletic) modals has been proposed for English (Fleisher 2008), the specifically deontic/epistemic modality that exists in all other uses is clearly different from the circumstantial modality *tough* predicates may carry. Conversely, deontic/epistemic modality is actually found in at least two other French infinitival constructions, the infinitival relative (13) and interrogative (14) clauses.

- (13) un endroit où aller
a place where go
'a place to go' = 'a place where one can/must go'
- (14) a. Où aller ?
where go
'Where to go?' = 'Where can/must one go?'
- b. Je ne sais pas où aller.
I NEG know not where go
'I do not know where to go.' = 'I do not know where one can/must go.'

All in all, we take modality as neither a defining nor an exclusive feature of the MOC.

2.3 Locality

French MOCs have been claimed since at least Kayne (1974; 1975) to be VP bounded: the dependency between the antecedent and the missing object is said not to hold across an embedding predicate like *essayer* below.

- (15) Abeillé et al. (1998: 4), our gloss
- a. * Le travail était facile à essayer de finir.
the work was easy to try to finish
'The assignment was easy to try to finish.'
- b. The assignment was easy to try to finish.

This contrasts with English, where this construction has been treated by Pollard & Sag (1994) as (weak) unbounded dependencies:

- (16) Grover (1995: 20)
Kim would be easy for you to persuade Lee to talk to.
- (17) Dalrymple et al. (2000: 16)
Mary is hard for me to believe Leslie kissed.

French tense auxiliaries, however, have been observed to be transparent to this dependency:

- (18) Abeillé et al. (1998: 24), our gloss
des gens utiles à avoir fréquenté
INDEF.PL people useful to have socialised with
'people useful to have known'

Abeillé et al. (1998) have related this transparency to complex predicate formation with tense auxiliaries, a phenomenon involving apparent argument structure sharing between two predicates, namely here the lexical participle and the tense auxiliary. Such argument sharing (or more appropriately raising) is suggested mainly by clitic climbing (19), in which pronominal arguments of the former are hosted by the latter.¹

- (19) Elle l' a lu.
 she 3SG.M has read
 'She has read it.'

Under this perspective, the transparency of clitic climbing verbs does not contradict the generalisation that missing objects are always strictly local: such apparently non-local missing objects are actually local raised objects of the upstairs auxiliary. This reasoning makes an important prediction: all and only clitic climbing verbs should be observed to license embedded missing objects. The first half of this statement is easily verified by looking at the rest of the limited class of French clitic climbing verbs, namely causative *faire*, *laisser* and the perception verbs *voir*, *entendre*, *regarder*, *écouter*.² As illustrated below for *faire* (20a) and *voir* (20b), natural examples from corpus can be found.

(20) FrWaC (2010) examples

- a. Example from cndp.fr
 les différentes compétences à faire acquérir aux élèves de cycle 1
 the different skills to make acquire to the students of cycle 1
 'the different skills to make the preschool pupils acquire'
- b. Example from free.fr
 donner leur avis sur ce problème et sur tous ceux qui leur
 give their opinion on this problem and on all those that DAT.3PL
 sembleraient intéressants à voir discuter
 would seem interesting to see discuss
 'to give their opinion on this problem and on all those that would seem interesting to see being discussed'

However, the second half of the prediction does not seem to prove empirically accurate. Abeillé et al. (1998) already mention the possibility of transparent modal, movement and aspectual verbs for some speakers:

- (21) Abeillé et al. (1998: 4), our glosses
- a. % un livre à devoir lire dès aujourd'hui
 a book to must read as soon as today
 'a book to have to read today'
- b. % une ville difficile à aller visiter en ce moment
 a town difficult to go visit nowadays
 'a town difficult to go to visit now'

¹ See, however, Abeillé & Godard (2002) and Aguila-Multner & Crysmann (2020) for a more comprehensive discussion.

² There is another case of clitic climbing in French found with *être* and other copular verbs, but their complementation pattern restricts them to predicative XPs, which never take DO complements in French, precluding their use in MOCs.

To the best of our knowledge, no speakers of modern French allow clitic climbing with such verbs. The two properties – transparency to a missing object dependency and clitic climbing – therefore appear to be actually distinct, though certainly overlapping.

To determine with more precision the extent of the class of non-climbing predicates that may embed a missing object infinitive, we conducted a corpus study on frWaC (Baroni et al. 2009), which is presented in the next section.

3 Corpus study

3.1 Method

The corpus study was carried out on FrWaC (Baroni et al. 2009), a lemmatised and PoS-tagged 1.3 billion word web-harvested corpus of French. We decided on the corpus precisely because of this large size, since we were looking not just for a particular, somewhat infrequent construction, but for particular sub-cases of this construction (those that involve an embedded missing object).

The corpus queries were conducted using Corpus Query Language (CQL – Jakubíček et al. 2010) expressions in the Sketch Engine.³ The basic search pattern for MOCs is illustrated in (22): essentially we searched for sequences of *à* followed by two infinitives. The use of optional search terms allowed us to cater for potentially intervening adverbs, pronominal clitics, and the markers *de* and *à* that precede some embedded infinitives (e.g. *finir de, commencer à*).

(22) **Basic CQL expression for the MOC:** [word = "à"] [tag = "ADV"]? [tag = "ADV"]? [tag = "PRO.*"]? [tag = "PRO.*"]? [tag = "VER:infi"] [lemma = "de" | lemma = "à"]? [tag = "ADV"]? [tag = "ADV"]? [tag = "PRO.*"]? [tag = "PRO.*"]? [tag = "VER:infi"]

ADV = Adverb, PRO.* = Personal pronoun, VER:infi = Infinitive verb

As the corpus does not contain structural information, it was not possible for the expressions to be sensitive to a missing object, and so the results of such queries contained many infinitival phrases marked by *à* that were not MOCs, such as can occur as complements of various predicates in French. The method we used was therefore to further restrict the search pattern by providing left context appropriate for the MOC, which is only licensed in four different contexts (recall Section 1). Leaving out object predication,⁴ we defined three different CQL search terms to provide left context corresponding to subject predicative, attributive, and *tough* use. Subject predicative context was defined by a disjunction of lemmata termed as predicative in the Lefff electronic lexicon (Sagot 2010) immediately before the MOC. Attributive context was simply defined by the presence of a noun immediately before the MOC. As for *tough* constructions, the query was built by prefixing a disjunction of *tough* adjectives (other categories of *tough* predicates were deemed too infrequent). Starting from the list provided by the Grande Grammaire du français (Abeillé & Godard 2021: Table VI-8), the list of adjectives was extended by making an inventory of *tough* adjectives by searching the corpus for sequences of an adjective followed by *à* then an infinitive. The full list of *tough* adjectives identified in that way is given below.

³ Kilgarriff et al. (2004; 2014). <https://www.sketchengine.eu/>.

⁴ Several cases of object predication were in fact returned while searching for attributive uses, all of them complements of *avoir*. As object predication is much more infrequent than the other contexts, and since cases in which it differs in surface form as a left context from attributive uses are a fortiori even rarer, we do not expect that a significant number of examples were missed by not creating a search expression specific to object predication contexts.

- (23) absurde, admirable, adorable, affreux, agréable, aisé, alléchant, amer, amusant, appétissant, âpre, ardu, assommant, atroce, balaise, banal, beau, bizarre, bon, bouleversant, bénéfique, bête, capital, cauchemardesque, chaud, cher, chiant, chouette, chronophage, comique, commode, complexe, compliqué, con, confortable, contraignant, cool, coriace, coton, coûteux, crucial, cruel, cuisant, curieux, dangereux, difficile, différent, dingue, distrayant, douloureux, doux, drôle, dur, dégoûtant, dégueulasse, délicat, délicieux, délirant, dément, déplaisant, dérangeant, déroutant, désagréable, désirable, désolant, économique, effrayant, effroyable, élémentaire, embarrassant, empoisonnant, encombrant, énervant, ennuyant, ennuyeux, énorme, épouvantable, éprouvant, épuisant, éreintant, essentiel, étonnant, étrange, évident, excellent, extraordinaire, fabuleux, facile, fascinant, fastidieux, fastoche, fatigant, flou, fondamental, formidable, fou, fragile, fun, gonflant, grand, grandiose, gratuit, grave, grisant, gros, génial, gênant, hallucinant, hasardeux, haut, hideux, hilarant, honteux, horrible, hyperfacile, idiot, idéal, illicite, immédiat, important, impossible, impressionnant, impropre, impératif, inaudible, inconfortable, incroyable, indispensable, infect, infernal, infâme, inintéressant, injuste, insoutenable, insupportable, intelligent, interminable, intolérable, intuitif, intéressant, inutile, irritant, joli, jouissif, joyeux, judicieux, laborieux, laid, lassant, lent, long, louable, lourd, lourdingue, léger, magique, magnifique, majestueux, malaisé, malcommode, malheureux, malin, marrant, mauvais, meilleur, merveilleux, mignon, moche, moindre, monstrueux, mortel, méchant, naze, net, nul, nécessaire, obligatoire, obscène, palpitant, paradoxal, passionnant, pathétique, perplexe, personnel, pertinent, petit, physique, pitoyable, plaisant, polluant, possible, poussif, pratique, primordial, prioritaire, précieux, préférable, prématuré, pénard, pénible, périlleux, raide, raisonnable, rapide, rare, ravissant, redoutable, reposant, ridicules, rigolo, risqué, rude, réaliste, réjouissant, sain, salubre, savoureux, sidérant, simple, simplissime, somptueux, souple, spectaculaire, splendide, stressant, subjugant, sublime, suffisant, super, superbe, surprenant, sympa, sympathique, sympatoche, tendre, terrible, terrifiant, toxique, triste, trivial, troublant, ultra-facile, ultra-simple, ultrarapide, urgent, utile, vexant, vital, vulgaire

Further restrictions were added to the queries to help identify the target construction more efficiently by filtering out the most common sources of noise (false positives). Firstly, personal pronouns preceding the infinitive that always express direct objects (*le, la, les*) were filtered out, as they indicate a direct object that is not missing but pronominalised. Secondly, the word directly after the second infinitive was constrained to not be a determiner, as determiners in this position very frequently introduced an expressed (i.e. non-missing) direct object. Thirdly, we identified, via successive refinement, some lexical items that were prone to yield an incommensurate number of false positives, and consequently excluded them from the search pattern. For example, the nouns *manière* and *façon* ('manner, way') were ruled out in the attributive search pattern, as they very frequently head a saturated (i.e. with no MO) *à*-infinitival complement (e.g. *de manière à trouver du travail*, 'in order to find work'); similarly, *dire* was excluded from the predicative search pattern as it was overly frequent because of its use in *c'est à dire* 'that is to say'.

These measures were not sufficient to suppress all false positives from the results, but they were effective enough to permit manual inspection of the results to identify legitimate instances of MOCs. Although these restrictions may have precluded some valid

examples from being turned up by the search patterns (false negatives), such as e.g. an unlikely use of *manière* as an antecedent noun modified attributively by a MOC, we feel this effect was minimal and did not greatly affect the number of results gathered. Further notable limitations of this study are to be related with the limitations of the chosen corpus: as a web-harvested corpus, frWaC often contains spelling mistakes, which in turn limited the efficiency of searching by tag. Although false positives returned because of spelling or tagging issues could easily be discarded when browsing the results, it is possible that such issues prevented valid examples of embedded missing objects from being properly retrieved.

3.2 Results

The first infinitive verb of each successfully identified missing object construction that was returned by the query was added to an inventory of transparent embedding predicates. These verbs are listed below, followed by (one of) the sentence(s) they occurred in. The numbers in brackets denote the amount of unique occurrences in the corpus.

(24) Movement verbs

a. *aller* (189) – example from allocine.fr

Un bon film à **aller** voir pour le plaisir des yeux.
a good film to go see for the pleasure of the eyes

‘A good film to go and watch for your visual pleasure.’

b. *venir* (61) – example from angers.fr

des panneaux de petites annonces à **venir** consulter sans
INDEF.PL signposts of little advertisements to come refer to without
modération
moderation

‘posting boards with advertisements to come and browse without moderation’

(25) Modal verbs

a. *devoir* (4) – example from probb.fr

Histoire de ne pas avoir trop d’ autocollants à **devoir** coller sur la
in order to NEG have too of stickers to have to stick on the
carrosserie
car’s body

‘So as to avoid having too many stickers to place on your car’

b. *pouvoir* (4) – example from yozone.fr

les deux Italiens ont tant de portes à **pouvoir** entrouvrir qu’ on
the two Italians have so many of doors to be able to half-open that we
peut s’attendre à de magnifiques surprises
can expect to INDEF.PL magnificent surprises

‘the two Italians have so many doors to be able to open that we can expect some beautiful surprises’

c. *manquer de* (3) – example from ird.fr

Les laboratoires, équipes ou départements et les chercheurs à ne pas
 the laboratories teams or departments and the researchers to NEG
manquer de visiter.

miss to visit

‘The labs, teams or departments and the researchers to be sure to visit.’

(26) Aspectual verbs

a. *finir de* (37) – example from greluche.fr

J’ ai aussi trois tonnes de billets à **finir** d’ écrire

I have also three tons of blog posts to finish to write

‘I also have a million blog posts to finish writing’

b. *commencer à* (1) – example from club.fr

les cartes et la frise à **commencer** à apprendre page 31

the maps and the timeline to start to learn page 31

‘the maps and the timeline to begin to memorise on page 31’

c. *continuer à/de* (3) – example from hardware.fr

J’ ai récupéré un projet de la fac à **continuer** de développer en html

I have retrieved a project of the uni to continue to develop in html
 php javascript.

php javascript

‘I have retrieved a project from uni to continue to work on in html php and javascript.’

(27) Cognition verbs

a. *savoir* (34) – example from ac-lille.fr

On créera aussi pour chaque élève un carnet orthographique où
 we will create also for each student a booklet orthographical where
 seront regroupés par ordre alphabétique les mots à **savoir** écrire et
 will be collected by order alphabetical the words to know write and
 les règles intuitives
 the rules intuitive

‘We will also create for each student a spelling notebook in which the words they need to know how to write will be collected in alphabetical order’

b. *apprendre à* (2) – example from forumpro.fr

C’ est le plus aisé à **apprendre** à utiliser et pour la mise en page c’

It is the most easy to learn to use and for the formatting it
 est nettement plus pratique

is clearly more practical

‘It’s the easiest one to learn to use and it’s a lot more practical for formatting’

c. *oublier de* (6) – example from *voyage.fr*

Voyage vous propose une petite liste de choses à ne pas **oublier** de faire
 Voyage 2PL.IO proposes a little list of things to NEG forget to do

‘Voyage provides you with a short list of things to not forget to do’

d. *penser à* (1) – example from *noosblog.fr*

Je réalisais accessoirement que je n’ avais toujours pas d’ information
 I realised coincidentally that I NEG had still no information
 sur ce qu’ on y mangeait, ni à quel prix, petits détails à **penser** à
 on what we there ate nor at what price little details to think to
 demander si toutefois j’ étais rappelée
 ask if however I was called back

‘I coincidentally realised that I still had no information on what could be had for food there, or at what price, little details I should not forget to ask about, if they even called me back’

(28) Volition verbs:⁵

a. *décider de* (1) – example from *club.fr*

ma question c’ est : comment trouver ce qu’ il y a de possible à **décider**
 my question it is how find what there is of possible to decide
 de faire
 to do

‘my question is: how to find out what possible things to decide to do there are’

b. *vouloir* (1) – example from *listesratures.fr*

La qualité, donc l’ appréciation des lecteurs, semble l’ objectif
 the quality so the appreciation of the readers seems the objective
 logique à **vouloir** atteindre.
 logical to want reach

‘Quality, i.e. readers’ esteem, seems like the logical goal to want to reach.’

⁵ Two examples involving *essayer* (‘try’) escaped the attributive queries for technical reasons, but were found while looking specifically for this verb:

(i) Example from *free.fr*

Tu dis : “la religion vise un idéal à **essayer** d’ atteindre”.
 you say the religion aims a ideal to try to reach

‘You say: “religion aims for an ideal to try to reach”.’

(ii) Example from *optimum-blog.net*

l’ économie est un truc à étudier (ou à **essayer** de comprendre)
 the economy is a thing to study or to try to understand

‘the economy is something to study (or to try to understand)’

c. *tenter de* (6) – example from jussieu.fr

Parmi les médicaments à **tenter** de diminuer puis d'arrêter
 among the medications to attempt to diminish then to stop
 de principe, on retient les médicaments à effet anticholinergique
 a priori we memorise the medications with effect anticholinergic

'Among the drugs to try to diminish and a priori to quit using, drugs with an anticholinergic effect stand out'

d. *oser* (1) – example from ens-lsh.fr

J'ai une fureur pour assister à ces audiences, qui, par fois, sont
 I have a furor to attend to these audiences which sometimes are
 à n' **oser** qualifier, comme cette, prescription rétrograde, ce voyage en
 to NEG dare describe like this prescription backward this trip in
 1744 qu' on voudrait faire faire aux hommes de 1832
 1744 that they would want have made to the men of 1832

'I am outraged to attend these audiences, which, sometimes, one must not dare to describe, like this backward prescription, this trip to 1744 that they would have the men of 1832 do'

e. *éviter de* (1) – example from free.fr

Le padding est à **éviter** de modifier
 the padding is to avoid to modify

'Padding is not to be modified'

These results show that the class of verbs that are transparent to the missing object dependency is highly heterogeneous. It contains many verbs without clitic climbing (24-28), and at least two more semantic types than previously thought (we find cognition (27) and volition (28) verbs in addition to the movement, modal and aspectual verbs recognised by [Abeillé et al.](#)). Furthermore, the verbs of this class exhibit various syntactic behaviours, including raising verbs like *pouvoir* contrasting with control verbs like *vouloir*, as well as different complement patterns regarding the embedded infinitive (e.g. bare VP with *vouloir*, marked by *à* with *penser*, or marked by *de* with *décider*). We therefore reject argument composition as the mechanism responsible for embedded missing objects, and will propose instead in section 6 an alternative analysis of such cases.

4 Previous analyses

4.1 Movement analysis

The first analysis of French MOCs (specifically in their *tough* use) is due to [Kayne \(1977\)](#) and it is essentially based on syntactic movement. Although he does note the passive-like effect associated with the construction, he still rejects an analysis along these lines on several grounds. Firstly, he claims that MOCs, unlike participial passives, do not license a *par/de*-phrase. While we concur on the impossibility of a *de*-phrase, this is likely due to the associated stative semantics of the verb, which seem incompatible with the agentive interpretation of the subjects of MOCs. *Par*-phrases, on the other hand, seem to us to be

entirely grammatical in this context, and corpus examples abound on frWaC: searching for a *tough* adjective + *à* + Vinf + *par*, we found 546 unique examples of agentive *par*-phrases, some of which are given below.

(29) FrWaC (2010) examples

a. Example from asso.fr

Ces mélanges se présentent sous une forme liquide, facile à
these mixes present under a form liquid easy to
assimiler par l' intestin, et bien tolérés par l' estomac.
assimilate by the intestine and well tolerated by the stomach

'These mixes appear in liquid form, easy to assimilate by the intestine, and well tolerated by the stomach.'

b. Example from allocine.fr

Michael Douglas a du tout de même être très difficile à manier
Michael Douglas has had to still be very difficult to handle
par David Fincher
by David Fincher

'Michael Douglas still must have been difficult for David Fincher to direct'.

c. Example from gouv.fr

Ce contrat est global et plus simple à gérer par les autorités
this contract is global and more simple to manage by the authorities
publiques
public

'This contract is global and simpler for public authorities to manage'

Secondly, [Kayne](#) provides the following contrast involving *se*-verbs:

(30) [Kayne \(1977: 337-338\)](#), our glosses and intended translation

a. Ce livre est difficile à se procurer
this book is difficult to REFL procure
'This book is difficult to procure.'

b. *Ce livre se_i sera procuré (par Jean_i)
this book REFL is acquired by Jean
(‘This book is acquired by Jean.’)

Transitive verbs with intrinsic or reflexive *se* like *se procurer* (‘to acquire’) can enter a MOC, but not become passive participles. We come back to this point in Section 6.4.

Finally, [Kayne](#) notes the possibility of long distance MOCs involving causative *faire*:

(31) [Kayne \(1977: 337-338\)](#), our gloss

Cette décision sera difficile à faire accepter au comité
this decision will be difficult to make accept to the committee

'This decision will be difficult to get the committee to accept.'

(32) * Cette décision a été fait(e) accepter au comité
this decision has been made accept to the committee
(‘This decision was made to be accepted by the committee.’)

Essentially, the author rightfully notes that MOCs can be non-local dependencies, while participial passives are not generally non-local (32). However, we can adduce two cases that hint at a more complex picture of the locality of passive-like operations in French. Firstly, medio-passive *se* can operate non-locally:

- (33) Grâce à la mode du cat-sitting, les chats se font plus facilement
 thanks to the trend of cat-sitting, the cats REFL make more easily
 garder.
 keep
 ‘Thanks to the trend of cat-sitting, it is easier to have your cats kept.’

Secondly, non-local participial passives are in fact found in French, although the class of verbs they can occur with is limited:

- (34) a. Ce monument sera fini de construire en 2023.
 this monument will be finished to build in 2023
 ‘This monument will be finished building in 2023.’
 b. Achevé d’imprimer le 11 décembre 1987
 finished to print the 11 december 1987
 ‘Printed on December 11 1987’ (standard phrasing in French edition notices)

The idea of a movement analysis suggested by the author furthermore raises the question of the boundedness of the dependency: although such an approach could certainly deal with examples like (31), it remains unexplained how a movement rule would correctly account for the peculiar locality of the construction that we uncovered in the previous section, and would likely overgenerate to an unbounded dependency.

4.2 Reduced relative analyses

In their analysis of French MOCs, Abeillé et al. (1998) capitalise on the restricted locality of the construction, claiming that only argument composition verbs like the tense auxiliaries *avoir/être*, as well as *faire/laisser* and perception verbs are transparent to the MOC. What these verbs have in common is that they permit clitic climbing, which Abeillé et al. regard as local morphological expression of raised arguments. The analysis of French *tough* constructions crucially builds on this: the lexical entry of *tough* adjectives have their external argument (selected via SUBJ or MOD) co-indexed with an unsaturated direct object valency of their \hat{a} -infinitival complement, as detailed in Figures 1–2.

Figure 1: *Facile* – attributive variant (from Abeillé et al. 1998: 7)

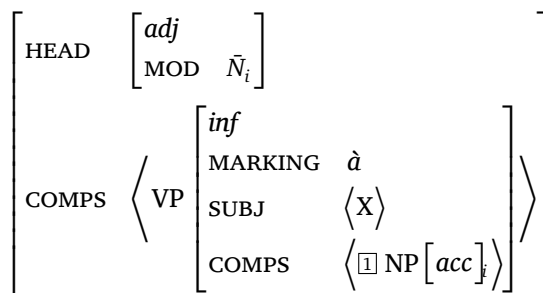
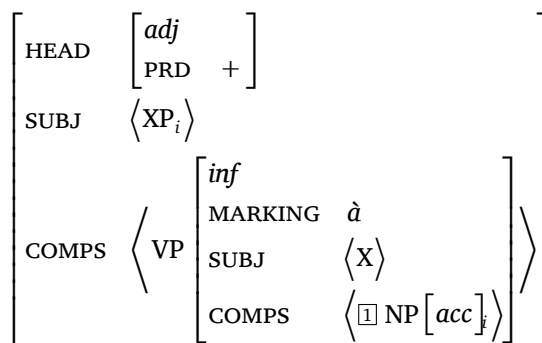


Figure 2: *Facile* – predicative variant (from Abeillé et al. 1998)

Attributive uses of the MOC receive a similar analysis. Abeillé et al. (1998) treat them as infinitival relative clauses, suggesting that there are three ways in which an antecedent noun can be related to an argument within the relative: either by way of an unbounded dependency (SLASH) or by direct co-indexation with an unsaturated SUBJ (*qui*-relatives) or COMPS valency (MOC).

There are several problems with this line of analysis: first, an infinitival relative analysis is intrinsically limited to attributive uses and *tough* constructions. Most crucially, this approach does not extend to subject predications, since the verbs used in these contexts, like copula *être* or *sembler*, are subject-to-subject raising verbs, not object-to subject raising verbs.

Second, as we have shown in Section 3, the class of verbs transparent to the MOC is by no means limited to clitic climbing or argument composition verbs, but comprises a larger class of VP-taking control and raising verbs, most of which do not exhibit the kind of behaviour that Abeillé & Godard (2002) associate with argument composition.

Third, the passive-like properties of the construction remain entirely unaccounted for. This holds most obviously for the realisation of the logical subject as an oblique *par*-phrase. But it also treats as a mere coincidence the fact that the external argument in a MOC is always the logical direct object.

- (35) une pétition à signer par tous les membres
 a petition to sign by all the members
 ‘a petition to be signed by all members’

4.3 Raising

4.3.1 A-movement in Romance

Giurgea & Soare (2007) discuss *tough* constructions in Romance, noting the formal similarities to predicative and attributive uses of the missing object construction. They suggest that these constructions are passive-like and should therefore be captured in terms of A-movement (raising).

Giurgea & Soare (2007), however, do not address the issue of locality: in the light of the data presented in Section 3, A-movement will have to cross intervening PRO subjects, essentially leading to principle A violations. Similarly, their proposal does not address the case of external control either. Thus, while we concur with their perspective of the missing object construction as an instance of grammatical function change, we do note that A-movement per se is insufficient to do full justice to internal and external control.

4.3.2 Missing object constructions in English

While English *tough* constructions are often regarded as unbounded dependencies (cf. Polard & Sag 1994), it has been noted that extraction from finite clauses may be subject to individual or dialectal variations (Hukari & Levine 1987). Studying a wide range of missing object constructions in British English, including *tough* constructions and parasitic gaps, Grover (1995) observes that non-locality of the dependency uniformly involves chains of control or raising verbs.

Consequently, Grover (1995) suggests that missing objects are removed from COMPS, akin to direct objects in true passives, but that they are promoted to secondary subject status: technically, this is implemented by appending the erstwhile direct object valency to the SUBJ list. Control and raising continue to target the first element of a complement's SUBJ list, but additionally structure-share (=inherit) the list remainder, thereby implementing a notion of secondary subject raising that goes piggy-back on existing control and raising dependencies.

Given the partial overlap between missing object constructions in English and French and the parallelism regarding locality conditions and control relations, our formal analysis of the French data will actually build on Grover's account.

5 Discussion: Subject status in the missing object construction

As we have seen in the previous sections, the French missing object construction shares some properties with standard lexical passives. First, the argument that remains unexpressed in situ is always the direct object. Second, the missing object behaves as the external argument of the construction, being co-referent with the antecedent noun in attributive use or the grammatical function being predicated on. Third, the downstairs logical subject is either omitted, or else realised by a *by*-phrase.

While the missing object behaves as the external argument of the construction as a whole, subject status within the construction does not reflect this: most notably, whenever a (subject) control verb intervenes (see Section 3 above), the logical subject of the control verb is co-referent with the logical subject of its infinitival complement, and crucially not with the missing object.

The logical subject of the verb whose object is locally missing is not only available to control within the construction, but can also be controlled from outside. E.g. with object predication, the missing object of the downstairs infinitive corresponds to the direct object of *donner* or *avoir*, while at the same time the logical subject of the *à*-infinitive is controlled by some other function of the object predication verb, i.e. by the subject in the case of *avoir*, or by the indirect object, in the case of *donner*. Similarly, control of the logical subject from outside was also found with *tough* predicates like *coûter* 'cost' and *prendre* 'take', cf. example (10c).

These observations show that we are confronted here with two notions of subject-hood being operative at the same time. One way we can make sense of this is in terms of a two-step passivisation, separating promotion of the direct object from demotion of the logical subject. More precisely, we shall assume that, in a first step, a lexical rule promotes the direct object to secondary subject in the sense of Grover (1995), without demotion of the logical subject, which stays available for control. At the top of the *à*-infinitive, the grammatical function change is finally concluded.

6 Analysis

We now propose a unified analysis of the MOC as an infinitival passive phrase able to act as a predicate, a modifier, or a selected complement (*tough* uses). This analysis is formalised in Head-driven Phrase Structure Grammar (HPSG – Pollard & Sag 1994). The chosen approach shall expose two of the construction’s core arguments, viz. not only the missing direct object but also the logical subject which needs to be available for external control. We also take care to account for the particular locality of the dependency independently of argument composition by allowing select lexical heads (i.e. verbs of the class of *aller* that were found to naturally occur in this position) to intervene transparently within the construction.

We rely specifically on the double subject approach suggested by Grover (1995) to encode the dependency with the missing object. Lexically, the direct object valency is promoted to secondary subject, which accounts not only for the absence of a local direct object complement, but also makes the direct object available to raising. This approach captures the fact that the MOC behaves externally like a passive, but internally maintains canonical control relations by intervening infinitives in non-local cases.

6.1 The bottom of the construction: A lexical rule for passive infinitives

We implement the first step in the passivisation by means of a lexical rule operating on valence lists (cf. Grover 1995). This rule, given in Figure 3, applies to a non-finite verb and blocks the realisation of its direct object by taking it off the COMPS list. As in Grover (1995), the dependency is established on the SUBJ list: the blocked DO becomes a secondary subject of the verb.⁶ This is enough to license downstairs infinitives or participles with an unexpressed object, while at the same time constraining the distribution of such passive verbs: as contexts for finite phrases require a single subject under standard HPSG theory, the double subject value effectively limits occurrences of passive infinitives to the special contexts we define below. This can be enforced by a constraint on head-subject phrases requiring the subject list being saturated to have exactly one element, as in Figure 4.

Figure 3: Lexical rule for \hat{a} -infinitival passives

$$\left[\begin{array}{l} \text{HD} \\ \text{VAL} \\ \text{ARG-ST} \end{array} \left[\begin{array}{l} \text{VFORM } \textit{nonfin} \\ \text{SUBJ } \langle \boxed{1} \rangle \\ \text{COMPS } \langle \boxed{2} \mid \boxed{3} \rangle \\ \langle \boxed{1}, \boxed{2} \mid \boxed{3} \rangle \end{array} \right] \right] \mapsto \left[\begin{array}{l} \text{VAL} \\ \text{COMPS} \end{array} \left[\begin{array}{l} \text{SUBJ } \langle \boxed{1}, \boxed{2} \rangle \\ \boxed{3} \end{array} \right] \right]$$

6.2 The top of the construction

As we have seen in the discussion of the empirical patterns, the missing object construction features in four different contexts: attributive use, subject predicatives, object

⁶ The possibility of having more than one element on the SUBJ valence list has also been explored for the treatment of the Double Nominative Construction in Korean (Lee 2003; Ko 2010; Ryu 2013). Similarly, Müller & Ørsnes (2013) propose an analysis of object shift in Danish that has multiple elements on the SPR list, alongside the subject.

Figure 4: Head-subject schema limiting the realisation of double subjects

$$\left[\begin{array}{l} \textit{hd-sbj-phr} \\ \text{HD-DTR} \quad \left[\text{SYNSEM|LOC|CAT|VAL|SUBJ} \quad \langle \boxed{1} \rangle \right] \\ \text{NON-HD-DTR} \quad \left[\text{SYNSEM} \quad \boxed{1} \right] \end{array} \right]$$

predication (e.g. *avoir* and *donner*), and *tough* constructions. While in the case of *tough* predicates and object predication verbs the missing object construction is specifically selected by the lexical entry, this is certainly not the case for the attributive use, and it is not necessarily the case for subject predicatives either, if we take absolute predicatives into account, cf. (36).

- (36) Avec cinq livres à lire, il est vraiment en retard.
 with five books to read he is really behind
 ‘With five books to read, he is really behind.’

We therefore propose that missing object constructions in French can be licensed in one of two ways: either lexically by a governing predicate, or constructionally.

This distinction is further corroborated by the differences regarding control: while the subject of the *à*-infinitive can undergo obligatory control from outside, this again appears to be limited to the cases where the missing object construction is specifically selected by a governing lexical head, as is the case for object predication verbs, and some *tough* predicates, cf. (10c), but crucially not for ordinary attributive and predicative uses. Furthermore, the fact that there is lexical variation among *tough* predicates with respect to control provides additional support for a lexical perspective.

Furthermore, we observed in Section 2.2 that missing object constructions differ also with respect to the modality they introduce: while attributive and predicative uses clearly involve necessity or possibility, this is not the case for *tough* predicates. Thus, it appears preferable to associate modality not with the *à*-infinitive directly, but rather make it a property of the context in which it is used.

Finally, there is no evidence that the marker *à* found at the top of the infinitival construction behaves any different from ordinary infinitival markers: in particular, modal infinitives do not constitute extraction islands, as shown below.⁷

- (37) Abeillé & Godard (2021: VI – 60d), our gloss and translation
 À qui est-ce que ce film est facile à montrer ?
 to who is it that this film is easy to show
 ‘Who is this film easy to show to?’

This contrasts with prepositions, which, unlike VP markers, are indeed islands (Abeillé et al. 2006). If the marker *à* is unlikely to be a preposition, the predicative and attributive properties are best conceived of as properties of the construction, rather than of its initial lexical marker.

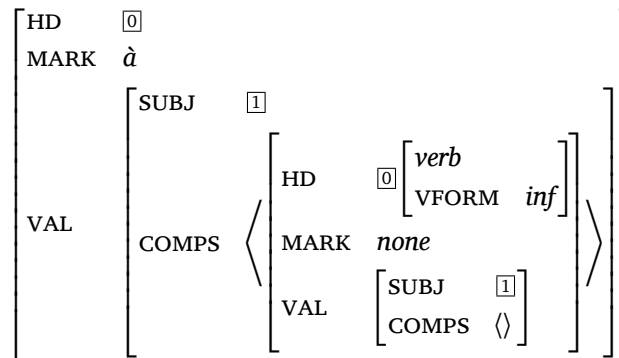
⁷ We thank one of the reviewers for having pointed this out to us.

6.2.1 The infinitival marker \grave{a}

Given the fact that missing object constructions are not extraction islands, thus patterning with ordinary \grave{a} -marked VPs, we propose that the marker found at the top of the MOC is indeed the ordinary infinitival VP marker.

We shall essentially follow [Abeillé et al. \(2006\)](#), treating the marker as a weak head, as depicted in Figure 5.

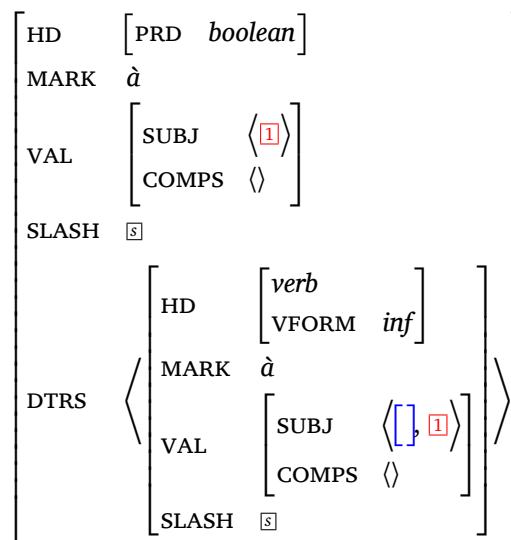
Figure 5: Lexical entry for VP marker \grave{a}



In a nutshell, the marker takes an infinitival VP complement, with which it shares the SUBJ and H(EA)D values. It differs from its bare VP complement, however, in terms of its MARK value.

6.2.2 Constructional licensing of MOCs: predicative and attributive uses

Figure 6: Constructional licensing by unary rule



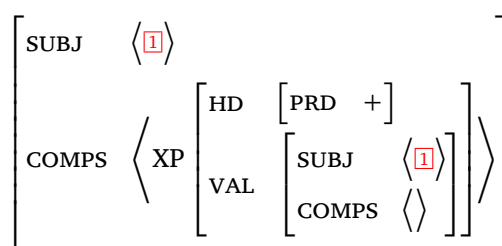
As we have observed above, attributive and subject predicative uses differ from *tough* constructions and object predications in two crucial respects: first, there is no specific lexical governor selecting the missing object construction and there is no control of the infinitive's primary subject from outside. Moreover, French VPs do not normally have attributive or predicative uses.

We therefore suggest that predicative and attributive uses are actually licensed by a specific construction, implemented by the unary rule detailed in Figure 6. The rule projects an XP[PRD *boolean*] from a double subject *à*-marked VP. Crucially, this rule ignores the primary subject of the VP daughter, and raises that daughter’s secondary subject to its own SUBJ value, thereby concluding the passivisation effect.

Subject predication

The predicative use of the MOC is straightforwardly derived under a standard HPSG account by subject-to-subject raising. A lexical entry for the copula is given in Figure 7; similar entries can be given to other copular verbs.

Figure 7: Lexical entry for the copula *être*



The tree in Figure 8 illustrates the analysis of a simple predicative instance of the construction, *ce livre est à lire* (‘this book is to be read’).

Attributive use

As in English, there is in French a systematic alternation between predicative and attributive uses for e.g. adjectives. We model this alternation by means of a unary rule deriving the latter from the former. This rule is given in Figure 9. The rule simply shifts the valence from SUBJ to MOD(IFIER). As the predicative/attribution alternation in general may have exceptions both ways (non-attributive predicates and non-predicative attributes), the feature PR(E)D(ICATIVE) controls which elements may undergo it, restricting application to signs for which PRD is defined and where the value unifies with *–*, which includes *boolean*.

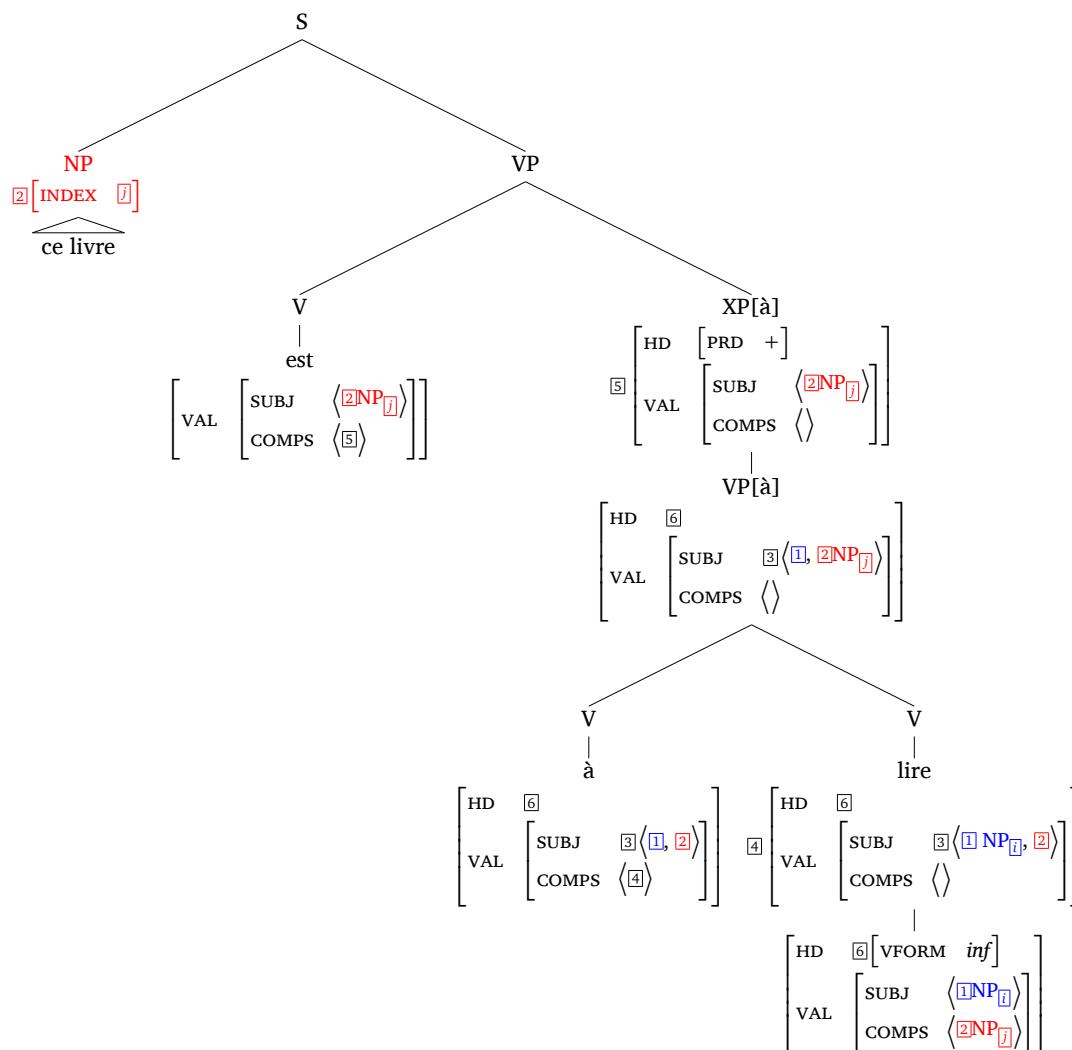
A sample derivation for a simple local attributive MOC, *livre à lire* ‘book to read’, is given in Figure 10.

6.2.3 Lexical licensing of MOCs

Lexical heads selecting for the missing object construction include object predication verbs, such as *avoir*, *donner*, or *laisser*, as well as *tough* predicates. In contrast to attributive and subject predicative uses, we also find control of the primary subject of the *à*-infinitive. Thus, we assume that in these constructions, the governing lexical item directly subcategorises for a double subject *à*-infinitive.

Object predication

We analyse object predication as involving both control and raising. A lexical entry for *avoir* is given in Figure 11. As can be seen in the entry, the subject of *avoir* exerts

Figure 8: Sample analysis for *ce livre est à lire*

control over the primary subject of the MOC, just like control is done within the MOC. The secondary subject of the *à*-infinitive is raised to object.

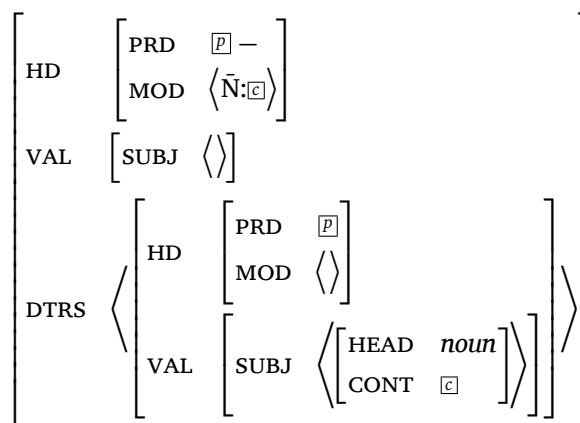
Donner works very similarly to *avoir* in this context, only it is the indirect object rather than the subject that acts as a controller for the primary subject of the *à*-infinitive (Figure 12).

The different raising and control relations are illustrated in Figure 13, which derives *avait ce livre à lire* 'had to read this book'.

Tough uses

We also analyse *tough* predicates in terms of raising of a secondary subject. A lexical entry for *facile* is given in Figure 14, and a sample derivation for *facile à lire* in Figure 15. Note that attributive uses of *tough* adjectives can be derived by the unary rule previously given (Figure 9).

As for *tough* predicates that introduce a controlling argument such as *coûter*, they can again specify the control relation between the controller and the MOC's primary subject (Figure 16), analogous to the object predication verbs we discussed above.

Figure 9: PRD-to-MOD unary rule

6.3 Boundedness

As was found in Section 3, a class of verbs may intervene in the dependency with the missing object, comprising both subject raising and subject control verbs. Our analysis is simply to treat such verbs as optional raisers of secondary subjects. Turning first to subject control verbs like *aller*, we can give an entry as in Figure 17, which displays the raising of an optional secondary subject. Note that at this point in the structure, the primary subject is still the logical subject, which means control of the subject functions as usual with no further specifications.

As for subject raising verbs like *pouvoir*, they simply raise their complement's entire subject list, whether simple or double (Figure 18). This extends quite naturally to the tense auxiliary *avoir*, with the sole difference of a participial and not infinitival verb form.

As for verbs that are opaque to the dependency, their entries simply require a complement with a singleton subject list, in keeping with standard HPSG.

Finally, some of the transparent verbs discovered embed not a bare infinitive, but one marked with *à* or *de*. This requires the minor adjustment of making the two markers subject-list raisers in the same way as *pouvoir*.

The trees in Figures 19–20 summarise the analysis of non-local missing objects with intervening control and raising verbs, respectively.

6.4 Reflexive transitives

We can now turn back to the issue concerning reflexive *se procurer* raised by Kayne (1977) against a passive analysis. The relevant contrast is repeated below from (30).

(38) Kayne (1977: 337-338), our glosses and intended translation

- a. Ce livre est difficile à se procurer
 this book is difficult to REFL procure
 'This book is difficult to procure.'
- b. *Ce livre se_i sera procuré (par Jean_i)
 this book REFL is acquired by Jean
 ('This book is acquired by Jean.')

In our view, the different behaviours of the two passives can ultimately be attributed to the difference in verb form required by the two contexts. While the infinitive form

used in the MOC can easily host pronominal clitics of any kind, French syncretic passive-past-perfective participles are morphologically defective, and can never combine with reflexive pronouns such as *se*. Of course, given the possibility of clitic climbing to the copula in the passive case, this fact alone does not predict the impossibility of (38b). However, facts from another voice phenomenon involving clitic climbing, namely the causative *faire*-construction, illustrate a peculiarity of the interaction of voice and reflexivity in French:

- (39) Il_i $s_{i/*j}$ ' est fait coiffer par Camille_j.
 he REFL is made do one's hair by Camille
 'He got his hair done by Camille.'
- (40) Il_i a fait $se_{j/*i}$ coiffer les enfants_j.
 he has made REFL do one's hair the children
 'He got the children to do each other's/their hair.'

Essentially, when *se* originates from the argument structure of a verb in a causative construction, its placement (climbing or local) dictates how it is interpreted: upstairs *se* is bound by the subject of the causative event, while downstairs *se* is bound by the lexical verb's subject. In other words, reflexive *se* is always locally interpreted, i.e. bound by the local subject of the verb it attaches to. We take the causative situation to be parallel to the participial passive case, and therefore expect that the only way for participial passives to express the binding of *se* by the logical subject as in (38b) is by downstairs realisation of *se*. This realisation is however ruled out by morphological defectiveness, whereas upstairs realisation would come with a different binding relying on the passive subject (e.g. the original DO), which explains the ungrammaticality of (38b). The MOC's passive, which, on the other hand, is an infinitive, can freely realise *se* locally downstairs, where it is accordingly bound by the lexical verb's original subject, accounting for (38a).

6.5 Missing sentential complements

Before we conclude, let us briefly address a particular case of the MOC in which the missing object is not an object NP, but a complement clause introduced by *que*:⁸

- (41) Ruwet (1976: 79-83), our gloss and translation
- Que Nixon ne soit pas impliqué [...] est difficile à croire.
 that Nixon NEG.be not involved is difficult to believe
- 'That Nixon is not involved [...] is difficult to believe.'

⁸ Another noteworthy case involves a missing object that is part of a V-NP idiom together with the infinitive:

- (i) Ruwet (1983: 42), our gloss and translation
- La soirée s'annonçait sinistre, mais, grâce à ce boute-en-train de Gaston, la glace a
 the evening promised to be sinister but thanks to this jolly fellow of Gaston the ice has
 été facile à briser.
 been easy to break
- 'The evening promised to be dreadful, but thanks to that jolly fellow Gaston the ice was easy to break.'

As Ruwet (1983: 40-47) notes, this property of idioms is shared with participial passives. Although we do not provide a sample analysis of such cases, it should be sufficiently clear that these cases are easily captured by the passive-like approach we defend here.

- (42) Que Nixon ne soit pas impliqué est à ne pas croire une seconde.
 that Nixon NEG.be not involved is to not believe a second
 ‘That Nixon is not involved is not to be believed for a second.’

This situation is in fact to be expected under a passive analysis, as *que*-complements are passivisable in French:

- (43) Qu’ il ne soit pas impliqué est encore cru par beaucoup.
 that he NEG.be not involved is still believed by many
 ‘That he is not involved is believed by many.’

The distribution of such missing sentential complements seems restricted compared to the missing direct objects: attributive uses are excluded, a fact readily derived by the description of the rule in Figure 9, which applies only to nominal heads; object predication uses are also excluded, which we attribute to selectional constraints for NP by the relevant verbs (e.g. *avoir*, *donner*; cf. Figures 11–12); subject predication is possible (42), and allowed by our underspecified description of the subject of copular verbs (Figure 7); as for *tough* predicates, they also take underspecified XP subjects (Figures 14–16), correctly allowing sentential complements.

7 Conclusion

In this paper we have presented an in-depth investigation of the French MOC. While we concur with previous research on the bounded status of the dependency, we extend beyond the attributive and *tough* uses to include predicative uses (both subject and object) as well as the possibility of control, providing a unified analysis of the construction across its distribution. Based on corpus data, we have refined the exact characterisation of the locality of the dependency to include a number of possible intervening verbs. Borrowing insights from Grover (1995) and Giurgea & Soare (2007), we have proposed an analysis of the MOC as an infinitival passive with two external arguments, which correctly accounts for the possible control relations within and into the construction.

Abbreviations

DAT = dative, F = feminine, INDEF = indefinite, INF = infinitive, M = masculine, NEG = negative, PART = partitive, PL = plural, REFL = reflexive, SG = singular

Funding information

This work has been supported by a doctoral grant from U Paris to Gabriel Aguila-Multner and has also benefited from a public grant overseen by the French National Research Agency (ANR) as part of the program “Investissements d’Avenir” (reference: ANR-10-LABX-0083). It contributes to the IdEx Université de Paris - ANR-18-IDEX-0001. Authors’ names are listed in alphabetical order.

Acknowledgements

Some of the ideas presented here have been presented at the 13th *Colloque de Syntaxe et Sémantique à Paris* (CSSP 2019). We are highly grateful to the audience of that event

for extensive discussion of the data and our proposal, in particular Marianne Desmets, Danièle Godard, Philip Miller, Adam Przepiorkowski, Laurent Roussarie and Manfred Sailer. Furthermore, a great many thanks go to Anne Abeillé and Jean-Pierre Koenig for discussing various aspects of this work at several occasions. All remaining errors are of course ours.

Competing interests

The authors have no competing interests to declare.

References

- Abeillé, Anne & Bonami, Olivier & Godard, Danièle & Tseng, Jesse. 2006. The syntax of French *à* and *de*: an HPSG analysis. In Saint-Dizier, Patrick (ed.), *Syntax and semantics of prepositions*, 147–162. Springer.
- Abeillé, Anne & Godard, Danièle. 2002. The syntactic structure of French auxiliaries. *Language* 78(3). 404–452.
- Abeillé, Anne & Godard, Danièle (eds.). 2021. *La grande grammaire du français*. Éditions Actes Sud.
- Abeillé, Anne & Godard, Danièle & Miller, Philip & Sag, Ivan. 1998. French bounded dependencies. In Dini, Luca & Balari, Sergio (eds.), *Romance in HPSG*, 1–54. Stanford: CSLI publications.
- Aguila-Multner, Gabriel & Crysmann, Berthold. 2020. French clitic climbing as periphrasis. *Linguisticæ Investigationes* 43(1). 23–61.
- Baroni, Marco & Bernardini, Silvia & Ferraresi, Adriano & Zanchetta, Eros. 2009. The WaCky wide web: a collection of very large linguistically processed web-crawled corpora. *Language resources and evaluation* 43(3). 209–226.
- Dalrymple, Mary & King, Tracy Holloway & Butt, Miriam. 2000. Missing-object constructions: Lexical and constructional variation. In *On-line proceedings of the LFG2000 Conference*.
- Fleisher, Nicholas. 2008. A crack at a hard nut: Attributive-adjective modality and infinitival relatives. In *Proceedings of the 26th West Coast Conference on Formal Linguistics*, vol. 91. 163–171.
- Giurgea, Ion & Soare, Elena. 2007. When are adjectives raisers? tough to get it. In Falk, Yehuda (ed.), *Proceedings of the Israel Association of Theoretical Linguistics (IATL) 23*. Tel Aviv University.
- Grover, Claire. 1995. *Rethinking some empty categories*: University of Essex dissertation.
- Hukari, Thomas E. & Levine, Robert D. 1987. Rethinking connectivity in unbounded dependency constructions. In Crowhurst, Megan (ed.), *Proceedings of the Sixth West Coast Conference on Formal Linguistics*. 91–102. Stanford: Stanford Linguistics Association.
- Jakubíček, Miloš & Kilgarriff, Adam & McCarthy, Diana & Rychlý, Pavel. 2010. Fast syntactic searching in very large corpora for many languages. *PACLIC* 741–47.
- Kayne, Richard. 1974. French relative *que*. *Recherches Linguistique de Vincennes II*, 40–61.
- Kayne, Richard. 1975. French relative *que*. *Recherches Linguistique de Vincennes III*, 27–92.
- Kayne, Richard. 1977. *Syntaxe du français: le cycle transformationnel*. Ed. du Seuil.

- Kilgarriff, Adam & Baisa, Vít & Bušta, Jan & Jakubíček, Miloš & Kovář, Vojtěch & Michelfeit, Jan & Rychlý, Pavel & Suchomel, Vít. 2014. The Sketch Engine: ten years on. *Lexicography* 7–36.
- Kilgarriff, Adam & Rychlý, Pavel & Smrž, Pavel & Tugwell, David. 2004. Itri-04-08 the sketch engine. *Information Technology* .
- Ko, Kil Soo. 2010. *La syntaxe du syntagme nominal et l'extraction du complément du nom en coréen: description, analyse et comparaison avec le français*: Paris 7 dissertation.
- Lee, Sun-Hee. 2003. Korean tough constructions and double nominative constructions. In Kim, Jong-Bok & Wechsler, Stephen (eds.), *The proceedings of the 9th international conference on head-driven phrase structure grammar*. 187–208. Stanford: CSLI Publications. <http://csli-publications.stanford.edu/HPSG/3/>.
- Müller, Stefan & Ørsnes, Bjarne. 2013. Towards an HPSG analysis of object shift in Danish. In Morrill, Glyn & Nederhof, Mark-Jan (eds.), *Formal Grammar: 17th and 18th International Conferences, FG 2012, Opole, Poland, August 2012, Revised Selected Papers, FG 2013, Düsseldorf, Germany, August 2013. Proceedings* (Lecture Notes in Computer Science 8036). 69–89. Berlin/Heidelberg/New York, NY: Springer Verlag. <http://hpsg.fu-berlin.de/~stefan/Pub/object-shift.html>.
- Pollard, Carl J. & Sag, Ivan A. 1994. *Head-driven phrase structure grammar*. Chicago: University of Chicago Press.
- Ruwet, Nicolas. 1976. Note sur la montée de l'objet. *Recherches linguistiques de Saint-Denis* IV. 185–208.
- Ruwet, Nicolas. 1983. Du bon usage des expressions idiomatiques dans l'argumentation en syntaxe générative. *Revue québécoise de linguistique* 13(1). 9–145.
- Ryu, Byong-Rae. 2013. Multiple case marking as case copying: A unified approach to multiple nominative and accusative constructions in Korean. In Müller, Stefan (ed.), *Proceedings of the 20th International Conference on Head-Driven Phrase Structure Grammar, Freie Universität Berlin*. 182–202. <http://csli-publications.stanford.edu/HPSG/2013/ryu.pdf>.
- Sagot, Benoît. 2010. The Lefff, a freely available and large-coverage morphological and syntactic lexicon for French. In *7th International Conference on Language Resources and Evaluation (LREC 2010)*.

Figure 10: Sample analysis for *livre à lire*

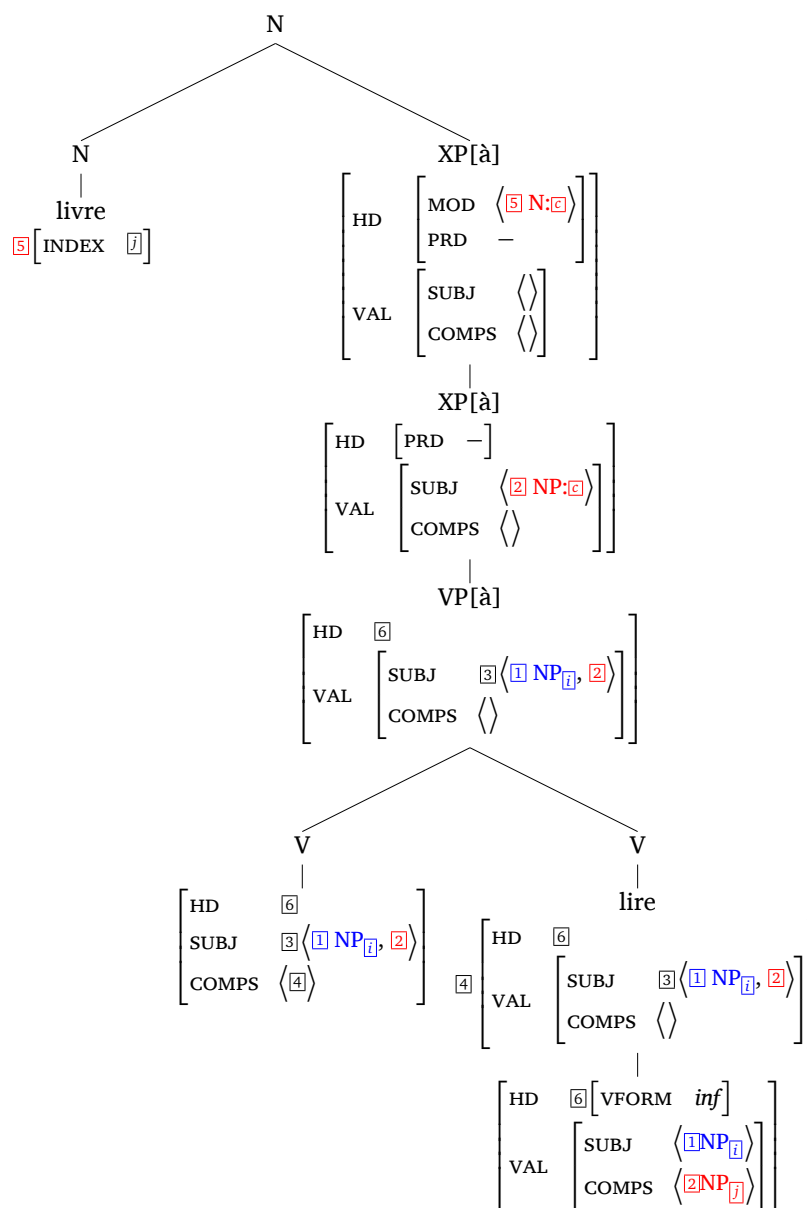


Figure 11: Entry for lexical *avoir*

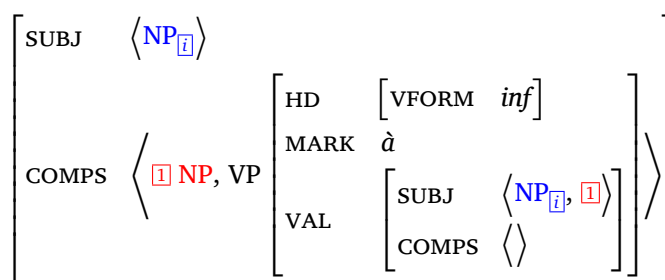


Figure 12: Entry for lexical *donner*

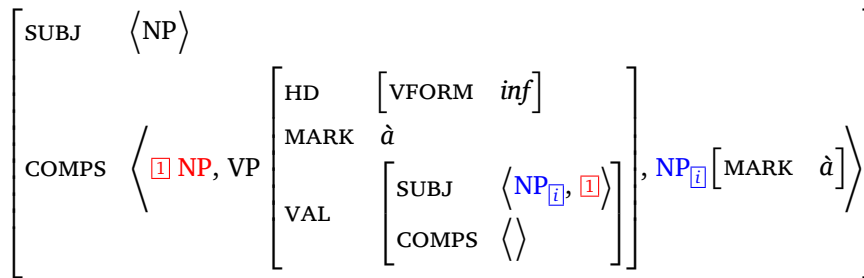


Figure 13: Sample analysis for *avait ce livre à lire*

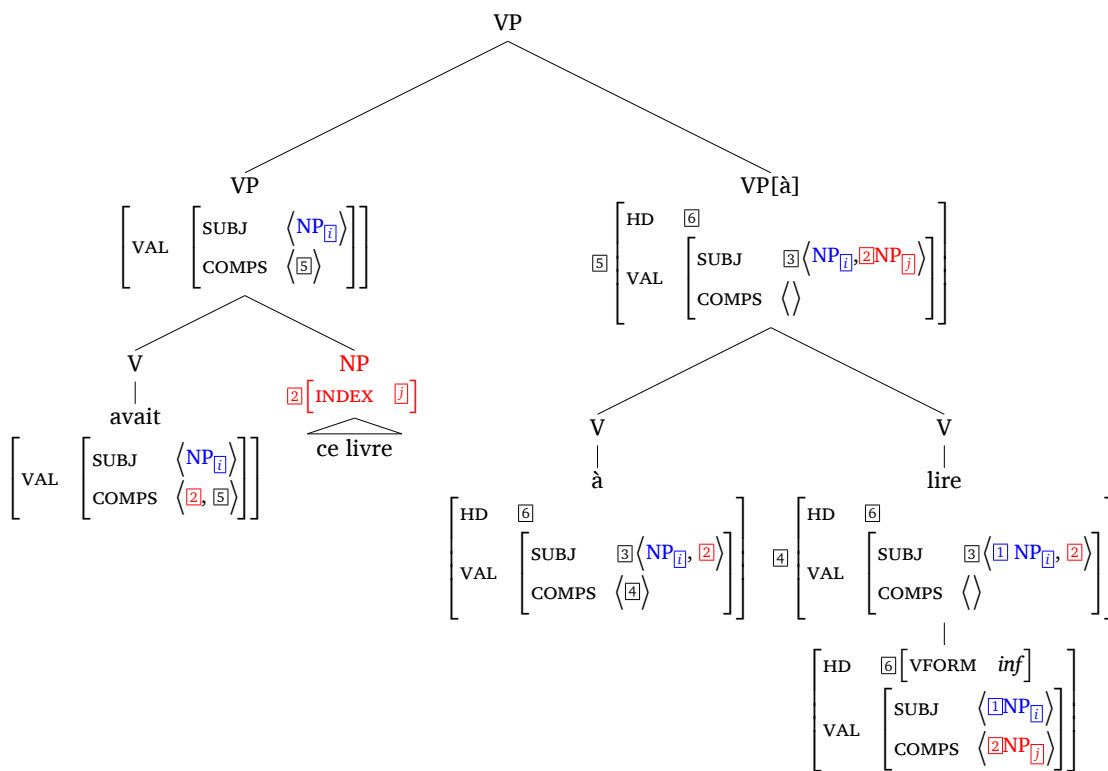


Figure 14: Lexical entry for a *tough* adjective

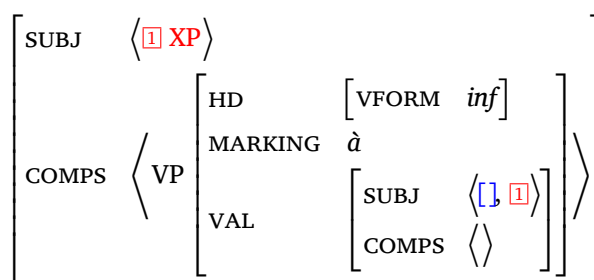


Figure 15: Sample analysis for *facile à lire*

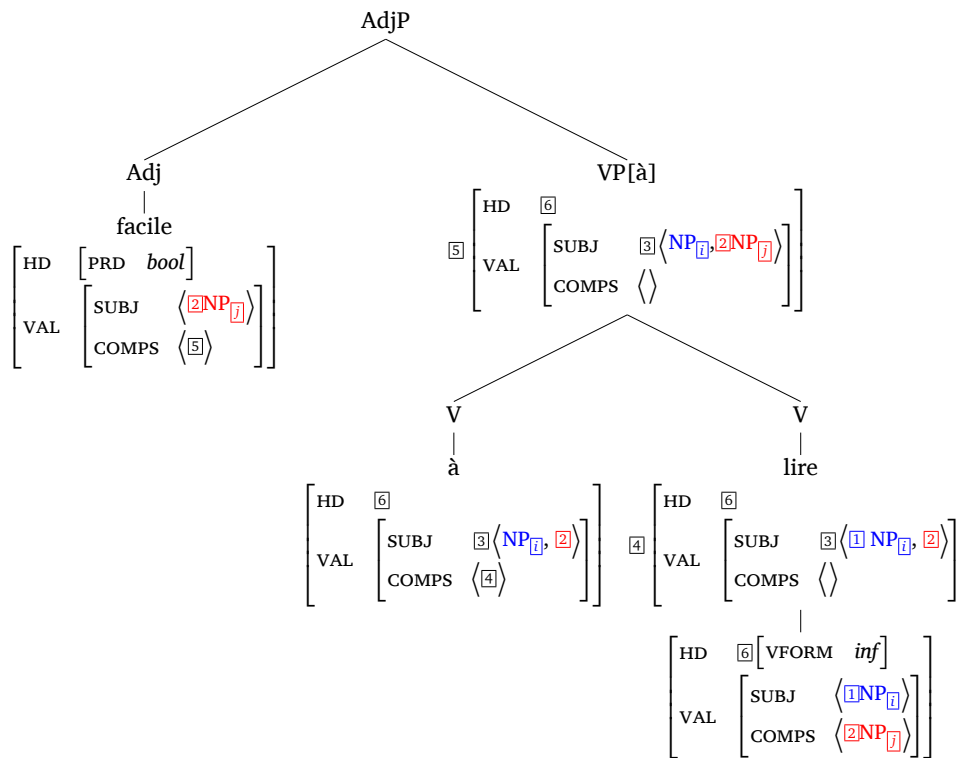


Figure 16: Lexical entry for *coûter*

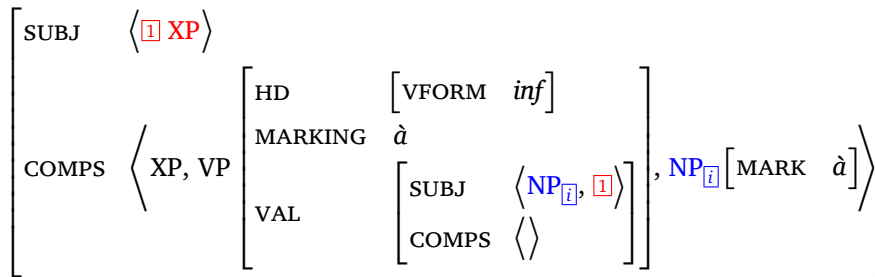


Figure 17: Lexical entry for a transparent subject control verb

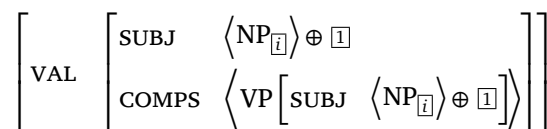


Figure 18: Lexical entry for a transparent subject raising verb

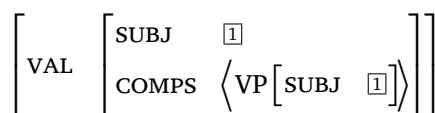
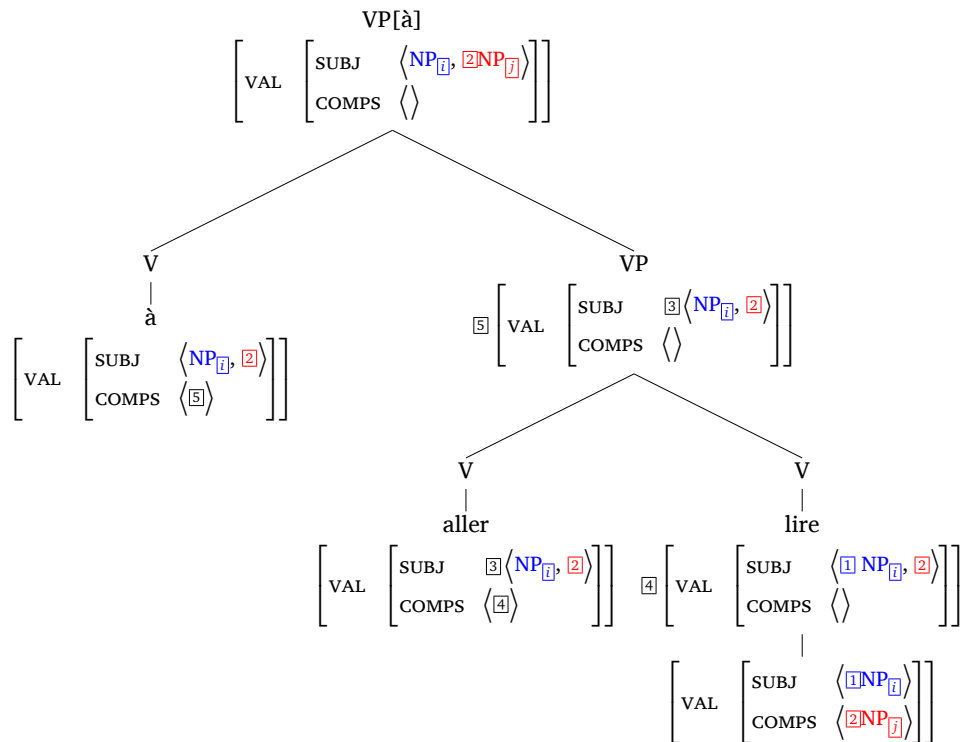


Figure 19: Sample analysis for *à aller lire***Figure 20:** Sample analysis for *à pouvoir lire*