Categorial Grammar and French Syntax

Richard Moot
• In this first talk about the principles of categorial grammars for analyzing the French language, I will talk about the basic building blocks.
• The goal is to give an analysis which is as simple as possible.
• We will use the annotated treebank of Paris VII as a guideline.
• I will start at the very beginning, with the basic building blocks (the atomic formulas)
• I will (informally) introduce the concepts of categorial grammar as we progress.
• Even though we start simple, we will quickly see there are some tough choices to make.

• I welcome any comments and suggestions for improvements!
## Atomic Formulas

<table>
<thead>
<tr>
<th>Case</th>
<th>Example</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>np</td>
<td><em>Paris, Milan, M. John Major, cette époque</em></td>
<td>noun phrase</td>
</tr>
<tr>
<td>n</td>
<td><em>mois, secrétaire d’Etat, qualité</em></td>
<td>noun or common noun</td>
</tr>
<tr>
<td>pp</td>
<td><em>à Milan, à son frère Michel, de contrefaçons</em></td>
<td>prepositional phrase</td>
</tr>
<tr>
<td>s</td>
<td><em>Il s’agissait de contrefaçons, Libération cherche à augmenter ses recettes</em></td>
<td>sentence</td>
</tr>
</tbody>
</table>
Atomic Formulas

- np
- n
- s
  - Sppart, Sppres, Sinf, Sainf, Sdeinf
- pp
  - ppà, ppde, pppar, ...

We will introduce a sort of "subtyping", some slightly more refined types to make useful selection restrictions on the arguments. These subtypes will be motivated by several examples on the slides which follow.
Complex Formulas

- All atomic formulas are formulas
- If A and B are formulas, then A/B and B\A are formulas as well

\[
\frac{A/B}{A} B \quad \frac{B}{A} B\backslash A
\]
Complex Formulas

A transitive verb looks first for a noun phrase to its right ("Marie" in this case) to form an intransitive verb np\s

Then it combines with a noun phrase to its left ("Jean" in this case) to form a sentence.
The table shows the formulas for different parts of speech in French:

<table>
<thead>
<tr>
<th>np</th>
<th>(np \s) / np</th>
<th>np / n</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jean</td>
<td>aime</td>
<td>la</td>
<td>logique</td>
</tr>
</tbody>
</table>

The determiner “la” looks for a noun to its right to form a noun phrase. Therefore, the two expressions “la logique” and “Marie” can be assigned the same formula: np.
### Complex Formulas

<table>
<thead>
<tr>
<th>np</th>
<th>(np\s)/pp</th>
<th>pp/np</th>
<th>np/n</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Il</td>
<td>travaillait</td>
<td>à</td>
<td>une</td>
<td>réforme</td>
</tr>
</tbody>
</table>

Some verbs select a prepositional argument (though sometimes the distinction between argument and modifier can be subtle, as we will see in the slides which follow).
**Complex Formulas**

<table>
<thead>
<tr>
<th>np</th>
<th>((np/s)/pp)</th>
<th>pp/np</th>
<th>np/n</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Il travaillait</td>
<td>à</td>
<td>une</td>
<td>table</td>
<td></td>
</tr>
</tbody>
</table>

Difference between arguments and modifiers
The difference between this analysis and the one on the previous slide is that here “à une table” is analyzed as a locative, the table is the place where the subject is working, whereas on the previous analysis, the table was the object being made by the subject.
**Complex Formulas**

<table>
<thead>
<tr>
<th>np</th>
<th>(np\s)/pp</th>
<th>pp/np</th>
<th>np/n</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Il</td>
<td>travaillait</td>
<td>à</td>
<td>une</td>
<td>réforme</td>
</tr>
</tbody>
</table>
Complex Formulas

<table>
<thead>
<tr>
<th></th>
<th>np</th>
<th>(np\s)/pp</th>
<th>pp/np</th>
<th>np/n</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>*</td>
<td>Il</td>
<td>travaillait</td>
<td>dans</td>
<td>une</td>
<td>réforme</td>
</tr>
</tbody>
</table>

perhaps this seems strange

The verb "travailler" does not allow a locative pp as an argument. Though "travailler dans une école" is certainly possible, I prefer not to analyse this by having "travailler" select a locative pp argument, since this argument is not "inherent" in the verb. Other verbs, such as "rester" or "aller" do have a good reason to select a pp argument.

pp's are generally difficult: they can be either arguments of the verb, adverbs (modifying the verb) or noun modifiers and the distinctions (especially between adverb and verb argument) are sometimes subtle. Fortunately, the functional annotation of the Paris VII corpus and resources such as Dicovalence help make the right decisions.
**Complex Formulas**

<table>
<thead>
<tr>
<th>np</th>
<th>(np\s)/ppà</th>
<th>ppà/np</th>
<th>np/n</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Il</td>
<td>travaillait</td>
<td>à</td>
<td>une</td>
<td>réforme</td>
</tr>
</tbody>
</table>

?! This seems strange

It makes sense to have the verb indicate which prepositions it select as arguments.
Complex Formulas

<table>
<thead>
<tr>
<th>np</th>
<th>(np\s)/ppₐ</th>
<th>ppₐ/np</th>
<th>np/n</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Il</td>
<td>travaillait</td>
<td>à</td>
<td>une</td>
<td>réforme</td>
</tr>
</tbody>
</table>

This multiplies the number of formulas for verbs, but is justified by

"travailler" combines with several prepositions: "à", "sur", "avec", "pour", "contre".

The difference between "travailler à" and "travailler sur" is a bit subtle: "travailler sur" suggest work on an item already in a more-or-less finished state, a work of refinement, whereas "travailler à" is more a work of creation.
### Complex Formulas

<table>
<thead>
<tr>
<th>np</th>
<th>(np\s)/pp\text{sur}</th>
<th>pp\text{sur}/np</th>
<th>np/n</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Il</td>
<td>travaillait</td>
<td>sur</td>
<td>une</td>
<td>réforme</td>
</tr>
</tbody>
</table>

This multiplies the number formulas for many verbs, but is justified by semantic differences.
Complex Formulas

<table>
<thead>
<tr>
<th>np</th>
<th>(np\s)/pp_avec</th>
<th>pp_avec/np</th>
<th>np/n</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Il</td>
<td>travaillait</td>
<td>avec</td>
<td>le</td>
<td>gouverment</td>
</tr>
</tbody>
</table>

This multiplies the number formulas for many verbs, but is justified by semantic differences
Complex Formulas

Example from the corpus
Complex Formulas

Example from the corpus

\texttt{augmente: (np/s)/pp_{de}}
Similarly, the verbs “être” and “avoir” combine with past participles and clitics. Unfortunately, the Paris VII corpus analyses “est arrivé” as a single VN constituent (in general the verb cluster and clitics are treated in this way). This means that some rebracketing is necessary to get the analysis displayed here.

<table>
<thead>
<tr>
<th>np</th>
<th>(np\s)/(np\s)</th>
<th>(np\ s)/pp_a</th>
<th>pp_a/np</th>
<th>np</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jean</td>
<td>est</td>
<td>arrivé</td>
<td>à</td>
<td>Paris</td>
</tr>
</tbody>
</table>
Similarly, the verbs “être” and “avoir” combine with past participles not with simple tensed verbs.
We can code this selection restriction in the formula for “est”
A more complex example

- Le raffermissement a été provoqué par des ventes de marks au profit du yen, à de nouvelles appréhensions sur le sort de l'ex-URSS - hypothèque qui ne cessera de peser sur la devise allemande dans les prochains mois - et à l'évolution de la conjoncture en Allemagne.
## Complex Formulas

<table>
<thead>
<tr>
<th>Formula</th>
<th>Cessera</th>
<th>De</th>
<th>Peser</th>
<th>Sur</th>
<th>...</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>(np \ s) / (np \ s_{deinf})</code></td>
<td>cessera</td>
<td>de</td>
<td>peser</td>
<td>sur</td>
<td>...</td>
</tr>
<tr>
<td><code>(np \ s_{deinf}) / (np \ s_{inf})</code></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><code>(np \ s_{inf}) / pp_{sur}</code></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><code>pp_{sur} / np</code></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><code>np</code></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Dès le 9 août dernier, M. Pierre Jaans, directeur de l'Institut monétaire luxembourgeois (IML), autorité bancaire centrale du Grand-Duché, avait menacé de demander à la Cour de Luxembourg de liquider la BCCI SA (Le Monde du 9 août 1991).
Complex Formulas

<table>
<thead>
<tr>
<th>((np \ s\text{inf}) / (np \ s\text{deinf})) / pp_à</th>
<th>pp_à</th>
<th>(np \ s\text{deinf}) / (np \ s\text{inf})</th>
<th>(np \ s\text{inf}) / np</th>
<th>np</th>
</tr>
</thead>
<tbody>
<tr>
<td>demander</td>
<td>à ...</td>
<td>de</td>
<td>liquider</td>
<td>...</td>
</tr>
</tbody>
</table>

More on these kinds of examples when I talk about semantics (subject control versus object control)

“Jean persuade Marie de venir” “Jean aide Marie à venir” (Marie vient)
“Jean aime venir” (Jean vient”)

“Jean demande de louer un livre” (Jean love un livre)
“Jean demande Marie de louer un”
Apparemment, cette reprise tarde à venir, si tant est qu'elle vienne et, surtout, élément capital sur les marchés de change, l'écart de taux d'intérêt entre les Etats-Unis et l'Allemagne s'est encore creusé cette semaine, passant de 6 % à 6,50 %.
## Complex Formulas

<table>
<thead>
<tr>
<th>np/n</th>
<th>n</th>
<th>(np \ s)/(np \ s_ainf)</th>
<th>(np \ s_ainf)/(np \ s_inf)</th>
<th>np \ s_inf</th>
</tr>
</thead>
<tbody>
<tr>
<td>cette</td>
<td>reprise</td>
<td>tarde</td>
<td>à</td>
<td>venir</td>
</tr>
</tbody>
</table>

"Il me tarde de aller à Paris" is also possible, but again the semantics - as well as the syntax - are quite different (eg. impersonal "il" only)
Present Participle

• The present participle is most frequently used in the context “en + (np \ s_{ppres})” (eg. “en voyageant”, “en franchissant les frontières”)

• Therefore (s \ s) / (np \ s_{ppres}) is one of the possible formula assignments to “en”
Other Infinitive Groups

- infinitive groups headed by the preposition “pour” are not assigned the formula \( (\text{np} \backslash s_{\text{pourinf}})/(\text{np} \backslash s_{\text{inf}}) \) but the formula \( (s \backslash s)/(\text{np} \backslash s_{\text{inf}}) \) and (more rarely) \( pp_{\text{pour}}/(\text{np} \backslash s_{\text{inf}}) \), eg. “nécessaire pour contrôler l’inflation”

- the case for “sans” is similar.
Other Infinitive Groups

- finally, the combination "commencer / finir par + infinitive" is analysed by assigning \( pp_{par} / (np \backslash s_{inf}) \) to the word "par", eg. "certains finissent par trouver des différences"
Other Infinitive Groups

Good question: why not use \( pp_\text{à} / (np \setminus s_{\text{inf}}) \) and \( pp_{\text{de}} / (np \setminus s_{\text{inf}}) \)?
Other Infinitive Groups

Good question:
why not use pp_{à} / (np / s_{inf}) and pp_{de} / (np / s_{inf}) ?

• in favor: a verb like “commencer” is assigned (np \ s) / pp_{a} which permits both “commence à lire” et “commence à un livre”

• there are quite a few verbs which follow this pattern
Good question: why not use `pp_{a}/(np \backslash s_{\text{inf}})` and `pp_{de}/(np \backslash s_{\text{inf}})`?

- problem: quite a few other verbs allow only an `à/\text{deinf}` argument, eg. “les entreprises cherchent à diminuer leur coûts” contre “*les entreprises cherchent à une réduction de leur coûts”
- so it makes sense to distinguish `(np \backslash s_{\text{ainf}})` from `pp_{a}`
The extracted lexicon

- Formula assignments to the present tense verb form “fait”
- 124 occurrences in the corpus with 19 different formulas assigned to it.

![Pie chart showing the distribution of formula assignments.](chart.png)
Adjectival uses of Participles

La société Dassault systèmes, filiale de Dassault aviation, spécialisée dans la conception et la fabrication assistées par ordinateur (CFAO)

“spécialisée” is a past participle, but it is used as an adjective—that is to say a noun modifier—in the current sentence. This means that the correct formula to assign it would be (n\n)/pp_dans.

These cases are very frequent. It seems tempting to treat them with a (non-logical) lexical rule; this is the solution used for the CCGbank. It has the additional advantage that adverbs can keep their normal ((np\s)/(np\s)) formulas, combine with the past participle, then apply the non-logical rule to transform the result to n\n.

Currently, the extracted grammar gives explicit (n\n) categories to past participles used as adjectives (and (n\n)/(n\n) to adverbs modifying them). Lexical ambiguity is reduced (and supertagger performance is improved) by adopting the CCGbank strategy.
Adjectival uses of Participles

Les 60 % du capital restants demeureront tchèques

“restants” = present participle used as an adjective

“demeureront” is a copula verb and therefore has as one of its possibilities to select an adjective - like “tchèques” in the current example - to its right.

the second example shows how “demeureront” can also select a past participle, which is an indication that $n\backslash n$ is a good formula in this context as well

Les reporters demeureront attachés à leur antenne

Question: does it make sense to analyse “être” the same way as “demeurer” that is to say replacing all \( (np\backslash s)/(np \backslash s_ppart) \) formulas for “être” by \( (np\backslash s)/(n\backslash n) \)?
Au demeurant, M. Tapie se dit convaincu qu'il sortira blanchi de l' affaire Toshiba ...
Adverbial uses of Participles - “une fois”

Une fois les réformes lancées

une fois acquise la question de " l' équilibrage structurel " et donc résolue celle du déficit financier par des mesures nouvelles
Beyond AB grammars

Structural Rules

- Even though it is permitted to assign many formulas to a single word, sometime we can generalize the pattern of this set of formulas by using structural rules.

- The most important (and the clearest) case, is the case of adverbs.
Beyond AB grammars

Structural Rules

... condamnent à l’unanimité le projet

• In the example above, “condamner” is a transitive verb taking “le projet” as its object argument.

• However, the adverb “à l’unanimité”, which modifies the verb is placed between the verb and its object.
Beyond AB grammars

Structural Rules

... condamnent à l’unanimité le projet

• What formula do we assign to “à l’unanimité”?

• Preferably, the same formula would work for “condamnent le projet à l’unanimité” as well
Beyond AB grammars

Structural Rules

... condamnent à l’unanimité le projet

• The multimodal solution is to say: “à l’unanimité” has type $s_1s$ where the mode 1 indicates that the adverb (optionally) “move”
• this means “à” has formula $(s_1s)/np$
• in addition we have $(s_1s)/np \vdash (s/s)/np$

The definition of movement is very liberal in the extracted grammar. It is basically: modify an “s” anywhere. This is like a logic where a commutative and a non-commutative implication live together. An interesting open question is: what are the right restrictions on movement?
Beyond AB grammars

Structural Rules

• In the resulting system a non-commutative and a commutative logic live together.

• The commutative implication $A \setminus_1 B$ is allows generalizations of the lexical formulas assigned to words.
• Even though we have seen formulas such as \((np \backslash s)/(np \backslash s_{\text{part}})\) which take a complex formula as their argument, we have not (yet) done anything interesting with this property.
Beyond AB grammars

Introduction Rules

\[ \frac{A}{A/B} [/I]^i \]

\[ \frac{A}{B/A} \backslash I]^i \]

[B] is the rightmost free hypothesis in the proof

[B] is the leftmost free hypothesis in the proof

“...” contains at least one other non-discharged formula

Each introduction rule discharges exactly one formula.

A unique integer \( i \) links the rule application with this formula.
### Beyond AB grammars

#### Introduction Rules

<table>
<thead>
<tr>
<th>n</th>
<th>(n\n)/(s\np)</th>
<th>np</th>
<th>(np\s)/(np\sppart)</th>
<th>(np\sppart)/np</th>
</tr>
</thead>
<tbody>
<tr>
<td>redaction</td>
<td>qu’</td>
<td>on</td>
<td>a</td>
<td>créée</td>
</tr>
</tbody>
</table>
### Beyond AB grammars

#### Introduction Rules

<table>
<thead>
<tr>
<th></th>
<th>(n\n)/(s/np)</th>
<th>np</th>
<th>(np\s)/(np\s\ppart)</th>
<th>(np\s\ppart)/np</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>qu’</td>
<td>on</td>
<td>a</td>
<td>crééée</td>
</tr>
<tr>
<td>redaction</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[
\text{créée} \quad \frac{(np \s_{\text{ppart}})/np}{np \s_{\text{ppart}}}[/E]
\]
## Beyond AB grammars

### Introduction Rules

<table>
<thead>
<tr>
<th></th>
<th>(n\n)/(s/np)</th>
<th>np</th>
<th>(np\s)/(np\sppart)</th>
<th>(np\sppart)/np</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>redaction</td>
<td>qu’</td>
<td>on</td>
<td>a</td>
<td>créée</td>
</tr>
</tbody>
</table>

### Example

```
(a np\sppart)/np np\sppart [E]
(np\s)/(np\sppart) np\sppart [E]
```
# Beyond AB grammars

## Introduction Rules

<table>
<thead>
<tr>
<th>n</th>
<th>(n\n)/(s/np)</th>
<th>np</th>
<th>(np\s)/(np\s_sppart)</th>
<th>(np\s_sppart)/np</th>
</tr>
</thead>
<tbody>
<tr>
<td>redaction</td>
<td>qu’</td>
<td>on</td>
<td>a</td>
<td>crééée</td>
</tr>
</tbody>
</table>

créée

\[
\begin{align*}
\text{créée} & \quad \frac{a}{(np\_s\_sppart)/np} \quad np \\
\text{on} & \quad \frac{(np\_s)/(np\_s\_sppart)}{np\_s\_sppart} \\
\text{np} & \quad \frac{np\_s}{np\_s} \\
\text{s} & \quad \text{[\textbackslash E]} \\
\end{align*}
\]
## Beyond AB grammars

### Introduction Rules

<table>
<thead>
<tr>
<th>n</th>
<th>(\text{n}/(\text{s}/\text{np}))</th>
<th>np</th>
<th>(\text{(np}/\text{s})/(\text{np}/\text{s}_{\text{ppart}}))</th>
<th>(\text{(np}/\text{s}_{\text{ppart}})/\text{np})</th>
</tr>
</thead>
<tbody>
<tr>
<td>redaction</td>
<td>qu’</td>
<td>on</td>
<td>a</td>
<td>crééée</td>
</tr>
</tbody>
</table>

\[
\begin{align*}
\text{np} & \frac{\text{np}}{\text{s}} \frac{\text{s}_{\text{ppart}}}{\text{np}} \\
\text{on} & \frac{\text{np}}{\text{s}} \frac{\text{s}_{\text{ppart}}}{\text{np}} \\
\text{s} & \frac{\text{s}}{\text{np}}
\end{align*}
\]
Beyond AB grammars

Introduction Rules

<table>
<thead>
<tr>
<th>n</th>
<th>(n\n)/(s,np)</th>
<th>np</th>
<th>(np\s)/(np\sppart)</th>
<th>(np\sppart)/np</th>
</tr>
</thead>
<tbody>
<tr>
<td>redaction</td>
<td>qu’</td>
<td>on</td>
<td>a</td>
<td>créée</td>
</tr>
</tbody>
</table>

creée

\[
\begin{array}{c}
\text{on} \quad \frac{(np\s)/(np\sppart)}{np\sppart} \quad np\sppart[/E] \\
\text{np} \quad \frac{(np\s)/(np\sppart)}{\text{np}s} \quad \text{np}s[/E] \\
\text{qu’} \quad \frac{(n\n)/(s,np)}{s\text{np}} \quad s\text{np}[/I]\text{I} \\
n\n\end{array}
\]
Introduction Rules

<table>
<thead>
<tr>
<th>n</th>
<th>((n \backslash n) / (s/np))</th>
<th>np</th>
<th>((np \backslash s) / (np \backslash s_{ppart}))</th>
<th>((np \backslash s_{ppart}) / np)</th>
</tr>
</thead>
<tbody>
<tr>
<td>redaction</td>
<td>qu’</td>
<td>on</td>
<td>a</td>
<td>crééée</td>
</tr>
</tbody>
</table>

\[
\text{créée} \quad \frac{(np \backslash s_{ppart}) / np}{\text{[np]}} \quad [\text{E}]
\]
\[
\text{a} \quad \frac{(np \backslash s_{ppart}) / np}{\text{[np]}} \quad [\text{E}]
\]
\[
\text{on} \quad \frac{(np \backslash s) / (np \backslash s_{ppart})}{\text{[E]}}
\]
\[
\text{np} \quad \frac{np \backslash s}{\text{[\text{E}]}}
\]
\[
\text{qu’} \quad \frac{s}{\text{[I]}} \quad [\text{I}]
\]
\[
\text{n} \quad \frac{(n \backslash n) / (s/np)}{\text{[E]}}
\]
\[
\text{n} \quad \frac{n \backslash n}{\text{[E]}}
\]
### Beyond AB grammars

#### Introduction Rules

<table>
<thead>
<tr>
<th>n</th>
<th>(n\n)/(s/np)</th>
<th>np</th>
<th>(np\s)/(np\s_{ppart})</th>
<th>(np\s_{ppart})/np</th>
<th>s_{1}s</th>
</tr>
</thead>
<tbody>
<tr>
<td>redaction</td>
<td>qu’</td>
<td>on</td>
<td>a</td>
<td>crééée</td>
<td>ensemble</td>
</tr>
</tbody>
</table>

*Note: The table represents rules for natural language processing, where `n`, `np`, `s`, `ppart`, and `s_{ppart}` are categories or tokens, and `qu’` is a special token. The redaction tokens are used to mark parts of the text that are to be redacted or removed.*
## Beyond AB grammars

### Introduction Rules

<table>
<thead>
<tr>
<th>n</th>
<th>(n\n)/(s/□np)</th>
<th>np</th>
<th>(np\s)/(np\s_{ppart})</th>
<th>(np\s_{ppart})/np</th>
<th>s\1s</th>
</tr>
</thead>
<tbody>
<tr>
<td>redaction</td>
<td>que</td>
<td>on</td>
<td>a</td>
<td>crééée</td>
<td>ensemble</td>
</tr>
</tbody>
</table>

**Note:** The symbols and expressions used in the table represent grammatical rules or categories. The table provides a structured view of the rules governing the structure of sentences in the AB grammar system.
Beyond AB grammars

Introduction Rules

\[
\frac{A}{A/B} [/I]^i \\
\frac{A}{B/A} [\backslash I]^i
\]

[B] is the rightmost free hypothesis in the proof

[B] is the leftmost free hypothesis in the proof

“...” contains at least one other non-discharged formula

Each introduction rule discharges exactly one formula.

A unique integer \(i\) links the rule application with this formula.
Beyond AB grammars

Introduction Rules

\[ \frac{A}{A/\Box \Box B} \quad [/I]^i \]

[B] is a free hypothesis occurring anywhere in the proof

there is at least one other non-discharged formula

Each introduction rule discharges exactly one formula.

A unique integer $i$ links the rule application with this formula.

Note that this rule is just a "sugared" version of a set of derivation steps in a slightly more complicated logic. Details can be found in eg. Moortgat (1997)
Beyond AB grammars
A More complex example

Sentence as we find it in the corpus. "dont" is a relative pronoun like "que" but which selects a sentence missing de "de" preposition (instead of a sentence missing an np like "que")

Note how "dont" is annotated as a "de-obj" argument, which is useful.
Beyond AB grammars

A more complex example

However, the "de" preposition belongs to "responsable" (some adjectives select for prepositions: "responsable de X" functions as an adjective just as "responsable")

Remark however, that there is no way to derive this from the annotation as it is given. Manual intervention (or at least verification!) is unfortunately necessary to assure the correct placement of the hypothetical preposition.
Beyond AB grammars
A more complex example
Beyond AB grammars
A more complex example
Beyond AB grammars

A more complex example
Beyond AB grammars

A more complex example
Beyond AB grammars

A more complex example
Beyond AB grammars

A more complex example
Beyond AB grammars

A more complex example
## BEYOND AB GRAMMARS
### Introduction Rules

<table>
<thead>
<tr>
<th>np</th>
<th>(np \s)/(np \s_deinf)</th>
<th>(np \s_deinf)/(np \s_inf)</th>
<th>np</th>
<th>np -(np \s_inf)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Il</td>
<td>convient</td>
<td>de</td>
<td>tout</td>
<td>reviser</td>
</tr>
</tbody>
</table>
### Beyond AB grammars

#### Introduction Rules

<table>
<thead>
<tr>
<th>np</th>
<th>(np (s))/(np (s_{deinf}))</th>
<th>(np (s_{deinf}))/(np (s_{inf}))</th>
<th>(np (s_x))/((np (s_x))/np)</th>
<th>(np (s_{inf}))/np</th>
</tr>
</thead>
<tbody>
<tr>
<td>Il</td>
<td>convient</td>
<td>de</td>
<td>tout</td>
<td>reviser</td>
</tr>
</tbody>
</table>

**Advantage of this solution:** the formula for “reviser” is a simple transitive verb form for an infinitive.

In addition we do not generate “Il convient de les exercises reviser” (the class of pronouns which can occur to the left of an embedded VP seems to be rather limited: “tout”, “rien”, “ce” and “beaucoup” are the only ones which I found in the extracted grammar.

The s\_X here is a first-order quantifier which quantifies over the different types of sentences even though I think only s\_inf and s\_ppart are possible here.
une lettre qu’il lui a adressé

il a adressé une lettre à M. Eltsine
Beyond AB grammars

Clitics

une lettre qu’ il lui a adressé

il a adressé une lettre à M. Eltsine

a: (np \ s)/ (np \ s_{ppart})

adressé: ((np \ s_{ppart})/ pp_{a})/ np
Beyond AB grammarS

Clitics

une lettre qu’ il lui a adressé
il a adressé une lettre à M. Eltsine

que: (n\n)/s/◊np

a: (np\s)/(np\spart)
adressé: ((np\spart)/ppa)/np
une lettre qu’ il lui a adressé

il a adressé une lettre à M. Eltsine

que: (n \n) / s / □np

lui: ((np \s_x) \ (np \s_x) / □pp_a)

a: (np \s) / (np \s_{ppart})

adressé: ((np \s_{ppart}) / pp_a) / np
Conclusions

• We have applied the basic principles of categorial grammar to the analysis of the French language, using an annotated corpus as a guide.