

## ***Une grammaire NooJ pour le syntagme verbal noyau en français***

François Trouilleux – Clermont Université, Université Blaise-Pascal, EA 999, LRL – Novembre 2012

Ce document fournit une documentation de base sur une grammaire NooJ pour le syntagme verbal noyau en français: la licence d'utilisation, un mode d'emploi sommaire de la grammaire et une série de captures d'écran de la grammaire graphique NooJ. L'article de référence qui présente cette grammaire est le suivant :

François Trouilleux, 2013. "A Description of the French Nucleus VP Using Co-occurrence Constraints", in *Formalising Natural Languages with NooJ. Selected Papers from the NooJ 2012 International Conference*. A. Donabedian, V. Khurshudian, M. Silberztein, eds. Cambridge Scholars Publishing, 2013.

### ***Licence***

La grammaire est distribuée selon les termes de la *GNU General Public License*, cf. <http://www.gnu.org/licenses/licenses.html#GPL>.

### ***Comment utiliser la grammaire avec NooJ***

La grammaire se présente sous la forme d'un fichier de projet NooJ nommé `French_Nucleus_VP.nop`. Pour ouvrir ce fichier, lancer NooJ ([www.nooj4nlp.net](http://www.nooj4nlp.net)) et sélectionner le menu **Project > Run Project**.

Le projet contient les fichiers suivants (utilisez le menu **Windows** de NooJ pour sélectionner les fenêtres) :

- `French_Nucleus_VP.nog`, la grammaire elle-même ;
- `FNVP_special.dic`, un petit dictionnaire de verbes à utiliser avec une priorité supérieure au dictionnaire `DM` ; les verbes de ce dictionnaire ont une information sur la transitivité qui est absente du `DM` (cf. <http://www.univ-bpclermont.fr/LABOS/lrl/spip.php?rubrique48>) ; la grammaticalité des chaînes pour lesquelles cet aspect est pertinent (*i.e.* la voie passive) ne devrait être testée qu'avec l'échantillon de verbes de `FNVP_special.dic` ;
- `FNVP_sample-text.not`, un court texte illustrant la façon dont la grammaire annote du texte ;
- `properties.def`, a fichier de propriétés NooJ ; il spécifie le vocabulaire utilisée dans les contraintes.

Pour naviguer dans les différents graphes de la grammaire, la fenêtre de la grammaire étant active, sélectionnez **GRAMMAR > Show structure**.

Pour afficher le contrat de la grammaire, *i.e.* un ensemble de chaînes grammaticales et agrammaticales qu'elle devrait admettre ou refuser, la fenêtre de la grammaire étant active, sélectionnez **GRAMMAR > Show contract**. Après avoir cliqué sur le bouton **Check**, toutes les chaînes devraient apparaître en vert. Si ce n'est pas le cas, merci de m'informer du problème, en m'indiquant la version de NooJ que vous avez utilisée (cf. **Info > About NooJ**).

Pour voir les annotations du texte exemple, cochez la case **Show Text Annotation Structure** dans la fenêtre du texte. Vous pouvez aussi exporter les annotations sous forme de balises XML: la fenêtre du texte étant active, sélectionnez **TEXT > Export annotated text as an XML document**. Vous devriez obtenir un pseudo fichier XML nommé `FNVP_sample-text.xml.txt`. Notez que tandis que la grammaire reconnaît *elle n'oublia pas* et *il était engagé*, elle annote *elle n'oublia* et *il était* comme des `NV`, *pas* comme un `GR`, et *engagé* comme un `NVPP` (une segmentation à la EASY). Le texte contient des exemples significatifs pour le traitement des participes passés :

- dans *Il serait, peut-être, deux ans parti*, la grammaire ne parvient pas à identifier *parti* comme un `NV` participe passé ; elle ne gère que des séquences plus simples entre l'auxiliaire et le participe ;

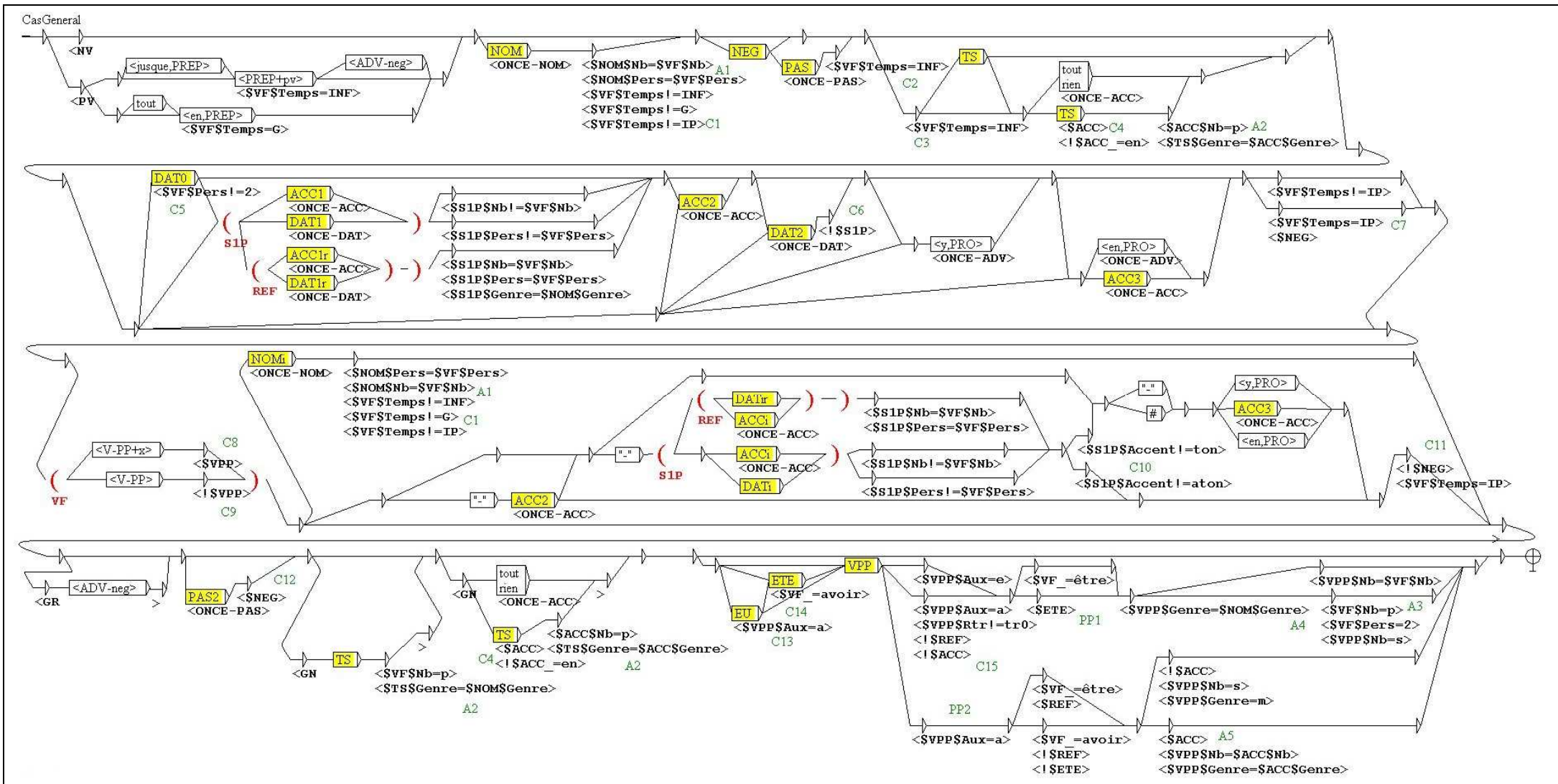
- dans *N'avez-vous pas crainte*, la grammaire, de façon correcte, n'identifie pas *crainte* comme un NV participe passé, parce qu'il n'y a pas de complément d'objet direct féminin singulier avant le verbe ;
- dans *leurs grands-parents défunts que les Liébard avaient connus*, la grammaire ne parvient pas à identifier *connus* comme un NV participe passé, parce qu'elle n'identifie pas le complément d'objet direct *que* avant le verbe ; ce type de cas est hors de portée de la grammaire ; seuls les pronoms clitiques COD sont accessibles à la grammaire.

Voir le manuel NooJ pour d'autres usages (analyser vos propres textes, modifier la grammaire, etc.).

### Captures d'écran commentées

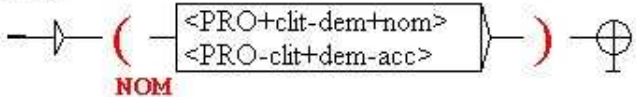

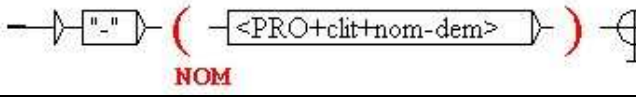
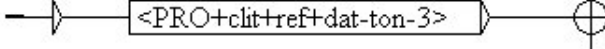
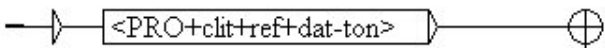
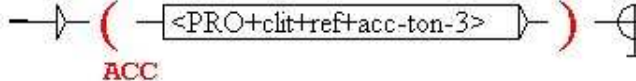
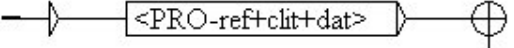
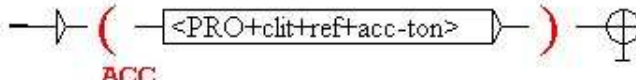
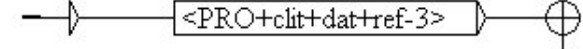
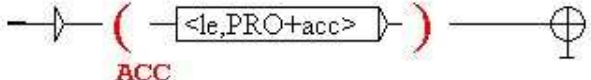
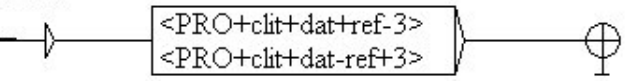
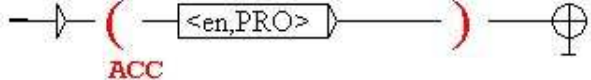
Les captures d'écran de la grammaire NooJ qui suivent ont un double but : (1) elles nous permettent de documenter quelques points non abordés dans l'article de référence, (2) elles fournissent un aperçu de la grammaire sans avoir à passer par un téléchargement du logiciel NooJ.

	<p>Graphe principal (axiome) Le sous-graphe CasGeneral est celui qui est documenté dans l'article.</p>
	<p>Trois cas particuliers :</p> <ul style="list-style-type: none"> <li>- les présentatifs : <i>en voici,</i> <i>le voilà ;</i></li> <li>- les séquences avec <i>ce</i> en fonction d'objet : <i>ce faisant ;</i></li> <li>- les séquences avec <i>ce</i> en fonction de sujet : <i>c'est,</i> <i>n'est-ce pas,</i> <i>c'aurait.</i></li> </ul>
	<p>Le graphe PrepPro ne fait pas partie de la définition du SVN. Il est inclus pour éviter que certains pronoms soient analysés comme des clitiques lors de l'analyse de textes (p.ex. dans <i>au dedans de lui grondait un orage</i>).</p>



Les prépositions au début du graphe ne sont pas commentées dans l'article. Deux contraintes leur sont associées :

- la préposition *en* exige un verbe au participe présent ;
- les prépositions marquées +pv exigent un verbe à l'infinitif (voir le DM pour le trait pv)

<p>NOM</p> 	<p>Les pronoms démonstratifs sont inclus dans le SV comme sujets.</p>	<p>DAT0</p> 	<p>Selon nous, les pronoms datif éthique sont toujours de 2e personne. Le DM contient les suites <i>te me</i>, <i>te nous</i> et <i>te vous</i> comme des datifs éthiques atomiques (+eth)</p>
<p>NOMi</p> 		<p>DAT1</p>  <p>DAT1r</p> 	
<p>ACC1</p> 	<p>Les pronoms « Slot 1 » avant le verbe ne doivent pas être marqués +ton (<i>moi</i>, <i>toi</i>). Les non réfléchis sont seulement de 1<sup>re</sup> ou 2<sup>e</sup> personne.</p>	<p>DAT2</p> 	
<p>ACC1r</p> 		<p>DATir</p> 	
<p>ACC2</p> 		<p>DATi</p> 	
<p>ACC3</p> 			
<p>ACCi</p> 