

# ATOLL - INRIA Rocquencourt

<http://atoll.inria.fr>

## DTD 2.0 pour les TAGs

Éric de la Clergerie

[Eric.De\\_La\\_Clergerie@inria.fr](mailto:Eric.De_La_Clergerie@inria.fr)

## DTD 2.0

- issue de la réunion de travail à Nancy en Avril 2002
- évolution principale :
  - formalisation TEI des structures de traits dans une DTD FS externe (fs001.dtd)
  - formalisation de bibliothèques de templates (morpho, syntaxe et autres)
  - remplacement de anchor par lexicalisation
  - normalisation de certains noms d'éléments et d'attributs
- essai de conversion automatique sur la petite TAG du français

# Bibliothèque de valeurs (DTD FS)

---

```
<fvLib type="fvLib">  
  <minus id="minus"/>  
  <sym id="pres" value="pres"/>  
  <sym id="moinshum" value="moinshum"/>  
  <sym id="3" value="3"/>  
  <vAlt id="ind|subj">  
    <sym value="ind"/>  
    <sym value="subj"/>  
  </vAlt>  
</fvLib>
```

---

# Bibliothèque de traits (DTD FS)

```
<fLib type="fLib">
  <f fVal="sing" id="npers:sing" name="npers"/>
  <f fVal="ind|inf" id="mode:ind|inf" name="mode"/>
  <f fVal="plus" id="poss:plus" name="poss"/>
  <f fVal="minus" id="inv:minus" name="inv"/>
  <f fVal="minus" id="comp:minus" name="comp"/>
  <f fVal="masc" id="gen:masc" name="gen"/>
  <f fVal="minus" id="quest:minus" name="quest"/>
  <f fVal="plus" id="wh:plus" name="wh"/>
  <f fVal="minus" id="poss:minus" name="poss"/>
  <f fVal="plus" id="comp:plus" name="comp"/>
  <f fVal="plus" id="quest:plus" name="quest"/>
</fLib>
```

# Bibliothèque de structures ou templates morpho (DTD FS)

---

```
<fsLib type="fsLib">
  <fs feats="gen:fem_num:sing" id="gen:fem@num:sing"/>
  <fs feats="mode:ind|inf" id="mode:ind|inf"/>
  <fs feats="poss:plus" id="poss:plus"/>
  <fs feats="npers:pl_num:sing_pers:3" id="npers:pl@num:sing@pers:3"/>
  <fs feats="mode:subj_num:sing" id="mode:subj@num:sing"/>
  <fs feats="mode:subj_num:pl" id="mode:subj@num:pl"/>
  <fs feats="restr:plushum_wh:plus" id="restr:plushum@wh:plus"/>
  <fs feats="num:sing_pers:3" id="num:sing@pers:3"/>
  <fs feats="npers:sing_num:pl_pers:3" id="npers:sing@num:pl@pers:3"/>
  <fs feats="mode:ind|subj_num:sing" id="mode:ind|subj@num:sing"/>
</fsLib>
```

---

# Morphologie

DTD 1.0 :

```
<morph lex="donne">
  <lemmeref cat="v" name="*DONNER*">
    <fs>
      <f name="num"><val>sing</val></f>
      <f name="mode"><val>ind</val><val>subj</val></f>
    </fs>
  </lemmeref>
</morph>
```

DTD 2.0 : (avec utilisation d'un template morpho)

```
<morph lex="donne">
  <lemmeref cat="v" name="*DONNER*">
    <fs copyof="mode:ind|subj@num:sing"/>
  </lemmeref>
</morph>
```

# Lemmes

DTD 1.0 :

---

```
<lemma cat="v" name="*DONNER*">
  <anchor tree_id="family[@name=tn1pn2]">
    <coanchor node_id="p_2" > <lex>à</lex> </coanchor>
    <equation node_id="np_0" type="top">
      <fs> <f name="restr"><val>plushum</val></f> </fs>
    </equation>
    <equation node_id="np_2" type="top">
      <fs> <f name="restr"><val>plushum</val></f> </fs>
    </equation>
  </anchor>
</lemma>
```

---

DTD 2.0 : disparus (implicites dans les lemmaref et les lexicalisations)

# Lexicalisation (DTD 2.0)

---

```
<!ELEMENT lexicalization ((family|tree)+,anchor+, equation*)>
```

---

```
<lexicalization>
  <family copyof="tn1pn2"/>
  <anchor node="v_">
    <lemmeref cat="v" name="*DONNER*" />
  </anchor>
  <anchor node="node[@name=p_2]">
    <lemmeref cat="p" name="à" />
  </anchor>
  <equation node="node[@name=np_0]" type="top">
    <fs copyof="restr:plushum" />
  </equation>
  <equation node="node[@name=np_2]" type="top">
    <fs copyof="restr:plushum" />
  </equation>
</lexicalization>
```

---



# Lexicalisation avec templates syntaxiques

```
<equationLib>
  <equation id="np0_plushum" node="node[@name=np_0]" type="top">
    <fs copyof="restr:plushum"/>
  </equation>
  <equation id="np2_plushum" node="node[@name=np_2]" type="top">
    <fs copyof="restr:plushum"/>
  </equation>
</equationLib>
```

```
<lexicalization type="syntaxe">
  <family copyof="tn1pn2"/>
  ...
  <equation copyof="np0_plushum"/>
  <equation copyof="np2_plushum"/>
</lexicalization>
```

# Arbres et familles

Similaires entre DTD 1.0 et 2.0 aux structures de traits près.

---

```
<family name="tn1pn2">
  <tree name="r0tn1pn2">
    ...
    <node adj="no" lex="qui" type="lex"/>
    <node adj="yes" cat="s" type="std">
      <narg type="bot">
        <fs feats="inv:minus">
          <f name="mode" varname="X2"/>
        </fs>
      </narg>
    <node adj="no" type="lex"/>
    ...
  </tree>
  ...
</family>
```

---

# Noeuds

Un attribut supplémentaire sur les noeuds pour les types d'adjonctions autorisés.

---

```
<!ENTITY % adj.kind '( wrapping | left | right | left_right)'\>
<!ELEMENT node (narg*,node*)>
<!ATTLIST node  cat    %cat.type;      #IMPLIED
                 lex    %lex.type;      #IMPLIED
                 type  %node.type;      #REQUIRED
                 name  CDATA            #IMPLIED
                 adjkind %adj.kind;     #IMPLIED
                 adj    %adj.type;      #IMPLIED>
```

---

# Autres bibliothèques

Possibilité de grouper en bibliothèques

- les familles,
- les lexicalisations,
- les entrées morphologiques
- les arbres.

---

```
<!ELEMENT morphLib (morph+)>
<!ATTLIST morphLib %a.Lib; >
<!ELEMENT lexicalizationLib (lexicalization+)>
<!ATTLIST lexicalizationLib %a.Lib; >
<!ELEMENT familyLib (family+)>
<!ATTLIST familyLib %a.Lib; >
<!ELEMENT treeLib (tree+)>
<!ATTLIST treeLib %a.Lib; >
```

---

**Note** les bibliothèques d'arbres ne sont pas des familles

# Métadata

---

```
<!-- remplace desc -->
<!ELEMENT adminGrp (admin)+>
<!ATTLIST admin type CDATA #REQUIRED>
<!ELEMENT admin ANY>
```

---

---

```
<morph lex="après">
  <adminGrp>
    <admin>EVDL April 8th 2002: added missing entry </admin>
  </adminGrp>
  <lemmaref cat="p" name="après"/>
</morph>
```

---

Discussion sur le statut des traits issus d'une méta-grammaire

# Élément chapeau

---

```
<!ELEMENT tagml (adminGrp?, ( tree|treeLib|
    family|familyLib|
    morph|morphLib|
    lexicalization|lexicalizationLib|
    fs|fsLib|
    equation|equationLib
    )*)>
```

---

---

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<!DOCTYPE tag PUBLIC "-//INRIA//DTD_TAG_2.01a//EN"
    "http://atoll.inria.fr/~clerger/tag201a.dtd,xml">
<tagml>
...
</tagml>
```

---

# Evolution

- Nouvelles DTD (FS 001a et TAG 201a) disponibles en lignes
- DTD alpha car en attente de retour
  - à terme, l'attribut node des éléments anchor et equation doit être un XPointer

---

```
<equation type="top" node="node[name=np_0]"> ... </equation>
```

---

- pouvoir mettre des équations dans les éléments tree en place des narg
- Outils ATOLL non encore mis en conformité