

Syntactic Reference Corpus of Medieval French
et logiciel d'annotation *NotaBene*

Nicolas Mazziotta

Universität Stuttgart, ILR

18 mai 2011, Meditext, Paris

Plan

Exposé en deux parties :

1. Présentation générale de SRCMF
2. Introduction aux fonctionnalités de NotaBene

SRCMF

Présentation du projet

Modèle syntaxique dépendancier

Outils et procédures

Ressources créées

NotaBene

Présentation rapide

Philosophie générale

Annotation syntaxique (SRCMF)

Perspectives : intégration TXM
(CoSToMeF)

1^{re} partie : Syntactic Reference Corpus of Medieval French

Présentation du projet

Modèle syntaxique dépendancier

Outils et procédures

Ressources créées

Présentation du projet

- Fiche descriptive du projet
- Champ d'études et objectif
- Équipe et répartition des tâches

Modèle syntaxique dépendanciel

- Fondements
- Annotation syntaxique

Outils et procédures

- Outils de communication/centralisation
- Outil d'annotation
- Procédures

Ressources créées

- Formats de sortie et exploitation
- Documentation
- Prospectif : CoSToMeF

Fiche descriptive du projet

Dénomination

Syntactic Reference Corpus of Medieval French (SRCMF)

Direction scientifique

- ▶ Sophie Prévost (CNRS LaTTiCe, Paris)
- ▶ Achim Stein (ILR, Stuttgart)

Sources de financement

- ▶ Subvention française : ANR
- ▶ Subvention allemande : DFG

Durée

Décembre 2008 – décembre 2011

Champ d'études et objectif

Analyse syntaxique de textes en ancien français et besoin d'une ressource

- ▶ Question de linguistique historique
- ▶ Français écrit en France du 9^e au 14^e siècle
- ▶ Comment les phrases sont-elles construites ?

Champ d'études et objectif

Analyse syntaxique de textes en ancien français et besoin d'une ressource

- ▶ Question de linguistique historique
- ▶ Français écrit en France du 9^e au 14^e siècle
- ▶ Comment les phrases sont-elles construites ?

**Nombreuses ressources électroniques syntaxiques
pour l'histoire de l'anglais, pas pour le français**

⇒ Possibilité de renouvellement de l'approche de la langue

Champ d'études et objectif

Analyse syntaxique de textes en ancien français et besoin d'une ressource

- ▶ Question de linguistique historique
- ▶ Français écrit en France du 9^e au 14^e siècle
- ▶ Comment les phrases sont-elles construites ?

Nombreuses ressources électroniques syntaxiques
pour l'histoire de l'anglais, pas pour le français

⇒ Possibilité de renouvellement de l'approche de la langue

Corpus de travail

Deux corpus de ca 3 millions de mots issus de textes de genres divers :

- ▶ *Base de Français Médiéval* (BFM, ENS Lyon)
- ▶ *Nouveau Corpus d'Amsterdam* (NCA, ILR Stuttgart)

Champ d'études et objectif

Annotation de corpus

- ▶ L'analyse décrit les ressources d'un point de vue syntaxique
- ▶ Elle est **encodée par un humain** (analyse automatique peu satisfaisante)
- ▶ Elle prend la forme d'une ressource électronique liée au corpus

Champ d'études et objectif

Annotation de corpus

- ▶ L'analyse décrit les ressources d'un point de vue syntaxique
- ▶ Elle est **encodée par un humain** (analyse automatique peu satisfaisante)
- ▶ Elle prend la forme d'une ressource électronique liée au corpus

Objectif

- ▶ Constituer une ressource pour la communauté scientifique, en unissant les deux corpus selon un modèle identique
- ⇒ Résultat : *Syntactic **Reference** Corpus*
- ▶ pour alimenter les **recherches** syntaxiques (linguistique théorique)
 - ▶ pour permettre de créer des outils de **traitement automatique** de l'ancien français (linguistique computationnelle)

Équipe et répartition des tâches

- ▶ Direction du projet
 - ▶ Sophie Prévost (CNRS LaTTiCe, Paris)
 - ▶ Achim Stein (ILR, Stuttgart)
- ▶ Contractuels (y compris temps partiels)
 - ▶ Béatrice Bischof (ILR, Stuttgart)
 - ▶ Julie Glikman (CNRS LaTTiCe)
 - ▶ Alexei Lavrentiev (ENS, Lyon)
 - ▶ Nicolas Mazziotta (ILR, Stuttgart)
 - ▶ Tom Rainsford (University of Cambridge)
- ▶ Collaborateurs réguliers
 - ▶ Céline Guillot (ENS, Lyon)
 - ▶ Serge Heiden (ENS, Lyon)
 - ▶ Christiane Marchello-Nizia (ENS, Lyon)
- ▶ Experts
 - ▶ Fernande Dupuis (UQAM, Montréal)
 - ▶ Richard Ingham (University of Birmingham)
 - ▶ Bernard Victorri (ENS, Paris)

Équipe et répartition des tâches

- ▶ Direction du projet
 - ▶ **Sophie Prévost** (CNRS LaTTiCe, Paris)
 - ▶ **Achim Stein** (ILR, Stuttgart)
- ▶ Contractuels (y compris temps partiels)
 - ▶ **Béatrice Bischof** (ILR, Stuttgart)
 - ▶ **Julie Glikman** (CNRS LaTTiCe)
 - ▶ **Alexei Lavrentiev** (ENS, Lyon)
 - ▶ **Nicolas Mazziotta** (ILR, Stuttgart)
 - ▶ **Tom Rainsford** (University of Cambridge)
- ▶ Collaborateurs réguliers
 - ▶ **Céline Guillot** (ENS, Lyon)
 - ▶ **Serge Heiden** (ENS, Lyon)
 - ▶ **Christiane Marchello-Nizia** (ENS, Lyon)
- ▶ Experts
 - ▶ **Fernande Dupuis** (UQAM, Montréal)
 - ▶ **Richard Ingham** (University of Birmingham)
 - ▶ **Bernard Victorri** (ENS, Paris)

Rôles

- ▶ **modélisation linguistique**
- ▶ annotation
- ▶ développement d'outils
- ▶ contrôle qualité

Équipe et répartition des tâches

- ▶ Direction du projet
 - ▶ Sophie Prévost (CNRS LaTTiCe, Paris)
 - ▶ Achim Stein (ILR, Stuttgart)
- ▶ Contractuels (y compris temps partiels)
 - ▶ **Béatrice Bischof** (ILR, Stuttgart)
 - ▶ **Julie Glikman** (CNRS LaTTiCe)
 - ▶ **Alexei Lavrentiev** (ENS, Lyon)
 - ▶ **Nicolas Mazziotta** (ILR, Stuttgart)
 - ▶ **Tom Rainsford** (University of Cambridge)
- ▶ Collaborateurs réguliers
 - ▶ Céline Guillot (ENS, Lyon)
 - ▶ Serge Heiden (ENS, Lyon)
 - ▶ Christiane Marchello-Nizia (ENS, Lyon)
- ▶ Experts
 - ▶ Fernande Dupuis (UQAM, Montréal)
 - ▶ Richard Ingham (University of Birmingham)
 - ▶ Bernard Victorri (ENS, Paris)

Rôles

- ▶ modélisation linguistique
- ▶ **annotation**
- ▶ développement d'outils
- ▶ contrôle qualité

Équipe et répartition des tâches

- ▶ Direction du projet
 - ▶ Sophie Prévost (CNRS LaTTiCe, Paris)
 - ▶ **Achim Stein** (ILR, Stuttgart)
- ▶ Contractuels (y compris temps partiels)
 - ▶ Béatrice Bischof (ILR, Stuttgart)
 - ▶ Julie Glikman (CNRS LaTTiCe)
 - ▶ Alexei Lavrentiev (ENS, Lyon)
 - ▶ **Nicolas Mazziotta** (ILR, Stuttgart)
 - ▶ Tom Rainsford (University of Cambridge)
- ▶ Collaborateurs réguliers
 - ▶ Céline Guillot (ENS, Lyon)
 - ▶ **Serge Heiden** (ENS, Lyon)
 - ▶ Christiane Marchello-Nizia (ENS, Lyon)
- ▶ Experts
 - ▶ Fernande Dupuis (UQAM, Montréal)
 - ▶ Richard Ingham (University of Birmingham)
 - ▶ Bernard Victorri (ENS, Paris)

Rôles

- ▶ modélisation linguistique
- ▶ annotation
- ▶ **développement d'outils**
- ▶ contrôle qualité

Équipe et répartition des tâches

- ▶ Direction du projet
 - ▶ **Sophie Prévost** (CNRS LaTTiCe, Paris)
 - ▶ **Achim Stein** (ILR, Stuttgart)
- ▶ Contractuels (y compris temps partiels)
 - ▶ Béatrice Bischof (ILR, Stuttgart)
 - ▶ Julie Glikman (CNRS LaTTiCe)
 - ▶ Alexei Lavrentiev (ENS, Lyon)
 - ▶ Nicolas Mazziotta (ILR, Stuttgart)
 - ▶ Tom Rainsford (University of Cambridge)
- ▶ Collaborateurs réguliers
 - ▶ Céline Guillot (ENS, Lyon)
 - ▶ Serge Heiden (ENS, Lyon)
 - ▶ Christiane Marchello-Nizia (ENS, Lyon)
- ▶ Experts
 - ▶ Fernande Dupuis (UQAM, Montréal)
 - ▶ Richard Ingham (University of Birmingham)
 - ▶ Bernard Victorri (ENS, Paris)

Rôles

- ▶ modélisation linguistique
- ▶ annotation
- ▶ développement d'outils
- ▶ **contrôle qualité**

Présentation du projet

- Fiche descriptive du projet
- Champ d'études et objectif
- Équipe et répartition des tâches

Modèle syntaxique dépendanciel

- Fondements
- Annotation syntaxique

Outils et procédures

- Outils de communication/centralisation
- Outil d'annotation
- Procédures

Ressources créées

- Formats de sortie et exploitation
- Documentation
- Prospectif : CoSToMeF

Fondements

Modèle dépendanciel vs analyse en constituants

En deux mots :

Dépendance

Constituants

Fondements

Modèle dépendanciel vs analyse en constituants

En deux mots :

Dépendance

Constituants

Unités liées : liens spécifiques

Unités groupées : hiérarchie de groupes

Fondements

Modèle dépendanciel vs analyse en constituants

En deux mots :

Dépendance

Constituants

Unités liées : liens spécifiques

Unités groupées : hiérarchie de groupes

Fonctions

Natures et distribution

Fondements

Modèle dépendanciel vs analyse en constituants

En deux mots :

Dépendance

Constituants

Unités liées : liens spécifiques

Unités groupées : hiérarchie de groupes

Fonctions

Natures et distribution

Le chat ←-suj- *tue* -obj→ *la souris*

Fondements

Modèle dépendanciel vs analyse en constituants

En deux mots :

Dépendance	Constituants
Unités liées : liens spécifiques	Unités groupées : hiérarchie de groupes
Fonctions	Natures et distribution
<i>Le chat</i> ←-suj- <i>tue</i> -obj→ <i>la souris</i>	[SN <i>Le chat</i>][SV <i>tue</i> [SN <i>la souris</i>]]

Fondements

Modèle dépendanciel vs analyse en constituants

En deux mots :

Dépendance	Constituants
Unités liées : liens spécifiques	Unités groupées : hiérarchie de groupes
Fonctions	Natures et distribution
<i>Le chat</i> ←-suj- <i>tue</i> -obj→ <i>la souris</i>	[SN <i>Le chat</i>][SV <i>tue</i> [SN <i>la souris</i>]]

Le verbe au-dessus de la hiérarchie

- ▶ Marques spécifiques de portée propositionnelle
- ▶ Autonomie relative plus grande

Annotation syntaxique

Distinction entre les plans

- ▶ Morphosyntaxe/sémantique/énonciation
- ▶ Interaction permanente

Annotation syntaxique

Distinction entre les plans

- ▶ Morphosyntaxe/sémantique/énonciation
- ▶ Interaction permanente
- ▶ Informations encodées **purement syntaxiques**
 - ▶ Critères prioritairement syntaxiques (relations)
 - ▶ Toutefois, variété éteinte ⇒ besoin des autres plans

Annotation syntaxique

Distinction entre les plans

- ▶ Morphosyntaxe/sémantique/énonciation
- ▶ Interaction permanente
- ▶ Informations encodées **purement syntaxiques**
 - ▶ Critères prioritairement syntaxiques (relations)
 - ▶ Toutefois, variété éteinte ⇒ besoin des autres plans

Options théoriques

- ▶ Définition « mécanique » de la phrase
Basée sur *1 verbe* ⇔ *1 proposition*
- ▶ Place seconde pour les mots-outils (pas « translatifs »)
- ▶ Traitement orthogonal de la coordination (quasi symétrique)

Présentation du projet

Fiche descriptive du projet
Champ d'études et objectif
Équipe et répartition des tâches

Modèle syntaxique dépendanciel

Fondements
Annotation syntaxique

Outils et procédures

Outils de communication/centralisation
Outil d'annotation
Procédures

Ressources créées

Formats de sortie et exploitation
Documentation
Prospectif : CoSToMeF

Outils de communication/centralisation

Équipe large et **géographiquement dispersée**

Outils de communication/centralisation

Équipe large et **géographiquement dispersée**

Forum

- ▶ Discussions linguistiques
- ▶ Aide à distance

Wiki

- ▶ Archivage des décisions
- ▶ Stockage raisonné de la documentation

Subversion (SVN)

- ▶ Centralisation des données
- ▶ Suivi des versions

Outil d'annotation

NotaBene

Outil d'annotation

NotaBene

- ▶ Encodage de l'annotation syntaxique sous une forme ergonomique
- ▶ Personnalisation de l'environnement
- ▶ Comparaison entre deux annotations parallèles
- ▶ Exportation des résultats dans différents formats

Outil d'annotation

NotaBene

- ▶ Encodage de l'annotation syntaxique sous une forme ergonomique
- ▶ Personnalisation de l'environnement
- ▶ Comparaison entre deux annotations parallèles
- ▶ Exportation des résultats dans différents formats

Voir deuxième partie de l'exposé

- ▶ Présentation du logiciel
- ▶ Démarche d'annotation

Procédures

Besoin de procédures

Annotation manuelle

- ▶ Erreurs aléatoires
- ▶ Prises de position inconscientes
- ▶ Inconsistance

But : **diminuer** le nombre d'erreurs (0 erreurs impossible)

Procédures

Besoin de procédures

Annotation manuelle

- ▶ Erreurs aléatoires
- ▶ Prises de position inconscientes
- ▶ Inconsistance

But : **diminuer** le nombre d'erreurs (0 erreurs impossible)

Étapes

1. Annotation parallèle aveugle (deux annotateurs humains)
2. Correction croisée aveugle (même équipe)
3. Contrôles et décisions finals (directeurs)

Présentation du projet

- Fiche descriptive du projet
- Champ d'études et objectif
- Équipe et répartition des tâches

Modèle syntaxique dépendanciel

- Fondements
- Annotation syntaxique

Outils et procédures

- Outils de communication/centralisation
- Outil d'annotation
- Procédures

Ressources créées

- Formats de sortie et exploitation
- Documentation
- Prospectif : CoSToMeF

Formats de sortie et exploitation

Formats de sortie et exploitation

En interne

NotaBene utilise XML et RDF (voir deuxième partie)

Formats de sortie et exploitation

En interne

NotaBene utilise XML et RDF (voir deuxième partie)

Principaux formats pour le corpus SRCMF

- ▶ « Tel quel » (annotation)
- ▶ Graphviz (visualisation)
- ▶ TigerXML (exploration)
- ▶ CoNLL (entraînement des parseurs)

Formats de sortie et exploitation – « Tel quel »

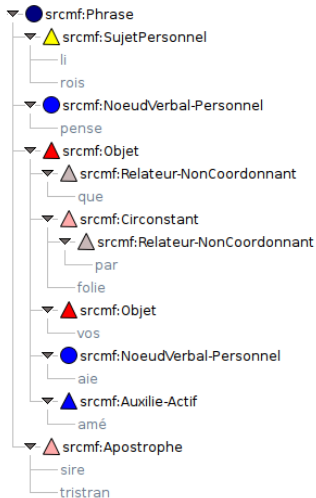
Annotation telle que présentée en interne

- ▶ Conventions de représentation à but **ergonomique**
 - ▶ Meilleur pour l'annotation que pour l'exploitation
- ⇒ Voir deuxième partie

Formats de sortie et exploitation – « Tel quel »

Annotation telle que présentée en interne

- ▶ Conventions de représentation à but ergonomique
 - ▶ Meilleur pour l'annotation que pour l'exploitation
- ⇒ Voir deuxième partie



Formats de sortie et exploitation – Graphviz

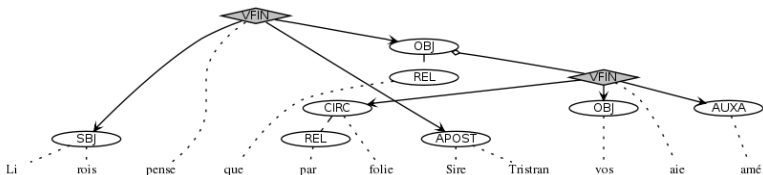
Graphviz et lecture humaine

- ▶ Outil de dessin de graphes (orientés ou non)
- ▶ Nombreuses options, rendu rapide

Formats de sortie et exploitation – Graphviz

Graphviz et lecture humaine

- ▶ Outil de dessin de graphes (orientés ou non)
- ▶ Nombreuses options, rendu rapide



Formats de sortie et exploitation – TigerXML

Format développé pour l'extraction de structures

- ▶ Fichiers XML faciles à générer/manipuler
- ▶ Peut encoder des structures dépendancielles/constituants
- ▶ Outil de **requête** avec syntaxe spécifique

Formats de sortie et exploitation – TigerXML

Format développé pour l'extraction de structures

- ▶ Fichiers XML faciles à générer/manipuler
- ▶ Peut encoder des structures dépendanciennes/constituants
- ▶ Outil de **requête** avec syntaxe spécifique

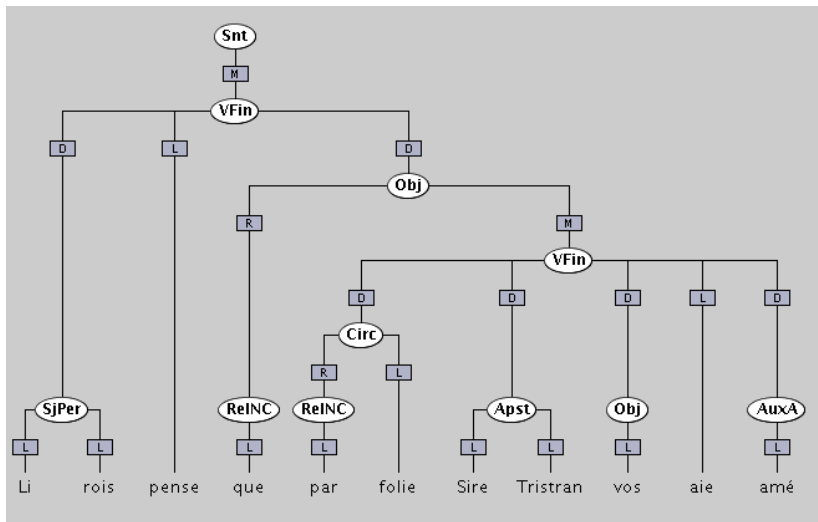
```
<graph root="_1263221221.22">
  <terminals>
    <t id="w26_00103" word="Li" />
    <t id="w26_00104" word="rois" />
    <t id="w26_00105" word="pense" />
    ...
  </terminals>
  <nonterminals>
    <nt id="_1263452434.16" cat="Obj">
      <edge label="R" idref="_1263453372.87" />
      <edge label="M" idref="_1263221225.03" />
    </nt>
    <nt id="_1263453372.87" cat="RelNC">
      <edge label="L" idref="w26_00106" />
    </nt>
    <nt id="_1263453407.18" cat="Obj">
      <edge label="L" idref="w26_00113" />
    </nt>
    ...
  </nonterminals>
</graph>
```

Formats de sortie et exploitation – TigerXML

Visualisation

Formats de sortie et exploitation – TigerXML

Visualisation



Formats de sortie et exploitation – TigerXML

Recherche

Syntaxe spécifique ; par exemple :

```
[cat="SjPer"] >M [cat="VFin"]
```

Tous les sujets dont le noeud principal est un verbe à un mode personnel

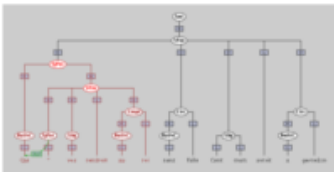
Formats de sortie et exploitation – TigerXML

Recherche

Syntaxe spécifique ; par exemple :

```
[cat="SjPer"] >M [cat="VFin"]
```

Tous les sujets dont le noeud principal est un verbe à un mode personnel



Formats de sortie et exploitation – CoNLL

Format optimisé pour le TAL

- ▶ Format tabulaire simple
- ▶ Permet d'**entraîner les parseurs**
- ▶ Moins expressif que RDF/TigerXML (perte)

Formats de sortie et exploitation – CoNLL

Format optimisé pour le TAL

- ▶ Format tabulaire simple
- ▶ Permet d'**entraîner les parseurs**
- ▶ Moins expressif que RDF/TigerXML (perte)

1	Li	-	-	-	-	-	-	2	-	Det	-	-	-	-
2	rois	-	-	-	-	morph=Nom	-	3	-	SjPer	-	-	-	-
3	pense	-	-	-	-	morph=VFin	-	0	-	ROOT	-	-	-	-
4	que	-	-	-	-	-	-	10	-	RelNC	-	-	-	-
5	par	-	-	-	-	-	-	6	-	RelNC	-	-	-	-
6	folie	-	-	-	-	morph=Nom	-	10	-	Circ	-	-	-	-
7	Sire	-	-	-	-	morph=Nom	-	10	-	Apst	-	-	-	-
8	Tristran	-	-	-	-	-	-	7	-	ModA	-	-	-	-
9	vos	-	-	-	-	-	-	10	-	Obj	-	-	-	-
10	aie	-	-	-	-	morph=VFin	-	3	-	Obj	-	-	-	-
11	amé	-	-	-	-	-	-	10	-	AuxA	-	-	-	-

Documentation

Fiche explicitant les choix d'annotation

Plan de la fiche

Remarques grammaticales

- \$1. Sujet nul
 - \$2. "Apposition", "épithète détachée", etc.
 - \$3. Sujet constitué uniquement d'un modifieur
- #### Remarques lexicales
- \$4. Sujet de "avenir"
 - \$5. Tout, trestuit...

Liste des fiches

Actant.xml.html
 Apet.xml.html
 AObj.xml.html
 ARfc.xml.html
 ASj.xml.html
 At.xml.html
 AuxA.xml.html
 AuxP.xml.html
 Aux_xml.html
 Circ.xml.html
 Cmpl.xml.html
 Conjoint.xml.html
 Coc.xml.html
 DiscoursDirect.xml.html
 Duplicata.xml.html
 Fonction.xml.html
 GpCoc.xml.html
 Groupe.xml.html
 Ignorer.xml.html
 Inrt.xml.html
 Intj.xml.html
 Lacune.xml.html
 Max.xml.html
 ModA.xml.html
 ModD.xml.html

Sujet personnel (SjPer)

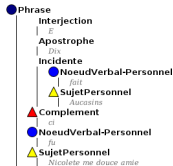
\$0.

Équivalent traditionnel

sujet, à l'exclusion du sujet dit "impersonnel" et du sujet dit "apparent"

Exemples-types

- (1) E ! Dix, fait Aucasins, ci fu Nicolette me douce amie, (AucR3 p. 27) [Masquer l'analyse](#)



- (2) A Deu, qui est sire du mont, | Cri ge merci, (TristBérM4 p. 67 v. 2185-2188) [\[Afficher l'analyse\]](#)

Remarques grammaticales

\$1. Sujet nul

Voir [→\\$1](#).

\$2. "Apposition", "épithète détachée", etc.

Les cas d'apposition et autres sont assimilés à des [→Mod](#) regroupés avec le sujet. Voir [→Mod](#), [→ModA](#) et [→ModD](#)

- (3) E ! Dix, fait Aucasins, ci fu Nicolette me douce amie, (AucR3 p. 27) [\[Afficher l'analyse\]](#)

Documentation

Plan de chaque fiche

Plan de la fiche

Remarques grammaticales

- \$1. Sujet nul
 - \$2. "Apposition", "épithète détachée", etc.
 - \$3. Sujet constitué uniquement d'un modifier
- #### Remarques lexicales
- \$4. Sujet de "avenir"
 - \$5. Tout, trestuit...

Liste des fiches

[Actant.xml.html](#)
[Apt.xml.html](#)
[AtObj.xml.html](#)
[AtRf.xml.html](#)
[AtSj.xml.html](#)
[At.xml.html](#)
[AuxA.xml.html](#)
[AuxP.xml.html](#)
[Aux.xml.html](#)
[Circ.xml.html](#)
[Cmpl.xml.html](#)
[Conjoint.xml.html](#)
[Coo.xml.html](#)
[DiscoursDirect.xml.html](#)
[Duplicata.xml.html](#)
[Fonction.xml.html](#)
[GpCoo.xml.html](#)
[Groupe.xml.html](#)
[Ignorer.xml.html](#)
[Insr.xml.html](#)
[Intj.xml.html](#)
[Lacune.xml.html](#)
[Max.xml.html](#)
[ModA.xml.html](#)
[ModD.xml.html](#)

Sujet personnel (SjPer)

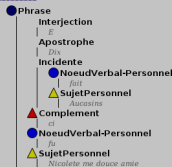
\$0.

Équivalent traditionnel

sujet, à l'exclusion du sujet dit "impersonnel" et du sujet dit "apparent"

Exemples-types

- (1) E ! Dix, fait Aucasins, ci fu Nicolette me douce amie, (AucR3 p. 27) [Masquer l'analyse](#)



- (2) A Deu, qui est sire du mont, | Cri ge merci, (TristBérM4 p. 67 v. 2186-2188) [\[Afficher l'analyse\]](#)

Remarques grammaticales

\$1. Sujet nul

Voir [→\\$1](#).

\$2. "Apposition", "épithète détachée", etc.

Les cas d'apposition et autres sont assimilés à des [→Mod](#) regroupés avec le sujet. Voir [→Mod](#), [→ModA](#) et [→ModD](#)

- (3) E ! Dix, fait Aucasins, ci fu Nicolette me douce amie, (AucR3 p. 27) [\[Afficher l'analyse\]](#)

Documentation

Détail de chaque exemple analysé (formats multiples)

Plan de la fiche

Remarques grammaticales

\$1. Sujet nul
\$2. "Apposition", "épithète détachée", etc.
\$3. Sujet constitué uniquement d'un modifieur
Remarques lexicales

\$4. Sujet de "avenir"
\$5. Tout, trestout...

Liste des fiches

- Actant.xml.html
- Apet.xml.html
- AtObj.xml.html
- AtRf.xml.html
- AtSj.xml.html
- At.xml.html
- AuxA.xml.html
- AuxP.xml.html
- Aux_xml.html
- Circ.xml.html
- Cmpl.xml.html
- Conjoint.xml.html
- Coc.xml.html
- DiscoursDirect.xml.html
- Duplicata.xml.html
- Fonction.xml.html
- GpCoc.xml.html
- Groupe.xml.html
- Ignorer.xml.html
- Instr.xml.html
- Intj.xml.html
- Lacune.xml.html
- Max.xml.html
- ModA.xml.html
- ModD.xml.html

Sujet personnel (SjPer)

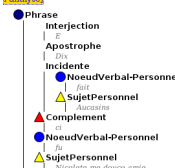
§0.

Équivalent traditionnel

sujet, à l'exclusion du sujet dit "impersonnel" et du sujet dit "apparent"

Exemples-types

(1) E | Dix, fait Aucasins, ci fu Nicolette me douce amie, (AucR3 p. 27) [Masquer l'analyse](#)



(2) A Deu, qui est sire du mont, | Cri ge merci, (TristBérM4 p. 67 v. 2185-2186) [\[Afficher l'analyse\]](#)

Remarques grammaticales

§1. Sujet nul

Voir →§1.

§2. "Apposition", "épithète détachée", etc.

Les cas d'apposition et autres sont assimilés à des →Mod regroupés avec le sujet.
Voir →Mod, →ModA et →ModD

(3) E | Dix, fait Aucasins, ci fu Nicolette me douce amie, (AucR3 p. 27) [\[Afficher l'analyse\]](#)

Prospectif : CoSToMeF

Computational Syntactic Tools for Medieval French

Prospectif : CoSToMeF

Computational Syntactic Tools for Medieval French

Projet soumis...

- ▶ Projet ANR/DFG soumis, appel 2011 (2012-2015)
- ▶ Serge Heiden (ENS Lyon) et Achim Stein (ILR Stuttgart)

Prospectif : CoSToMeF

Computational Syntactic Tools for Medieval French

Projet soumis...

- ▶ Projet ANR/DFG soumis, appel 2011 (2012-2015)
- ▶ Serge Heiden (ENS Lyon) et Achim Stein (ILR Stuttgart)

Application d'outils TAL aux corpus afr.

Prospectif : CoSToMeF

Computational Syntactic Tools for Medieval French

Projet soumis...

- ▶ Projet ANR/DFG soumis, appel 2011 (2012-2015)
- ▶ Serge Heiden (ENS Lyon) et Achim Stein (ILR Stuttgart)

Application d'outils TAL aux corpus afr.

- ▶ Outils
 - ▶ Parseurs, intégration à TXM, etc.
 - ▶ Interface web

Prospectif : CoSToMeF

Computational Syntactic Tools for Medieval French

Projet soumis...

- ▶ Projet ANR/DFG soumis, appel 2011 (2012-2015)
- ▶ Serge Heiden (ENS Lyon) et Achim Stein (ILR Stuttgart)

Application d'outils TAL aux corpus afr.

- ▶ Outils
 - ▶ Parseurs, intégration à TXM, etc.
 - ▶ Interface web
- ▶ Utilisation
 - ▶ « gold standard » ⇒ entraînement des parseurs
 - ▶ Analyses quantitatives et qualitatives
 - ▶ Description linguistique/évaluation des modèles

Prospectif : CoSToMeF

Computational Syntactic Tools for Medieval French

Projet soumis...

- ▶ Projet ANR/DFG soumis, appel 2011 (2012-2015)
- ▶ Serge Heiden (ENS Lyon) et Achim Stein (ILR Stuttgart)

Application d'outils TAL aux corpus afr.

- ▶ Outils
 - ▶ Parseurs, intégration à TXM, etc.
 - ▶ Interface web
- ▶ Utilisation
 - ▶ « gold standard » ⇒ entraînement des parseurs
 - ▶ Analyses quantitatives et qualitatives
 - ▶ Description linguistique/évaluation des modèles
- ▶ Enrichissement
 - ▶ Par traitement manuel
 - ▶ Par traitement (semi-)automatique

2^e partie : Logiciel *NotaBene* dans SRCMF

Présentation rapide

Philosophie générale

Annotation syntaxique (SRCMF)

Perspectives : intégration TXM (CoSToMeF)

Présentation rapide

Origine et développement

Coup d'œil sur l'interface

Philosophie générale

Annotations multiples – Vues multiples

Unités et modèle de données

Annotation syntaxique (SRCMF)

Vue arborescente

Annotations

Comparaisons

Perspectives : intégration TXM (CoSToMeF)

Présentation rapide



Philosophie générale



Annotation syntaxique (SRCMF)



Perspectives

Origine et développement

Origine et développement

Recherche doctorale : 2001-2007

- ▶ Recherche sur les relations entre ponctuation médiévale et syntaxe en afr. (corpus de chartes originales liégeoises en français)
- ▶ Besoin d'une **annotation exhaustive** des relations syntaxiques
- ▶ Logiciel Perl dédié



Origine et développement

Recherche doctorale : 2001-2007

- ▶ Recherche sur les relations entre ponctuation médiévale et syntaxe en afr. (corpus de chartes originales liégeoises en français)
- ▶ Besoin d'une **annotation exhaustive** des relations syntaxiques
- ▶ Logiciel Perl dédié

Postdoctorat ENS-Lyon (équipe BFM) : 2008

- ▶ Besoin d'ajouter/corriger des annotations
 - ▶ Création d'un prototype Python
 - ▶ Développement des capacités d'annotation de la syntaxe (absent BFM)
- ⇒ **Naissance de SRCMF**



Origine et développement

Recherche doctorale : 2001-2007

- ▶ Recherche sur les relations entre ponctuation médiévale et syntaxe en afr. (corpus de chartes originales liégeoises en français)
- ▶ Besoin d'une **annotation exhaustive** des relations syntaxiques
- ▶ Logiciel Perl dédié

Postdoctorat ENS-Lyon (équipe BFM) : 2008

- ▶ Besoin d'ajouter/corriger des annotations
 - ▶ Création d'un prototype Python
 - ▶ Développement des capacités d'annotation de la syntaxe (absent BFM)
- ⇒ **Naissance de SRCMF**

SRCMF : 2009-2012

Cf. supra : développement de l'outil en fonction des spécificités du projet

Coup d'œil sur l'interface

Trois zones principales

Projets – Annotation – Terminologie

The screenshot displays the SRM (SRM) software interface, which is used for managing linguistic projects, performing annotations, and managing terminology. The interface is divided into several main panels:

- Left Panel (Projets):** A tree view showing a list of projects and datasets. The 'Sophie - 1' project is selected, and its datasets (beroul 01/17 to 17/17) are listed with their status (e.g., 'Open RW', 'Closed').
- Top Panel (Tree view - Sophie - 1/beroul 01/17):** A hierarchical tree view of the annotation structure. It shows nodes like 'srcmf:hierarchy', 'srcmf:Phrase', 'srcmf:SujetPersonnel', and 'srcmf:NoeudVerbal-Personnel'. The tree is expanded to show sub-nodes like 'pense', 'srcmf:Objet', and 'srcmf:Circonstant'.
- Bottom Panel (Text view - Sophie - 1/beroul 01/17):** A text editor showing the annotated text. The text is:


```

      20] rois pense que par folie ,
      21] sire Tristan , vos aie amé ,
      22] Mais Dex plevis ma loiauté ,
      23] Qui sor mon cors mete flaele ,
      24] S' onques fors cil qui m' ot pucele
      25] Out m' amistié encor nul jor ! [p. 2]
      26] Se li felon de cest' enor ,
      27] Por qui jadis vos combatistes
      
```
- Right Panel (Panneau de contrôle):** A control panel for terminology. It shows a list of terms (e.g., 'srcmf:AttributSujet', 'srcmf:Regime', 'srcmf:Complément*', 'srcmf:Objet', 'srcmf:Reflechi', 'srcmf:ReflexifRenforce', 'srcmf:Sujet', 'srcmf:SujetImpersonnel', 'srcmf:SujetPersonnel', 'srcmf:Circonstant', 'srcmf:Forclusif', 'srcmf:Negation', 'srcmf:Parenthese', 'srcmf:Apostrophe', 'srcmf:Incident', 'srcmf:interjection', 'srcmf:Innocer').

Coup d'œil sur l'interface

Projets

Organisation des fichiers par utilisateur

The screenshot displays the SRCMF software interface, which is used for syntactic annotation. The interface is divided into several main sections:

- Projets (Projects):** A sidebar on the left shows a list of projects organized by user. Under 'Sophie - 1', there are two datasets: 'beroul 01/17' (Open RW) and 'beroul 02/17' (Closed). Other users listed include Alexei, Bea, Chefs, Julie, and PMor.
- Tree view - Sophie - 1/beroul 01/17:** The central pane shows a hierarchical tree structure of the text. The root node is 'n srcmf:hierarchy'. It branches into 'mie' and 'srcmf:Phrase'. 'srcmf:Phrase' further branches into 'li' and 'srcmf:SujetPersonnel'. 'li' branches into 'rois' and 'pense'. 'rois' branches into 'srcmf:NoeudVerbal-Personnel' and 'srcmf:Objet'. 'pense' branches into 'srcmf:Objet'. 'srcmf:Objet' branches into 'que' and 'srcmf:Circonstant'. 'que' branches into 'srcmf:Relateur-NonCoordonnant'. 'srcmf:Circonstant' branches into 'srcmf:Relateur-NonCoordonnant'.
- Text view - Sophie - 1/beroul 01/17:** The bottom pane shows the corresponding text with line numbers 20 to 27. The text is:


```

20|  rois pense que par folie ,
21|  Sire Tristran , vos aie amé ;
22|  Mais Dex plevis ma loiauté ,
23|  Qui sor mon cors mete flaele ,
24|  S' onques fors cil qui m' ot pucele
25|  Out m' amistié encor nul jor ! [p. 2]
26|  Se li felon de cest' enor ,
27|  Por qui jadis vos combatistes
      
```
- Panneau de contrôle (Control Panel):** On the right, there is a control panel with tabs for 'Terminologie' and 'Configuration de l'extension'. Below it, there is a section for 'Classes par portée' (Classes by scope) with a dropdown menu set to 'Literals'. A search box labeled 'Suj' is present. Below the search box, a list of classes is shown, including 'srcmf:AttributSujet', 'srcmf:Regime', 'srcmf:'Complément*', 'srcmf:Objet', 'srcmf:Reflechi', 'srcmf:ReflexifRenforce', 'srcmf:Sujet', 'srcmf:SujetImpersonnel', and 'srcmf:SujetPersonnel'. The 'srcmf:Sujet' class is currently selected.

Coup d'œil sur l'interface

Zone d'annotation

Différentes vues (ici texte/arbre) pour poser des étiquettes

The screenshot displays the SRCMF software interface, which is used for syntactic annotation. It is divided into several panels:

- Left Panel (Projets):** A list of projects and datasets, including 'beroul 01/17' which is currently open.
- Top Panel (Options):** Contains search and annotation tools like 'Search', 'Search Annotations', and 'Search Words'.
- Center Panel (Tree view):** Shows a hierarchical tree structure of the sentence. The root node is 'n' (srcmf:hierarchy). It branches into 'mie', 'li', 'rois', 'pense', 'que', and 'par folie'. Further sub-nodes include 'srcmf:SujetPersonnel', 'srcmf:NoeudVerbal-Personnel', 'srcmf:Objet', 'srcmf:Relateur-NonCoordonnant', and 'srcmf:Circonstant'.
- Bottom Panel (Text view):** Shows the original text of the sentence with corresponding annotations:


```

      20|  rois pense que par folie ;
      21|  Sire Tristan , vos aie amé ;
      22|  Mais Dex plevis ma loiauté ,
      23|  Qui sor mon cors mete flaele ,
      24|  S' onques fors cil qui m' ot pucele
      25|  Out m' amistié encor nul jor ! [p. 2]
      26|  Se li felon de cest' enor ,
      27|  Por qui jadis vos combatistes
      
```
- Right Panel (Panneau de contrôle):** Contains a 'Terminologie' section and a 'Classes par portée' section. The 'Classes par portée' section shows a list of classes for the selected node, such as 'srcmf:AttributSujet', 'srcmf:Regime', 'srcmf:Complément*', 'srcmf:Objet', 'srcmf:Reflechi', 'srcmf:ReflexifRenforce', 'srcmf:Sujet', 'srcmf:SujetImpersonnel', and 'srcmf:SujetPersonnel'.

Coup d'œil sur l'interface

Zone de terminologie

Hierarchie d'étiquettes

The screenshot displays the SRM (SRM) software interface, which is used for linguistic annotation. The interface is divided into several main sections:

- Left Panel (Projets):** A list of projects and datasets. The 'Sophie - 1' project is selected, showing a list of datasets from 'beroul 01/17' to 'beroul 17/17'. The 'beroul 01/17' dataset is currently open.
- Tree view (Top Center):** A hierarchical tree structure showing the organization of the text. The root is 'n srcmf:hierarchy', which branches into 'mie' and 'srcmf:Phrase'. Under 'srcmf:Phrase', there is a sub-tree for 'srcmf:SujetPersonnel', which further branches into 'rois', 'pense', and 'que'. 'pense' branches into 'srcmf:NoeudVerbal-Personnel' and 'srcmf:Objet'. 'srcmf:Objet' branches into 'srcmf:Relateur-NonCoordonnant' and 'que'. 'que' branches into 'srcmf:Circonstant' and 'srcmf:Relateur-NonCoordonnant'.
- Text view (Bottom Center):** A window showing the text being annotated. The text is:


```

20] | rois pense que par folie ,
21]   Sire Tristan , vos aie amé ,
22]   Mais Dex plevis ma loiauté ,
23]   Qui sor mon cors mete flaele ,
24]   S' onques fors cil qui m' ot pucele
25]   Out m' amistié encor nul jor ! [p. 2]
26]   Se li felon de cest' enor ,
27]   Por qui jadis vos combatistes
      
```
- Right Panel (Panneau de contrôle):** A control panel for the terminology. It shows a 'Terminologie' section with a 'Configuration de l'extension' dropdown set to 'Tree'. Below this, there is a 'Classes par portée' section with a dropdown set to 'Literals'. The main part of the panel shows a search for 'Suj' and a list of classes:
 - 17 srcmf:AttributSujet
 - 18 srcmf:Regime
 - 19 srcmf:'Complément'
 - 20 srcmf:Objet
 - 21 srcmf:Reflechi
 - 22 srcmf:ReflexifRenforce
 - 23 srcmf:Sujet
 - 24 srcmf:SujetImpersonnel
 - 25 srcmf:SujetPersonnel
 - 26 srcmf:Circonstant
 - 27 srcmf:Forclusif
 - 28 srcmf:Negation
 - 29 srcmf:Parenthese
 - 30 srcmf:Apostrophe
 - 31 srcmf:Incidente
 - 32 srcmf:interjection
 - 33 srcmf:Innocer

Présentation rapide

Origine et développement

Coup d'œil sur l'interface

Philosophie générale

Annotations multiples – Vues multiples

Unités et modèle de données

Annotation syntaxique (SRCMF)

Vue arborescente

Annotations

Comparaisons

Perspectives : intégration TXM (CoSToMeF)

Annotations multiples – Vues multiples

Principes généraux

- ▶ Multiplicité des schémas d'annotation
- ▶ Multiplicité des intérêts

Annotations multiples – Vues multiples

Principes généraux

- ▶ Multiplicité des schémas d'annotation
- ▶ Multiplicité des intérêts

⇒ Concrètement

Annotations multiples – Vues multiples

Principes généraux

- ▶ Multiplicité des schémas d'annotation
- ▶ Multiplicité des intérêts

⇒ Concrètement

1. Annotations de tous types :
 - ▶ morphologiques
 - ▶ syntaxiques
 - ▶ énonciatives

Annotations multiples – Vues multiples

Principes généraux

- ▶ Multiplicité des schémas d'annotation
- ▶ Multiplicité des intérêts

⇒ Concrètement

1. Annotations de tous types :
 - ▶ morphologiques
 - ▶ syntaxiques
 - ▶ énonciatives
2. Théories différentes :
 - ▶ dépendance
 - ▶ constituants

Annotations multiples – Vues multiples

Principes généraux

- ▶ Multiplicité des schémas d'annotation
- ▶ Multiplicité des intérêts

⇒ Concrètement

1. Annotations de tous types :
 - ▶ morphologiques
 - ▶ syntaxiques
 - ▶ énonciatives
2. Théories différentes :
 - ▶ dépendance
 - ▶ constituants

Besoins

- ▶ Flexibilité
- ▶ Réduction du bruit

Annotations multiples – Vues multiples

Dans SRCMF : synchronisation de visualisations multiples du même extrait ;
ici :

Annotations multiples – Vues multiples

Dans SRCMF : synchronisation de visualisations multiples du même extrait ;
ici :

- ▶ Arbre
- ▶ Concordance
- ▶ Texte courant

Annotations multiples – Vues multiples

Dans SRCMF : synchronisation de visualisations multiples du même extrait ;
ici :

- ▶ Arbre
- ▶ Concordance
- ▶ Texte courant

The screenshot displays the SRCMF interface with three main components:

- Tree View (Left):** A hierarchical tree structure showing annotations. A red circle highlights a node labeled 'ab' at line 720.
- Grid View (Top Right):** A table with four columns: 'nature', 'Contexte gauche', 'Pivot', and 'Contexte droit'. A red circle highlights the 'ab' pivot cell in the row corresponding to line 720.
- Text View (Bottom):** A text editor showing the source text. A red circle highlights the word 'ab' in the line 'gran compannie dels Judeus'.

Red arrows indicate the synchronization between these views: one arrow points from the 'ab' node in the tree to the 'ab' pivot in the grid, and another points from the 'ab' in the grid to the 'ab' in the text view.

nature	Contexte gauche	Pivot	Contexte droit
	, ab les femnes pres a	a	parler . * Audez ,
	. " Cum el perveng a	a	Golgota , davan la porta
	en la cruz , dos	a	sos laz penden lasruns :
	i diz . El mor a	a	tort . ren non fornsist
	non fornsist ; mais non	a	dreit per colpas granz esmes
	fel tuhas...	ab	gran compannie dels Judeus .
	li plus redre garde	ab	les femnes pres a parler
	quar il lo fel...	ab	vin , nostrae senior lo
	en cest di	ab	me venras in paradís .
	ortment lo vant il	acusand	, la soa mort mult
	soned , judeu f	acusent	, el se tais :
	per mal , P aurelia	ad	serv semper saned . Liadens
	el se tais :	ad	un responde non denat

Text view - nicolas - 1/passion_01

```

142) e anc orar cum ei aneed ,
126) s'ort sudor dunques suded ,
127) des cum lo sags a terra current
128) sa sudor las sanctas gutes . 104]
129) sos fidels cum repades
130) m benlemle los s'orted .
131) ludes ias aproismed
132) gran compannie dels Judeus .
  
```

Unités et modèle de données

Unités et modèle de données

Utilisation de standards du web sémantique

- ▶ XML (web « syntaxique »), identification des individus à annoter.

Unités et modèle de données

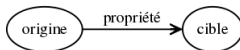
Utilisation de standards du web sémantique

- ▶ XML (web « syntaxique »), identification des individus à annoter.
- ▶ *Resource Description Framework* (RDF), annotations sous forme de graphes (ensembles de triplets) :

Unités et modèle de données

Utilisation de standards du web sémantique

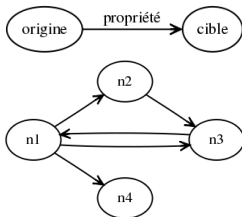
- ▶ XML (web « syntaxique »), identification des individus à annoter.
- ▶ *Resource Description Framework* (RDF), annotations sous forme de graphes (ensembles de triplets) :



Unités et modèle de données

Utilisation de standards du web sémantique

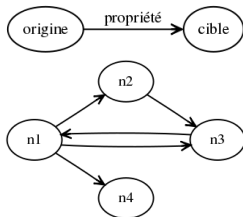
- ▶ XML (web « syntaxique »), identification des individus à annoter.
- ▶ *Resource Description Framework* (RDF), annotations sous forme de graphes (ensembles de triplets) :



Unités et modèle de données

Utilisation de standards du web sémantique

- ▶ XML (web « syntaxique »), identification des individus à annoter.
- ▶ *Resource Description Framework* (RDF), annotations sous forme de graphes (ensembles de triplets) :



- ▶ *OWL Ontology Web Language* (OWL) : hiérarchie de classes basé sur RDF, permet le raisonnement logique.

Unités et modèle de données

Application

Unités et modèle de données

Application

- ▶ **Contrainte générale** : si un seul constituant est présent dans un triplet, c'est l'origine qui pointe vers lui.

Unités et modèle de données

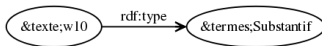
Application

- ▶ Contrainte générale : si un seul constituant est présent dans un triplet, c'est l'origine qui pointe vers lui.
- ▶ Trois relations fondamentales :

Unités et modèle de données

Application

- ▶ Contrainte générale : si un seul constituant est présent dans un triplet, c'est l'origine qui pointe vers lui.
- ▶ Trois relations fondamentales :
 1. Catégorisation



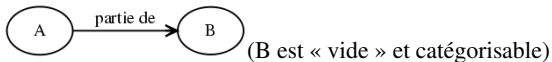
Unités et modèle de données

Application

- ▶ Contrainte générale : si un seul constituant est présent dans un triplet, c'est l'origine qui pointe vers lui.
- ▶ Trois relations fondamentales :
 1. Catégorisation



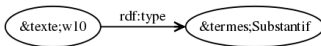
2. Groupement/délimitation



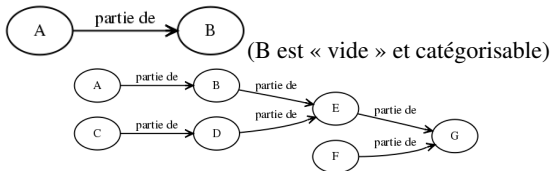
Unités et modèle de données

Application

- ▶ Contrainte générale : si un seul constituant est présent dans un triplet, c'est l'origine qui pointe vers lui.
- ▶ Trois relations fondamentales :
 1. Catégorisation



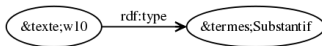
2. Groupement/délimitation



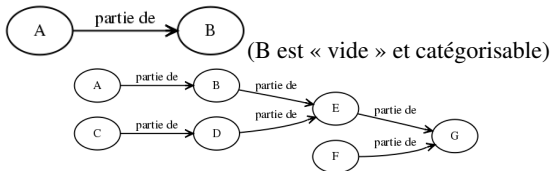
Unités et modèle de données

Application

- ▶ Contrainte générale : si un seul constituant est présent dans un triplet, c'est l'origine qui pointe vers lui.
- ▶ Trois relations fondamentales :
 1. Catégorisation



2. Groupement/délimitation



3. Liaison

Présentation rapide

Origine et développement

Coup d'œil sur l'interface

Philosophie générale

Annotations multiples – Vues multiples

Unités et modèle de données

Annotation syntaxique (SRCMF)

Vue arborescente

Annotations

Comparaisons

Perspectives : intégration TXM (CoSToMeF)

Vue arborescente

Vue arborescente

Objectif : ergonomie

- ▶ Nombreuses annotations à poser rapidement
- ▶ Présentation en vue hiérarchique arborescente
- ▶ **Vue de travail** (non théorique)

Vue arborescente

Objectif : ergonomie

- ▶ Nombreuses annotations à poser rapidement
- ▶ Présentation en vue hiérarchique arborescente
- ▶ **Vue de travail** (non théorique)
- ▶ Représentation des dépendances ; en deux mots :

Vue arborescente

Objectif : ergonomie

- ▶ Nombreuses annotations à poser rapidement
- ▶ Présentation en vue hiérarchique arborescente
- ▶ **Vue de travail** (non théorique)
- ▶ Représentation des dépendances ; en deux mots :
 - ▶ Les RONDS représentent des structures, les TRIANGLES spécifient leur dépendance

Vue arborescente

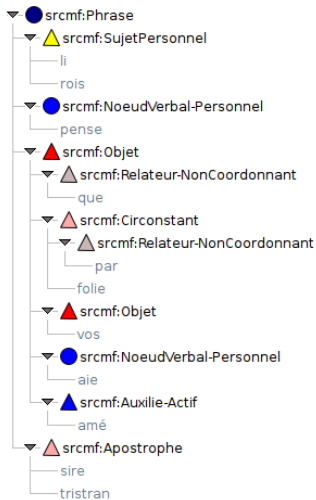
Objectif : ergonomie

- ▶ Nombreuses annotations à poser rapidement
- ▶ Présentation en vue hiérarchique arborescente
- ▶ **Vue de travail** (non théorique)
- ▶ Représentation des dépendances ; en deux mots :
 - ▶ Les RONDS représentent des structures, les TRIANGLES spécifient leur dépendance
 - ▶ À chaque niveau inférieur à la phrase, tous les nœuds dépendent du ROND

Vue arborescente

Objectif : ergonomie

- ▶ Nombreuses annotations à poser rapidement
- ▶ Présentation en vue hiérarchique arborescente
- ▶ **Vue de travail** (non théorique)
- ▶ Représentation des dépendances ; en deux mots :
 - ▶ Les RONDS représentent des structures, les TRIANGLES spécifient leur dépendance
 - ▶ À chaque niveau inférieur à la phrase, tous les nœuds dépendent du ROND
 - ▶ le ROND est le nœud principal (= tête) de la structure (sous-arbre)



Annotations

Démarche

Annotations

Démarche

1. Délimitation de la phrase

Annotations

Démarche

1. Délimitation de la phrase
2. Identification du nœud verbal qui fonde la phrase

Annotations

Démarche

1. Délimitation de la phrase
2. Identification du nœud verbal qui fonde la phrase
3. Identification, délimitation, caractérisation des dépendants du nœud verbal

Annotations

Démarche

1. Délimitation de la phrase
2. Identification du nœud verbal qui fonde la phrase
3. Identification, délimitation, caractérisation des dépendants du nœud verbal
4. Récursivement avec toutes les propositions

Annotations

Démarche

1. Délimitation de la phrase
2. Identification du nœud verbal qui fonde la phrase
3. Identification, délimitation, caractérisation des dépendants du nœud verbal
4. Récursivement avec toutes les propositions

Il s'agit d'une démarche théorique

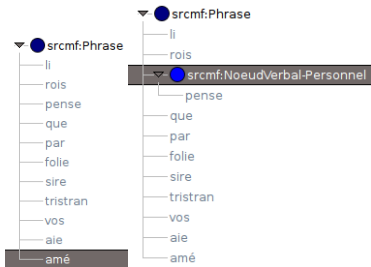
- ▶ On peut avancer par essai/erreur
- ▶ Un annotateur expérimenté peut brûler les étapes

Annotations



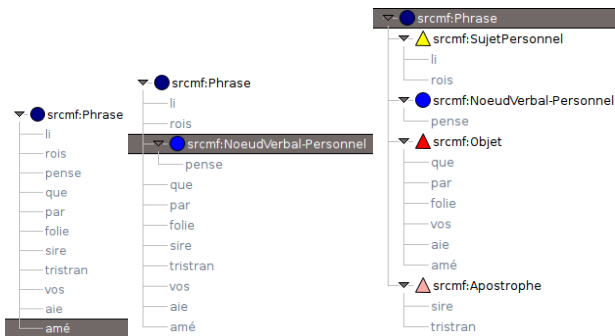
1. La phrase est délimitée

Annotations



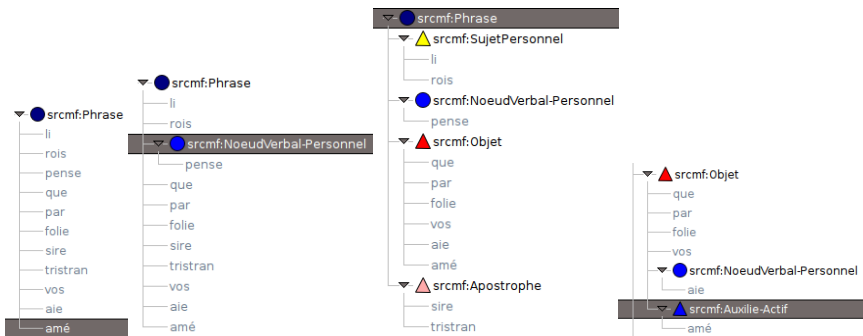
1. La phrase est délimitée
2. Le nœud verbal est identifié

Annotations

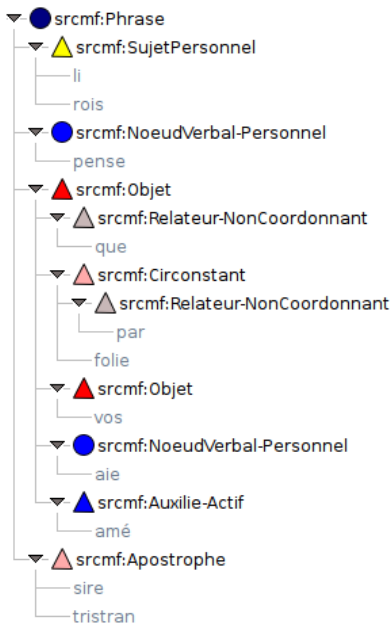


1. La phrase est délimitée
2. Le nœud verbal est identifié
3. Les dépendants du nœud verbal sont délimités et identifiés

Annotations



1. La phrase est délimitée
2. Le nœud verbal est identifié
3. Les dépendants du nœud verbal sont délimités et identifiés
4. ... et récursivement avec toutes les subordonnées

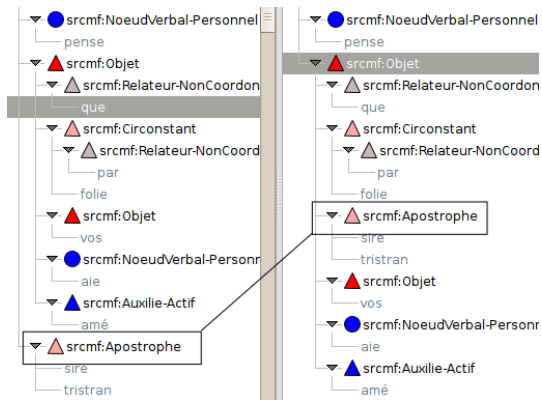


Comparaisons

Cf. étapes de la procédure

Comparaisons

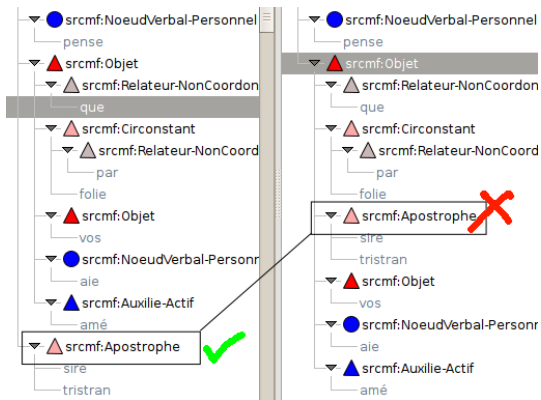
Cf. étapes de la procédure



- ▶ Comparaison d'analyses parallèles \Rightarrow divergences

Comparaisons

Cf. étapes de la procédure



- ▶ Comparaison d'analyses parallèles \Rightarrow divergences
- ▶ Choix de la bonne analyse

Présentation rapide

Origine et développement

Coup d'œil sur l'interface

Philosophie générale

Annotations multiples – Vues multiples

Unités et modèle de données

Annotation syntaxique (SRCMF)

Vue arborescente

Annotations

Comparaisons

Perspectives : intégration TXM (CoSToMeF)

Perspectives : intégration TXM (CoSToMeF)

Dialectique annotation/recherche

- ▶ Même environnement intégrant annotation et recherche
- ▶ Correction et enrichissement par de nouvelles questions

Perspectives : intégration TXM (CoSToMeF)

Dialectique annotation/recherche

- ▶ Même environnement intégrant annotation et recherche
- ▶ Correction et enrichissement par de nouvelles questions

Techniquement

- ▶ Développement d'un modèle de données commun décrivant la relation texte/annotation
- ▶ Réécriture Java Eclipse RPC

Merci

`nicolas.mazziotta@ulg.ac.be`