



Annotation manuelle ou semi-automatique de corpus avec Glozz

Yann Mathet, Antoine Widlöcher

Laboratoire GREYC - Université de Caen

*Journée d'étude « annotation de haut niveau »
ILF, Paris, le 16 janvier 2013*



1. Annotation de textes ?

- Nous entendons par annotation :
 - La création d'éléments placés sur le texte (blocs de texte)
 - Ou placés au-dessus d'autres éléments (relations entre ces éléments, structures les regroupant, etc.)
 - La caractérisation de chacun de ces éléments (type, structure de traits)
- L'annotation peut être :
 - Manuelle
 - Automatique
 - Semi-automatique

Exemple 1 : les chaînes de référence

Lorem **ipsum** dolor sit amet, consectetur adipiscing elit. Sed **non** risus.
 Suspendisse lectus tortor, dignissim sit **amet**, adipiscing nec, ultricies sed,
 dolor. Cras **elementum** ultrices diam. Maecenas ligula massa, varius a,
 semper congue euismod non, mi. Proin porttitor, orci **nec** nonummy
 molestie, enim est eleifend mi, **non** fermentum **diam** nisi sit amet erat. Duis
 semper. Duis arcu massa, scelerisque vitae, consequat in, pretium a, **enim**,
 Pellentesque congue. Ut in risus volutpat libero pharetra tempor. Cras
 vestibulum bibendum augue. **Praesent** egetas leo in pede. Praesent blandit
 odio eu enim, **Pellentesque** sed dui ut augue blandit sodales. Vestibulum
 ante **ipsum** primis in faucibus orci luctus et ultrices posuere cubilia Curae;
 Aliquam nibh. Mauris ac **mauris** sed pede pellentesque fermentum.



Exemple 2 : segmentation thématique, transitions

Dans cent ans, les pauvres historiens auront bien du mal à cerner notre époque ! Où donneront-ils de la tête ? Comment choisiront-ils entre cette multitude de documents _ imprimés, enregistrés, filmés, microfilmés _ que nous produisons à longueur de semaine ? Peut-être devrait-on déjà sélectionner, à leur intention, quelques oeuvres plus précieuses que d'autres, parce que plus précises, plus riches, ou simplement plus authentiques.

(ID=9)

Ces petits bijoux ne sont pas forcément des documents à grand tirage. Il faut parfois aller les chercher très loin du tintamarre médiatique. Le livre de Pierre Marthelot appartient à cette catégorie. Et s'il retient l'attention aujourd'hui, c'est sans doute parce qu'il tombe _ bien involontairement _ en pleine actualité politique.

Alors que la guerre scolaire montre à nouveau le bout de son nez, comment rester insensible à la chronique familiale et autobiographique de ce fils et petit-fils d'instituteurs, grandi dans le giron de l'école publique, nourri de morale laïque, mais qui se convertit inopinément au catholicisme à l'âge de quatorze ans et deviendra un jour président de la Paroisse universitaire ? A lui seul, le drame provoqué par cette conversion, dans une famille admirable où l'on ne triche pas, aurait suffi à faire un très bon livre. Mais, au-delà de sa personne, Pierre Marthelot a entrepris de raconter toute une " dynastie d'écolâtres ".



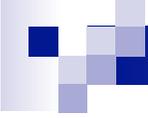
2. Contraintes

- Multi-annotateurs
- Multi-échelles
- Variété « topologique » (sporadicité, imbrication, recouvrement)
- Variété « structurelle » (unités, relations, schémas)
- Différentes vues, différents paradigmes de représentation des annotations
- Processus d'annotation et d'interrogation simultanés



3. Campagne d'annotation

- Détermination du corpus et des formats
- Définition des objets étudiés
- Définition du modèle d'annotation
- Proposition de « feuilles de style » pour la visualisation des annotations
- Procédures de recherche et de vérification prédéfinies
- Rédaction du manuel d'annotation
- Analyse de l'accord inter-annotateurs
- Collecte et diffusion des données (corpus et annotations)



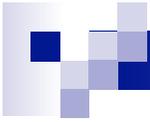
4. Le méta-modèle U-R-S et ses instanciations

- Méta-modèle générique
- Déclinable en une infinité de modèles spécifiques (relevant chacun d'une campagne donnée)
- Basé sur 3 catégories : Unités, Relations, Schémas
- Chaque catégorie peut donner lieu à la création de différents types (unité de type nom, de type adjectif, relation de type sujet, de type complément, etc.)
- Chaque type se voit associer un ensemble de traits prédéfinis à remplir ensuite individuellement pour chaque élément (genre = masculin ou féminin, nombre = singulier ou pluriel, etc.)
- Il est possible de créer des groupes permettant de regrouper différents types au sein d'une même « famille ». Par ex. un groupe syntaxe et un groupe sémantique.

4.1. Variété topologique et structurelle

Unités

lectus tortor, dignissim sit amet, adipiscing nec, ultricies sed, dolor. Cras elementum ultrices diam. **Maecenas** ligula massa, **varius a, semper** congue, euismod non, mi. *Proin porttitor, orci nec nonummy molestie, enim est eleifend mi, non fermentum diam nisl sit amet erat.* Duis **semper**. Duis arcu massa, scelerisque vitae, consequat in, pretium a, enim. **Pellentesque congue.** **Ut in risus** volutpat libero pharetra tempor. Cras vestibulum bibendum augue. Praesent egestas leo in pede. Praesent **blandit odio** eu enim. Pellentesque sed dui ut augue blandit sodales. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Aliquam nibh. Mauris ac mauris sed pede pellentesque fermentum. Maecenas **adipiscing ante non diam sodales hendrerit.** Ut velit mauris, egestas sed, gravida nec, ornare ut, mi. Aenean **ut orci vel massa suscipit pulvinar.** Nulla sollicitudin. Fusce **varius, ligula non** tempus aliquam, nunc turpis ullamcorper nibh, in tempus sapien eros vitae ligula. Pellentesque rhoncus nunc et augue. Integer id felis.



Relations

Lorem ipsum dolor sit amet, **consectetur** adipiscing elit. Sed non risus. Suspendisse lectus tortor, dignissim sit amet, adipiscing nec, ultricies sed, dolor. Cras elementum ultrices diam. Maecenas ligula massa, varius a, semper congue, euismod non, mi. Proin porttitor, orci nec **nonummy** molestie, enim **est eleifend mi, non fermentum** diam nisl sit

Schémas

Lorem ipsum dolor sit amet, **consectetur** adipiscing elit. Sed **non** risus. Suspendisse lectus tortor, dignissim sit amet, adipiscing nec, ultricies sed, dolor. Cras elementum ultrices diam. Maecenas ligula massa, varius a, semper congue, euismod non, **mi** Proin porttitor, orci nec **nonummy** molestie, enim **est eleifend mi, non fermentum** diam nisl sit amet erat. Duis semper. Duis arcu massa, scelerisque vitae, consequat in, pretium a, enim. Pellentesque congue. Ut in risus volutpat libero pharetra tempor. Cras vestibulum bibendum augue. **Praesent** egesta leo in **pede. Praesent** blandit odio eu enim. Pellentesque



4.2. Structures de traits

Structure de traits associée au type « nom » :

Feature name	Possible values
Catégorie	<i>pronom, nom propre, nom commun</i>
Nombre	<i>singulier, pluriel</i>

Instanciation pour l'unité annotée « Paul » :

Feature name	value
Catégorie	<i>nom propre</i>
Nombre	<i>singulier</i>

4.3. Instanciation d'un modèle

<annotationModel>

<units>

```
<type name="Nom">  
  <featureSet>  
    <feature name="Catégorie">  
      <possibleValues default = "">  
        <value>pronom</value>  
        <value>nom propre</value>  
        <value>nom commun</value>  
      </possibleValues>  
    </feature>  
    <feature name="Nombre">  
      <possibleValues default = "">  
        <value>singulier</value>  
        <value>pluriel</value>  
      </possibleValues>  
    </feature>  
  </featureSet>  
</type>
```

The screenshot shows a software interface with three panels: 'Units', 'Relations', and 'Schemas'. The 'Units' panel lists: Nom, Unité discursive, Sujet, Expression référentielle, Élément initial, and Verbe. The 'Relations' panel lists: Inclusion, Appartenance, Dépendance, and Tête. The 'Schemas' panel lists: Structure information and Coréférence. Below these panels is a table with 'Feature name' as a header. The table contains two rows: 'Catégorie' and 'Nombre'. A dropdown menu is open for the 'Nombre' feature, showing 'singulier' and 'pluriel' options. Red arrows point from the XML code on the left to the corresponding elements in the interface: from 'Nom' to the 'Nom' unit, from 'Catégorie' to the 'Catégorie' feature name, and from 'singulier' and 'pluriel' to the dropdown menu.

(...)

</units>

(...)

4.4. Principe des annotations déportées (stand-off)

es aventures d'un gascon désargenté de 18 ans
se lie d'amitié avec Athos Porthos et Aramis
tre hommes vont s'opposer au premier mi
ents, dont la belle et mystérieuse Milady de
de France, Anne d'Autriche. Avec ses nom
nesques, Les Trois mousquetaires est l'exem

ID=57
Unit
Type = expr. réf
Début = 191
Fin = 196

ID=118
Relation
Type = Appartenance
ID-Début = 57
ID-Fin = 62

ID=62
Unit
Type = expr. réf
Début = 517
Fin = 540

extrait xml : une unité

```
<unit id="ymathet_1315780133989">
  <metadata>
    <author>ymathet</author>
    <creation-date>1315780133989</creation-date>
    <lastModifier>n/a</lastModifier>
    <lastModificationDate>0</lastModificationDate>
  </metadata>
  <characterisation>
    <type>Expression référentielle</type>
    <featureSet>
      <feature name="Support"/>
      <feature name="Catégorie"/>
    </featureSet>
  </characterisation>
  <positioning>
    <start>
      <singlePosition index="191"/>
    </start>
    <end>
      <singlePosition index="196"/>
    </end>
  </positioning>
</unit>
```

ID=57
Unit
Type = expr. réf
Début = 191
Fin = 196



5. Visualisation

- La représentation exhaustive des annotations devient vite surchargée
- Il faut s'efforcer de ne montrer que l'information utile,
- et la représenter d'une façon adaptée à la tâche en cours.
- On peut notamment proposer :
 - différentes « feuilles de style »
 - la création de groupes de types
 - différentes stratégies de visualisation



5.1. Feuilles de style

- Différentes feuilles de style peuvent être créées pour un modèle d'annotation donné, comme autant de points de vue sur un même matériau.
- Couleur par type : une couleur peut être associée à chaque type. Eventuellement, une même couleur peut être utilisée pour différents types, ou un certain dégradé de couleurs pour un certain paradigme.
- Spécificité des schémas : 4 modes distincts
- Visibilité par type : chaque type peut être caché ou montré. On peut ainsi se focaliser sur une partie des données.

Couleurs par type : exemple

was a French painter, printmaker, draughtsman, and illustrator, whose immersion in the colourful and theatrical life of fin de siècle Paris yielded an œuvre of exciting, elegant and provocative images.

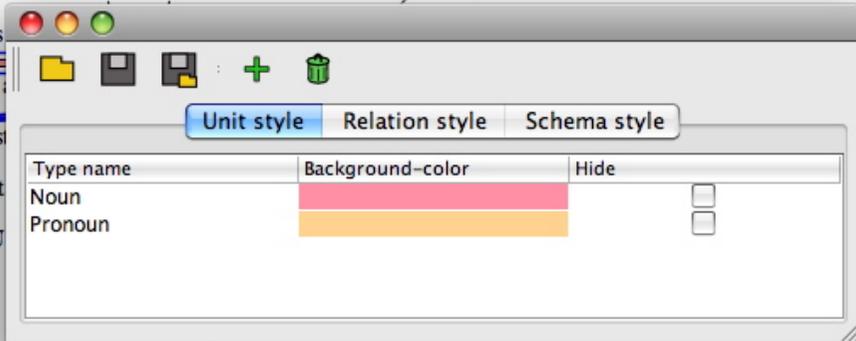
Lautrec is known as one of the Post-Impressionist painters of the Post-Impressionist movement.

A new record was set for the painting, which was sold for \$22.4 million USD.

Biography

Youth

Henri Marie Raymond de Toulouse-Lautrec-Montfa was born in Albi, Tarn in the Midi-Pyrénées région of France, the firstborn child of Comte Alphonse and Comtesse Adèle de Toulouse-Lautrec. He was therefore a member of an aristocratic family (descendants of the



Type name	Background-color	Hide
Noun	Red	<input type="checkbox"/>
Pronoun	Orange	<input type="checkbox"/>

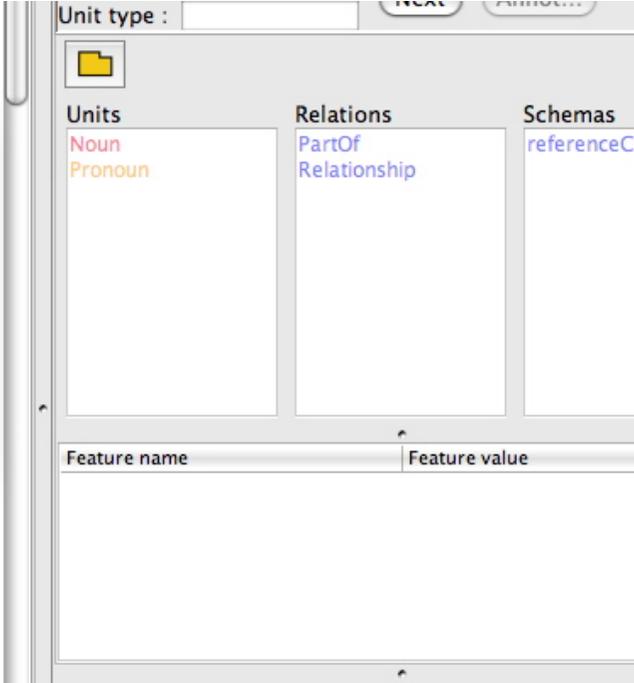
Unit type : [dropdown]

Units
Noun
Pronoun

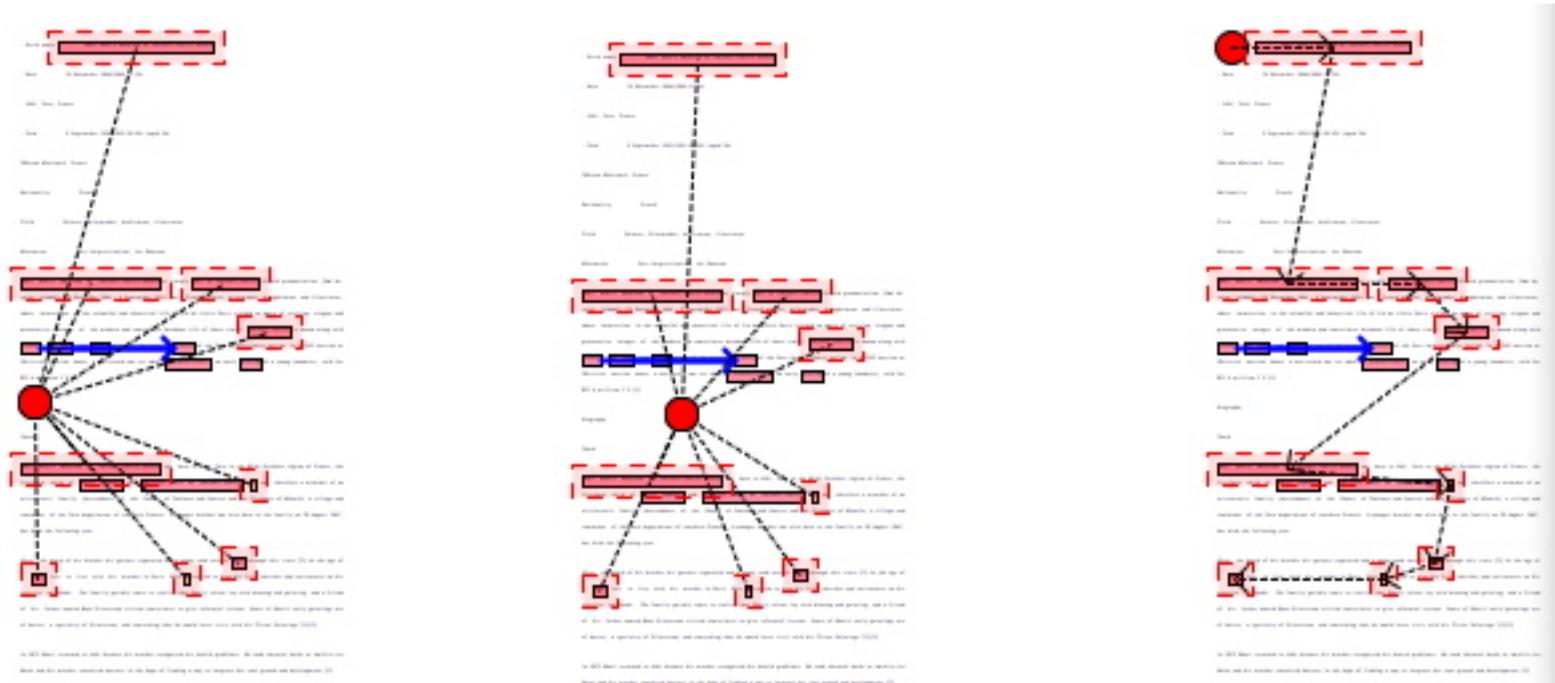
Relations
PartOf
Relationship

Schemas
referenceCh

Feature name | Feature value



Styles de schémas : 3 exemples





5.2. Visualisation par groupes

- Chaque groupe peut être montré ou caché, permettant de se focaliser sur certains paradigmes prévus par le modèle
- Ex : groupe syntaxe, groupe sémantique, etc.



5.3. Modes de visualisation

- Il est possible de déroger partiellement ou totalement aux feuilles de style par l'usage de modes particuliers
- Individuellement : possibilité d'associer une couleur particulière à un élément particulier, au cas par cas
- Possibilité de combiner ou non avec l'emploi d'une feuille de style
- Selon des particularités topologique : couleurs par sous-graphes connexes. Ex. : chaque chaîne de référence dispose d'une couleur distincte

Mode « co-référence »

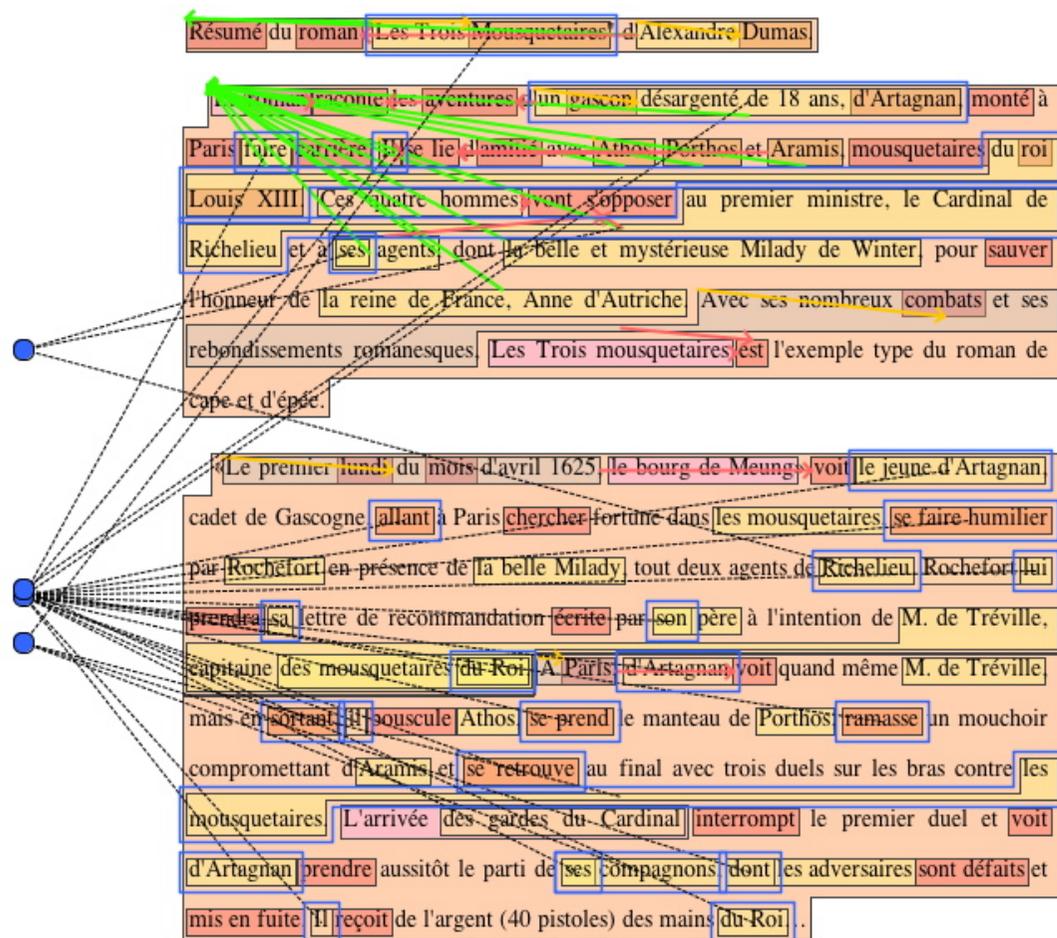
Lorem ipsum dolor sit amet, consectetur adipiscing elit. Sed non risus. Suspendisse lectus tortor, dignissim sit amet adipiscing nec, ultricies sed, dolor. Cras elementum ultrices diam. Maecenas ligula massa, varius a, semper congue euismod non, mi. Proin porttitor, orci nec nonummy molestie, enim est eleifend mi, non fermentum diam nisi sit amet erat. Duis semper. Duis arcu massa, scelerisque vitae, consequat in, pretium a, enim Pellentesque congue. Ut in risus volutpat libero pharetra tempor. Cras vestibulum bibendum augue. Praesent egestas leo in pede. Praesent blandit odio eu enim, Pellentesque sed dui ut augue blandit sodales. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Aliquam nibh. Mauris ac mauris sed pede pellentesque fermentum.

Sans

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Sed non risus. Suspendisse lectus tortor, dignissim sit amet adipiscing nec, ultricies sed, dolor. Cras elementum ultrices diam. Maecenas ligula massa, varius a, semper congue euismod non, mi. Proin porttitor, orci nec nonummy molestie, enim est eleifend mi, non fermentum diam nisi sit amet erat. Duis semper. Duis arcu massa, scelerisque vitae, consequat in, pretium a, enim Pellentesque congue. Ut in risus volutpat libero pharetra tempor. Cras vestibulum bibendum augue. Praesent egestas leo in pede. Praesent blandit odio eu enim, Pellentesque sed dui ut augue blandit sodales. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Aliquam nibh. Mauris ac mauris sed pede pellentesque fermentum.

Avec

Ajustement d'une vue (1)

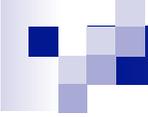


Ajustement d'une vue (2)

- Résumé du roman **Les Trois Mousquetaires** d'**Alexandre Dumas**
- Le roman raconte les aventures d'un **gascon désargenté** de 18 ans, d'**Artagnan** monté à Paris **faire carrière**. Il se lie d'amitié avec **Athos**, **Porthos** et **Aramis**, **mousquetaires du roi Louis XIII**. Ces quatre hommes vont s'opposer au premier ministre, le Cardinal de **Richelieu** et à **ses agents**, dont la belle et mystérieuse **Milady de Winter** pour sauver l'honneur de **la reine de France, Anne d'Autriche**. Avec ses nombreux combats et ses rebondissements romanesques, **Les Trois mousquetaires** est l'exemple type du roman de cape et d'épée.
- «Le premier lundi du mois d'avril 1625, le bourg de Meung» voit **le jeune d'Artagnan**, cadet de Gascogne **aller à Paris** chercher fortune dans **les mousquetaires** **se faire humilier** par **Rochefort** en présence de **la belle Milady** tout deux agents de **Richelieu**, **Rochefort** lui prendra **sa lettre de recommandation** écrite par **son père** à l'intention de **M. de Tréville**, capitaine **des mousquetaires du Roi**. A Paris, **d'Artagnan** voit quand même **M. de Tréville**, mais en **sortant** il **trouve** **Athos**, **se prend** le manteau de **Porthos**, **ramasse** un mouchoir compromettant d'**Aramis** et **se retrouve** au final avec trois duels sur les bras contre **les mousquetaires**. L'arrivée **des gardes du Cardinal** interrompt le premier duel et voit **d'Artagnan** prendre aussitôt le parti de **ses compagnons** dont **les adversaires** sont défaits et mis en fuite. Il **reçoit** de l'argent (40 pistoles) des mains **du Roi**.

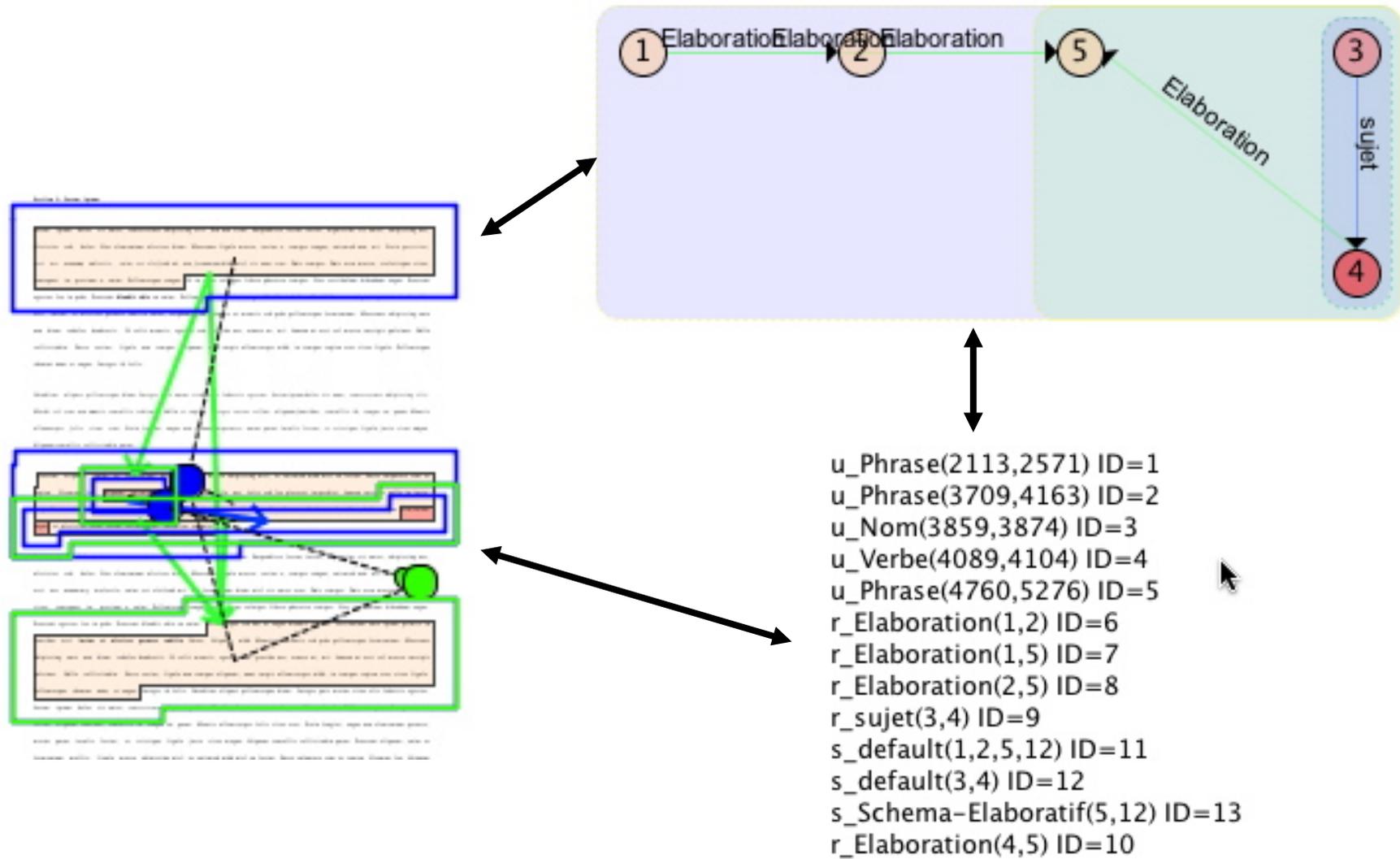
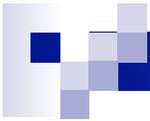
Ajustement d'une vue (3)

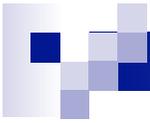
- Résumé du roman "Les Trois Mousquetaires" d'Alexandre Dumas.
- Le roman raconte les aventures d'un gascon désargenté de 18 ans, d'Artagnan, monté à Paris faire carrière. Il se lie d'amitié avec Athos, Porthos et Aramis, mousquetaires du roi Louis XIII. Ces quatre hommes vont s'opposer au premier ministre, le Cardinal de Richelieu et à ses agents, dont la belle et mystérieuse Milady de Winter, pour sauver l'honneur de la reine de France, Anne d'Autriche. Avec ses nombreux combats et ses rebondissements romanesques, Les Trois mousquetaires est l'exemple type du roman de cape et d'épée.
- «Le premier lundi du mois d'avril 1625, le bourg de Meung» voit le jeune d'Artagnan, cadet de Gascogne, aller à Paris chercher fortune dans les mousquetaires. se faire humilier par Rochefort en présence de la belle Milady, tout deux agents de Richelieu. Rochefort lui prendra sa lettre de recommandation écrite par son père à l'intention de M. de Tréville, capitaine des mousquetaires du Roi. A Paris, d'Artagnan voit quand même M. de Tréville, mais en sortant, il souscrite Athos, se prend le manteau de Porthos, ramasse un mouchoir compromettant d'Aramis et se retrouve au final avec trois duels sur les bras contre les mousquetaires. L'arrivée des gardes du Cardinal interrompt le premier duel et voit d'Artagnan prendre aussitôt le parti de ses compagnons, dont les adversaires sont défaits et mis en fuite. Il reçoit de l'argent (40 pistoles) des mains du Roi.



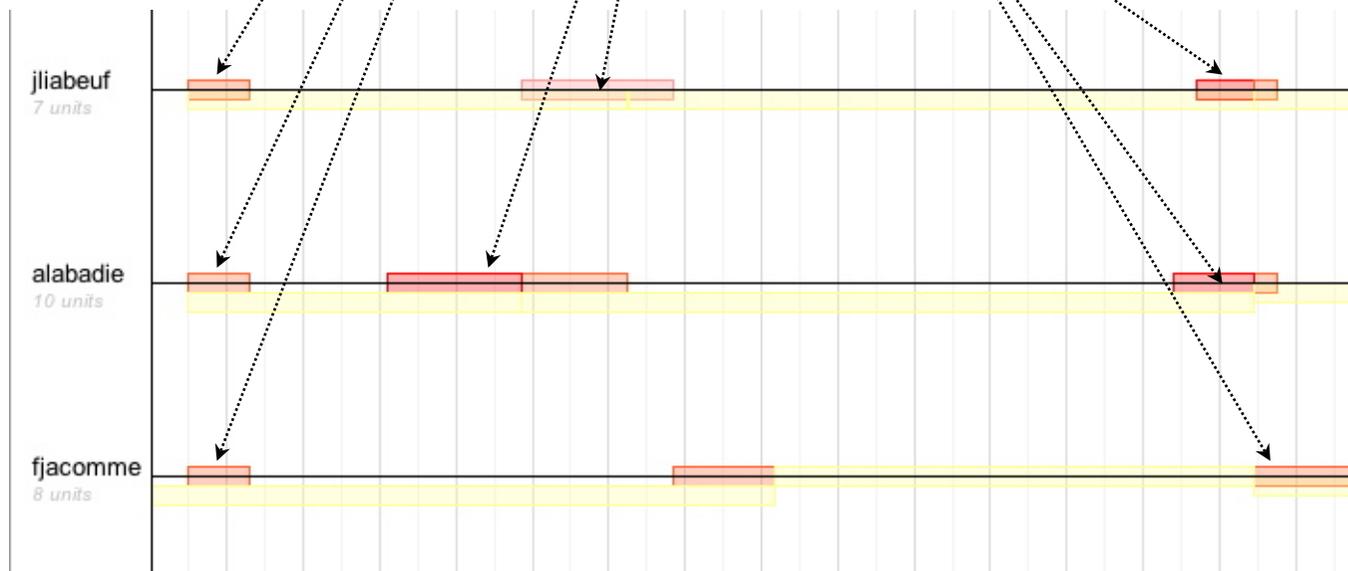
6. Multiplier les paradigmes de représentation

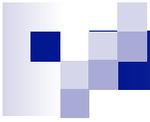
- Aux spécificités d'une tâche ou d'une sous-tâche d'annotation peut correspondre tel ou tel paradigme de représentation
- Représentation « sur le texte », pratique pour annoter les éléments de type Unité, mais ne permet pas toujours de bien appréhender les structures riches (schémas imbriqués, relations entre schémas, annotateurs multiples, etc.)
- Représentation sous forme de « prédicats »
- Représentation sous forme de graphes
- Représentation sous forme linéaire et multi-annotateurs



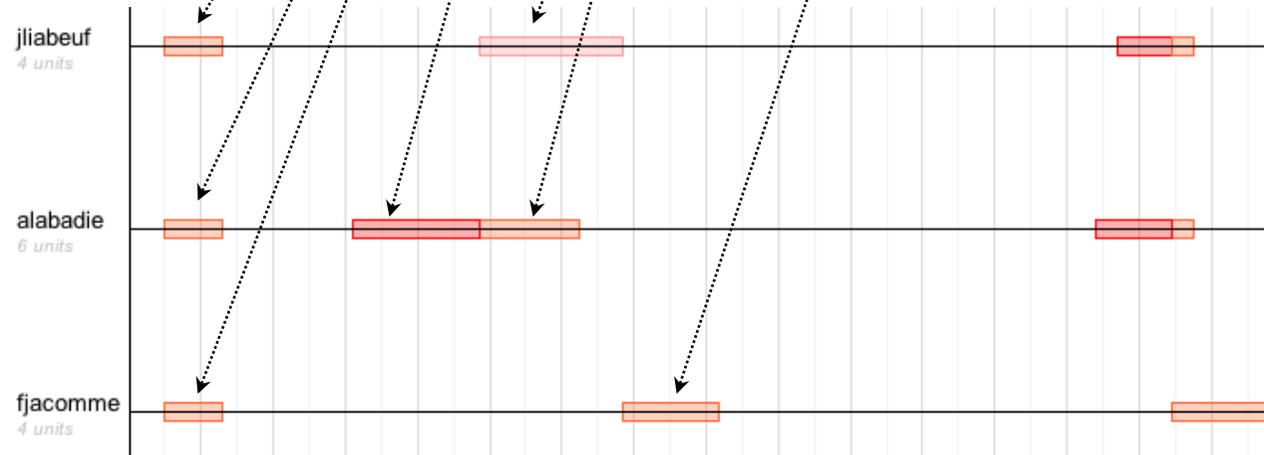


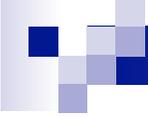
UN LIVRE> A bonne écoleUNE DYNASTIE D ECOLATRES Dans cent ans, les
pauvres historiens auront bien du mal à cerner notre époque ! Où donneront-ils de la tête ?
Comment choisiront-ils entre cette multitude de documents _ imprimés, enregistrés, filmés,
microfilmés _ que nous produisons à longueur de semaine ? Peut-être devrait-on déjà
sélectionner, à leur intention, quelques oeuvres plus précieuses que d'autres, parce que plus
précises, plus riches, ou simplement plus authentiques. Ces petits bijoux ne sont pas
forcément des documents à grand tirage. Il faut parfois aller les chercher très loin du
tintamarre médiatique. Le livre de Pierre Marthelot appartient à cette catégorie. Et s'il
retient l'attention aujourd'hui, c'est sans doute parce qu'il tombe _ bien involontairement _
en pleine actualité politique. Alors que la guerre scolaire montre à nouveau le bout de son
nez, comment rester insensible à la chronique familiale et autobiographique de ce fils et





UN LIVRE > A bonne école UNE DYNASTIE D ECOLATRES Dans cent ans, les
pauvres historiens auront bien du mal à cerner notre époque ! Où donneront-ils de la tête ?
Comment choisiront-ils entre cette multitude de documents _ imprimés, enregistrés, filmés,
microfilmés _ que nous produisons à longueur de semaine ? Peut-être devrait-on déjà
sélectionner, à leur intention, quelques oeuvres plus précieuses que d'autres, parce que plus
précises, plus riches, ou simplement plus authentiques. Ces petits bijoux ne sont pas
forcément des documents à grand tirage. Il faut parfois aller les chercher très loin du
tintamarre médiatique. Le livre de Pierre Marthelot appartient à cette catégorie. Et s'il
retient l'attention aujourd'hui, c'est sans doute parce qu'il tombe bien involontairement _
en pleine actualité politique. Alors que la guerre scolaire montre à nouveau le bout de son





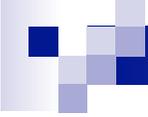
7. Outils d'interrogation et de vérification

- Processus d'annotation et d'interrogation sont souvent corrélés :
 - faire apparaître des structures à un niveau n pour bâtir des annotations au niveau $n+1$
 - Vérifier à différents moments que les annotations produites sont conformes au manuel d'annotation
- Intégration de GlozzQL au sein de Glozz, basé sur un formalisme de contraintes



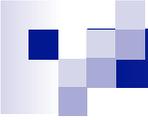
7.1. GlozzQL : présentation

- Création incrémentale de requêtes
- Requêtes variées, pouvant notamment porter sur :
 - l'auteur
 - Les traits
 - Les configurations topologiques
 - Combinaison de plusieurs contraintes
- Retour direct vers les résultats au sein même de l'interface d'annotation
- Panier permettant de collecter différents résultats, pour ensuite les enregistrer ou les supprimer des annotations en cours



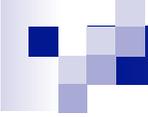
7.2 GlozzQL : quelques exemples de requêtes

- A. Recherche « full-text »
- B. Recherche par valeurs de traits (singulier vs pluriel, passé vs futur, etc.)
- C. Recherche par auteur (l'auteur pouvant être un processus automatique)
- D. Combinaison de deux requêtes (ex. : full-text + auteur)
- E. Recherches « profondes » au sein de relations et de schémas: vérification de l'absence d'un élément constitutif indispensable à une structure, ou réciproquement de la présence d'un élément interdit
- F. Recherche de relations orientées vers le haut ou vers le bas : détections de relations de type « cataphore » orientées vers le haut, ou « anaphore » orientées vers le bas



7.3 GlozzQL : utilisation du panier

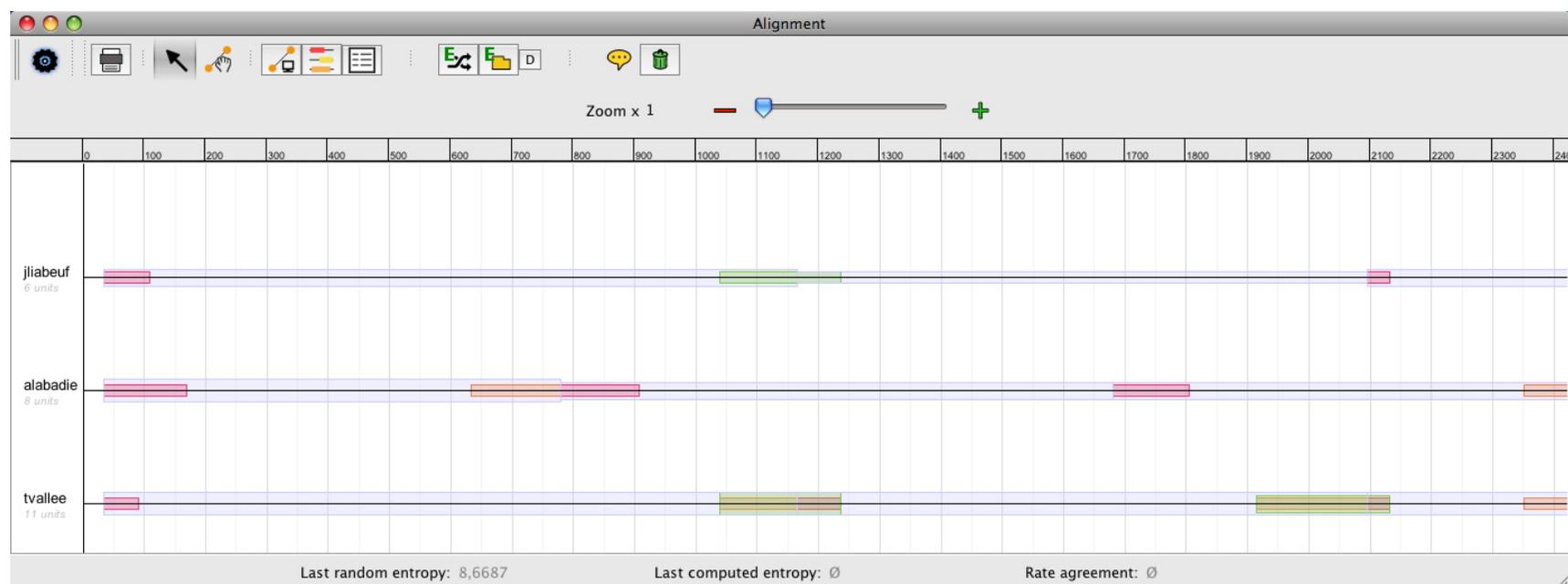
- 1. Sauvegarde des résultats d'une ou de plusieurs requêtes au sein d'un fichier spécifique. Ex. : regrouper toutes les annotations d'un auteur donné, toutes les annotations ayant trait au passé ou au futur, au singulier ou au pluriel, etc.
- 2. Retrait des résultats d'une ou de plusieurs requêtes des annotations actuellement utilisées. Ex. : filtrer des données
- À l'avenir :
 - appliquer automatiquement une modification sur l'ensemble des données du panier
 - créer automatiquement des annotations nouvelles sur la bases des données du panier (par application d'un patron).



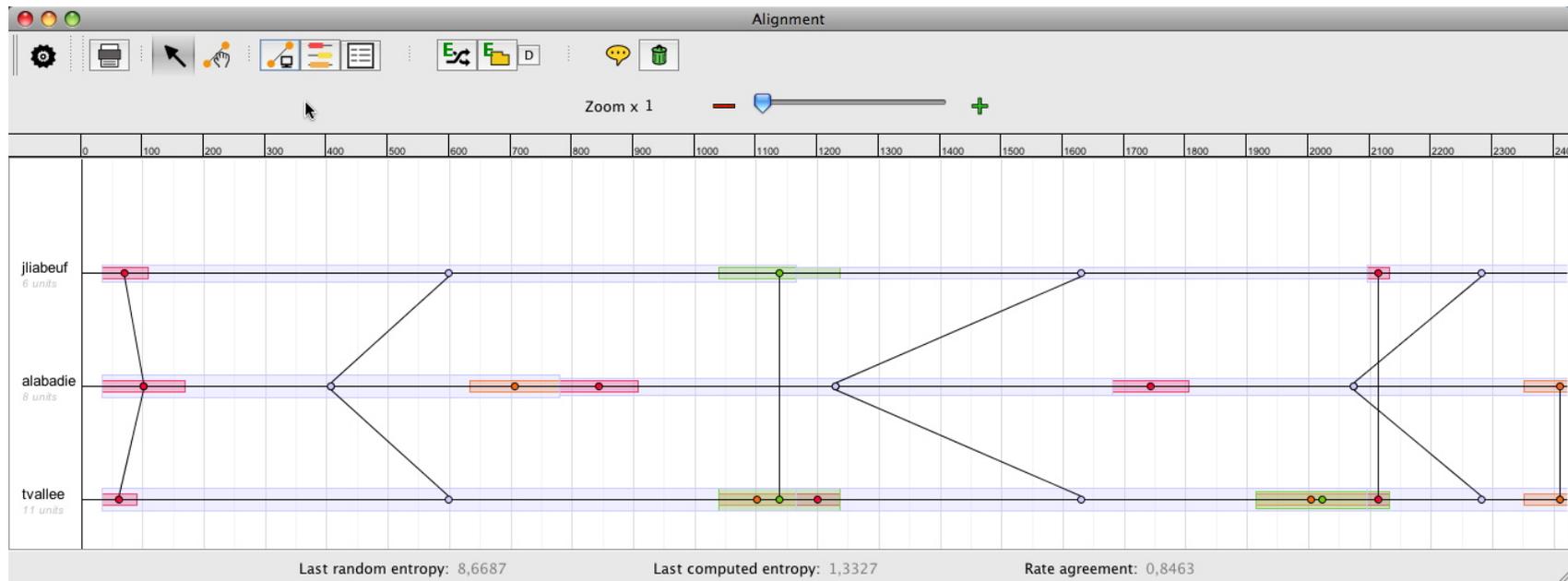
8. Alignement, Accord inter-annotateurs

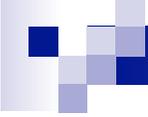
- Plusieurs annotateurs exécutent une même tâche d'annotation sur des textes identiques
- Les annotations ne se correspondent que partiellement...
- Comment mesurer l'accord ?
- Comment apparier (aligner) les productions ?
- Que signifie une valeur d'accord donnée ?
- Quid de ces questions dans le cadre d'une campagne d'annotation ?

Exemple : 3 annotateurs



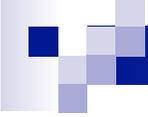
Alignement automatique associé :





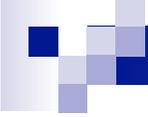
8.2 Exigences

- Généraliser alignement et mesure d'accord sur une grande variété de campagnes d'annotation implique :
 - prise en compte de structures complexes (recouvrement d'unités, emboîtement, etc.)
 - nombre d'annotateurs non limité à deux
 - pas d'annotation de référence (l'ensemble crée la référence)
 - prise en compte non binaire des différences de caractérisation (ex. : catégories proches = sanction faible) et de positionnement
 - prise en compte du hasard et des spécificités du corpus annoté
 - repérage des annotateurs « consensuels » et « non consensuels »
 - mise à disposition d'une « grille de lecture » explicite
- La méthode doit être capable de traiter alignement ET mesure d'accord, l'un ne pouvant aller sans l'autre



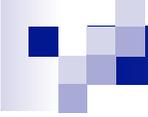
8.3. Utilisation dans une campagne d'annotation

- Mesurer d'accord pour :
 - Juger de la pertinence de la tâche
 - Vérifier la qualité du manuel d'annotation
 - Faire apparaître le plus tôt possible les divergences de certains annotateurs par rapport aux autres : le cas échéant, mieux les former ; s'il est trop tard, écarter certaines annotations
- Obtenir un alignement pour créer une annotation de référence :
 - la référence est créée consensuellement
 - Chaque annotation peut être pondérée individuellement par l'accord qui la concerne (annotations plus ou moins sûres)
- Inconvénient : coût en annotateurs --> s'adapter aux ressources disponibles...



8.4 Procédure possible

- On dispose par ex. de 5 annotateurs et 100 textes
- Chaque annotateur prend connaissance du manuel d'annotation
- On fait tout d'abord travailler les 5 annotateurs sur quelques textes.
- On mesure l'accord global. Trop faible : revoir la tâche ou le manuel
- On mesure ensuite le degré de consensualité de chaque annotateur :
 - pour un annotateur donné, sur chaque texte, on compare l'accord des 5 annotateurs avec l'accord des 4 annotateurs autres
 - Lorsque l'accord augmente en supprimant cet annotateur, il s'agit d'un annotateur non consensuel
 - À l'inverse, lorsqu'il diminue, c'est que cet annotateur est consensuel (la présence de ses annotations renforce l'ensemble)
- Lorsqu'un annotateur est trop peu consensuel, revoir avec lui en détail la procédure d'annotation.
- On peut alors distribuer les 100 textes de sorte qu'ils soient vus chacun par deux (ou trois) annotateurs, en faisant varier les couples.



Perspectives : Expériences psycho-linguistiques ?

- Chronologie de la création des annotations : différentes stratégies selon :
 - le type de corpus
 - l'annotateur
 - ...
- Vitesse de création des annotations. Peut dépendre :
 - de la phase d'annotation
 - de l'annotateur
 - ...