

# Études sur les chaînes de référence : approches computationnelles

Silvia Federzoni

Sous la direction de Cécile Fabre et Lydia-Mai Ho-Dac  
CLLE, CNRS & Université de Toulouse

**UE SLT0113T**

Thématiques actuelles de la recherche en TAL

6 décembre 2021



- 1 Objet d'étude : la continuité référentielle
- 2 Contexte et objectifs de la thèse
- 3 Données et méthode
- 4 Résultats
- 5 Discussion

# Objet d'étude : la continuité référentielle

Cette maison est hantée.

Elle habitait dans cette maison depuis longtemps.

> longtemps.

Il marcha marcha : Il se retourna en entendant ce grand bruit.

grand

bruit. Maintenant ça fait très peur à tout le monde. Dans la rue il y a presque personne qui ose sortir, il y a que des gens courageux qui sortent dehors.

Depuis cette aventure, les enfants ne sortent plus la nuit.

# Objet d'étude : la continuité référentielle

Cette maison est hantée.

Elle habitait dans cette maison depuis longtemps.

> longtemps.

Il marcha marcha : Il se retourna en entendant ce grand bruit.

grand

bruit. Maintenant ça fait très peur à tout le monde. Dans la rue il y a presque personne qui ose sortir, il y a que des gens courageux qui sortent dehors.

Depuis cette aventure, les enfants ne sortent plus la nuit.



# Objet d'étude : la continuité référentielle

Cette maison est hantée.

Elle habitait dans cette maison depuis longtemps.

> longtemps.

Il marcha marcha : Il se retourna en entendant ce grand bruit.

grand

bruit. Maintenant ça fait très peur à tout le monde. Dans la rue il y a presque personne qui ose sortir, il y a que des gens courageux qui sortent dehors.

Depuis cette aventure, les enfants ne sortent plus la nuit.



# Objet d'étude : la continuité référentielle

Cette maison est hantée.

Elle habitait dans cette maison depuis longtemps.

> longtemps.

Il marcha marcha : Il se retourna en entendant ce grand bruit.

grand

bruit. Maintenant ça fait très peur à tout le monde. Dans la rue il y a presque personne qui ose sortir, il y a que des gens courageux qui sortent dehors.

Depuis cette aventure, les enfants ne sortent plus la nuit.



# Objet d'étude : la continuité référentielle

Cette maison est hantée.

Elle habitait dans cette maison depuis longtemps.

> longtemps.

Il marcha marcha. Il se retourna en entendant ce grand bruit.

grand

bruit. Maintenant ça fait très peur à tout le monde. Dans la rue il y a presque personne qui ose sortir, il y a que des gens courageux qui sortent dehors.

Depuis cette aventure, les enfants ne sortent plus la nuit.



# Objet d'étude : la continuité référentielle

**Il était une fois une petite fille qui habitait dans une maison hantée.**

Elle habitait dans cette maison depuis longtemps.

> longtemps.

Il marcha marcha : Il se retourna en entendant ce grand bruit.

grand

bruit. Maintenant ça fait très peur à tout le monde. Dans la rue il y a presque personne qui ose sortir, il y a que des gens courageux qui sortent dehors.

Depuis cette aventure, les enfants ne sortent plus la nuit.



# Objet d'étude : la continuité référentielle

**Il était une fois une petite fille qui habitait dans une maison hantée.**

Elle habitait dans cette maison depuis longtemps.

> longtemps.

Il marcha marcha : Il se retourna en entendant ce grand bruit.

grand

bruit. Maintenant ça fait très peur à tout le monde. Dans la rue il y a presque personne qui ose sortir, il y a que des gens courageux qui sortent dehors.

Depuis cette aventure, les enfants ne sortent plus la nuit.



# Objet d'étude : la continuité référentielle

Il était une fois **une petite fille** qui habitait dans **une maison hantée**.

Elle habitait dans **cette maison** depuis longtemps.

> longtemps.

Il marcha marcha : Il se retourna en entendant ce grand bruit.

grand

bruit. Maintenant ça fait très peur à tout le monde. Dans la rue il y a presque personne qui ose sortir, il y a que des gens courageux qui sortent dehors.

Depuis cette aventure, les enfants ne sortent plus la nuit.



## Chaînes de référence

Suite d'expressions d'un texte entre lesquelles l'interprétation établit une identité de référence (Corblin, 1995)

### Chaîne de référence

**Mme. Seton** est-**elle** aussi connue chez-nous qu'**elle** le devrait être ?  
Je ne le crois pas : et, m'aidant de **ses** historiens, je vais donner aux lecteurs de la Revue Canadienne une biographie de **cette femme**,  
**l'une des plus accomplies, plus aimantes et des plus aimées qui**  
aient jamais foulé la terre. **Elle** a beaucoup écrit; et, autant que possible  
je **la** laisserai parler **elle-même**.

# Definitions : chaînes de référence et maillons

## Chaînes de référence

Suite d'expressions d'un texte entre lesquelles l'interprétation établit une identité de référence (Corblin, 1995)

Chaîne de  
référence

**Mme. Seton** est-**elle** aussi connue chez-nous qu'**elle** le devrait être ?  
Je ne le crois pas : et, m'aidant de **ses** historiens, je vais donner aux  
lecteurs de la Revue Canadienne une biographie de **cette femme**,  
**l'une des plus accomplies, plus aimantes et des plus aimées qui**  
aient jamais foulé la terre. **Elle** a beaucoup écrit; et, autant que possible  
je **la** laisserai parler **elle-même**.

**Référent** : entité extra-linguistique dont on parle

**Types de référent** : être humain, entité abstraite ou concrète, événement

## Chaîne de référence

**Mme. Seton** est-**elle** aussi connue chez-nous qu'**elle** le devrait être ?  
Je ne le crois pas : et, m'aidant de **ses** historiens, je vais donner aux lecteurs de la Revue Canadienne une biographie de **cette femme**, **l'une des plus accomplies, plus aimantes et des plus aimées** qui aient jamais foulé la terre. **Elle** a beaucoup écrit; et, autant que possible je **la** laisserai parler **elle-même**.

**Maillons** : expressions linguistiques signalant le **référent** et porteuse d'une **valuer instructionnelle**

- leur succession contribue à créer des **liens de cohésion** (Halliday & Hasan, 1976)
- mécanisme fondamental dans l'**organisation** et l'**interprétation** du discours

**Maillons** : expressions linguistiques signalant le **référent** et porteuse d'une **valuer instructionnelle**

- leur succession contribue à créer des **liens de cohésion** (Halliday & Hasan, 1976)
- mécanisme fondamental dans l'**organisation** et l'**interprétation** du discours

- En linguistique cognitive : théorie de l'accessibilité (Ariel, 2001) et théorie du centrage (Walker, Joshi, & Prince, 1998)
  - Principes qui régissent l'interprétation et la construction de la continuité référentielle
  
- En psycholinguistique (Gundel, Hedberg, & Zacharski, 2019 ; Kaiser & Fedele, 2019 ; Roberts, 2019 ; Salazar Orvig, 2019)
  - Facteurs qui influencent le choix des expressions référentielles (Fossard et al., 2018 ; Vogels, Krahmer, & Maes, 2013)

# Les chaînes de référence : différentes approches

- En linguistique descriptive (Charolles, 2002 ; Cornish, 2000 ; Kleiber, 1994, 2002 ; Schnedecker, 2005)
  - Description des variations selon les genres textuels ou le type de référent (Grishina & Stede, 2015 ; Kunz, Lapshinova-Koltunski, & Martinez, 2016 ; Schnedecker & Longo, 2012)
  - Observations fines, petits corpus, prise en compte d'un seul type de référent ou peu de paramètres pour caractériser les chaînes
- En traitement automatique du langage (TAL) (Landragin, 2018 ; Longo, 2013 ; Poesio, Pradhan, Recasens, Rodriguez, & Versley, 2016 ; Wilkens, Oberle, Landragin, & Todirascu, 2020)
  - Détection automatique des chaînes de référence et des maillons (De Marneffe, Recasens, & Potts, 2015 ; Mitkov, 2014 ; Recasens & Hovy, 2010)
  - Observations systématiques, diffusion de gros corpus annotés

Délimitation des chaînes

Chaînes de  
référence :  
au moins  
**trois** maillons

Délimitation des chaînes

Conception des maillons

## Conception Charolles

**Hector de Gribelin** rencontra dans ce monde une jeune fille noble et pauvre comme **lui**, et l'épousa [...] Mais voilà que, vers le printemps, un travail supplémentaire fut confié à **l'employé** par son chef, et **il** reçut une gratification extraordinaire de trois cents francs

Chaînes de référence :  
au moins  
**trois** maillons

Délimitation des chaînes

Conception des maillons

## Conception Ariel

**Hector de Gribelin** rencontra dans ce monde une jeune fille noble et pauvre comme **lui**, et l'épousa [...]. Mais voilà que, vers le printemps, un travail supplémentaire fut confié à **l'employé** par **son** chef, et **il** reçut une gratification extraordinaire de trois cents francs

Chaînes de référence :  
au moins  
**trois** maillons

Délimitation des chaînes

Conception des maillons

## Conception Landragin

**Hector de Gribelin** rencontra dans ce monde une jeune fille noble et pauvre comme **lui**, et **l'épousa** [...] Mais voilà que, vers le printemps, un travail supplémentaire fut confié à **l'employé** par **son** chef, et **il** reçut une gratification extraordinaire de trois cents francs

Chaînes de référence :  
au moins  
**trois** maillons

# Une définition qui ne fait pas consensus

Délimitation des chaînes

Conception des maillons

Délimitation des maillons

Chaînes de  
référence :  
au moins  
**trois** maillons

[Zsa Zsa, who slap a security guard]<sub>{zsazsa}</sub>

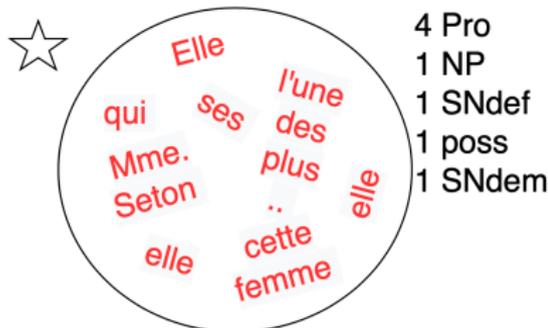
[Zsa Zsa]<sub>{zsazsa}</sub>, who slap a security guard

[Une bouteille d'eau]<sub>{b}</sub>

[Une bouteille d'[eau]<sub>{eau}</sub>]<sub>{b}</sub>

**Mme. Seton** est-elle aussi connue chez-nous qu'**elle** le devrait être ? Je ne le crois pas : et, m'aidant de **ses** historiens, je vais donner aux lecteurs de la Revue Canadienne une biographie de **cette femme, l'une des plus accomplies, plus. aimantes et des plus aimées qui** aient jamais foulé la terre. **Elle** a beaucoup écrit ; [...]

Les pauvres gens vivaient péniblement des petits appointements **du mari**. **Hector de Gribelin** avait été élevé en province, dans le manoir paternel, par un vieil abbé précepteur..Puis, à vingt ans, on lui avait cherché une position, et **il** était entré, commis à quinze cents francs, au ministère de la Marine. **Il** avait échoué sur cet écueil comme tous ceux [...]. **Ses** trois premières années de bureau furent horribles. **Il** avait retrouvé quelques amis de **sa** famille, vieilles gens attardés [...] ; et **il** s'était fait un cercle de connaissances.



**Mme. Seton** est-elle aussi connue chez-nous qu'elle le devrait être ? Je ne le crois pas : et, m'aidant de **ses** historiens, je vais donner aux lecteurs de la Revue Canadienne une biographie de **cette femme, l'une des plus accomplies, plus aimantes et des plus aimées qui** aient jamais foulé la terre. **Elle** a beaucoup écrit ; [...]

Les pauvres gens vivaient péniblement des petits appointements **du mari**. **Hector de Gribelin** avait été élevé en province, dans le manoir paternel, par un vieil abbé précepteur. Puis, à vingt ans, on lui avait cherché une position, et **il** était entré, commis à quinze cents francs, au ministère de la Marine. **Il** avait échoué sur cet écueil comme tous ceux [...]. **Ses** trois premières années de bureau furent horribles. **Il** avait retrouvé quelques amis de **sa** famille, vieilles gens attardés [...] ; et **il** s'était fait un cercle de connaissances.



- ⇒ Important de rester dans la linéarité, analyser comment les maillons s'enchaînent
- Analyser les stratégies de continuité référentielle que réalisent les CR
  - Analyser l'influence de l'organisation globale du discours sur ces stratégies
  - S'intéresser aux facteurs qui peuvent influencer les choix référentiels (genre textuel, type de référent (Longo & Todirascu, 2014))
  - Confronter la théorie de l'accessibilité et du centrage à la réalité des usages en corpus

- ⇒ Important de rester dans la linéarité, analyser comment les maillons s'enchaînent
- Analyser les stratégies de continuité référentielle que réalisent les CR
  - Analyser l'influence de l'organisation globale du discours sur ces stratégies
  - S'intéresser aux facteurs qui peuvent influencer les choix référentiels (genre textuel, type de référent (Longo & Todirascu, 2014))
  - Confronter la théorie de l'accessibilité et du centrage à la réalité des usages en corpus

- ⇒ Important de rester dans la linéarité, analyser comment les maillons s'enchaînent
- Analyser les stratégies de continuité référentielle que réalisent les CR
  - Analyser l'influence de l'organisation globale du discours sur ces stratégies
  - S'intéresser aux facteurs qui peuvent influencer les choix référentiels (genre textuel, type de référent (Longo & Todirascu, 2014))
  - Confronter la théorie de l'accessibilité et du centrage à la réalité des usages en corpus

- ⇒ Important de rester dans la linéarité, analyser comment les maillons s'enchaînent
- Analyser les stratégies de continuité référentielle que réalisent les CR
  - Analyser l'influence de l'organisation globale du discours sur ces stratégies
  - S'intéresser aux facteurs qui peuvent influencer les choix référentiels (genre textuel, type de référent (Longo & Todirascu, 2014))
  - Confronter la théorie de l'accessibilité et du centrage à la réalité des usages en corpus

- ⇒ Important de rester dans la linéarité, analyser comment les maillons s'enchaînent
- Analyser les stratégies de continuité référentielle que réalisent les CR
  - Analyser l'influence de l'organisation globale du discours sur ces stratégies
  - S'intéresser aux facteurs qui peuvent influencer les choix référentiels (genre textuel, type de référent (Longo & Todirascu, 2014))
  - Confronter la théorie de l'accessibilité et du centrage à la réalité des usages en corpus

⇒ Important de croiser plusieurs ressources

- Analyser la variété des chaînes de référence
- Tirer partie de l'hétérogénéité des ressources pour aboutir à une définition partagée

⇒ Important de croiser plusieurs ressources

- Analyser la variété des chaînes de référence

- Tirer partie de l'hétérogénéité des ressources pour aboutir à une définition partagée

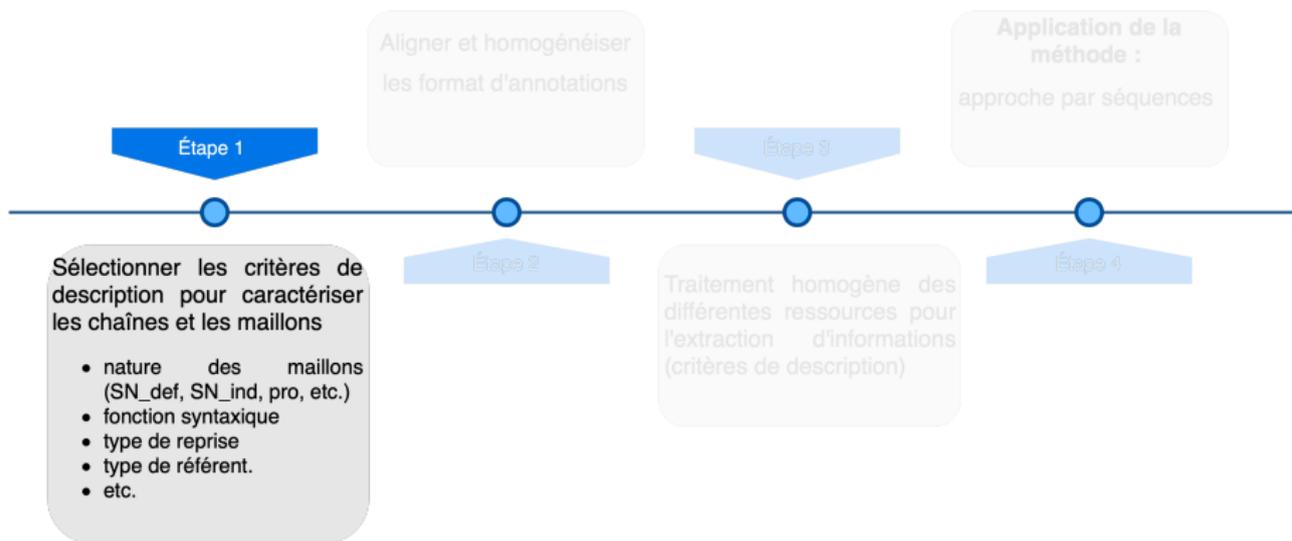
- ⇒ Important de croiser plusieurs ressources
- Analyser la variété des chaînes de référence
  - Tirer partie de l'hétérogénéité des ressources pour aboutir à une définition partagée

- ⇒ Important de croiser plusieurs ressources
- Analyser la variété des chaînes de référence
  - Tirer partie de l'hétérogénéité des ressources pour aboutir à une définition partagée

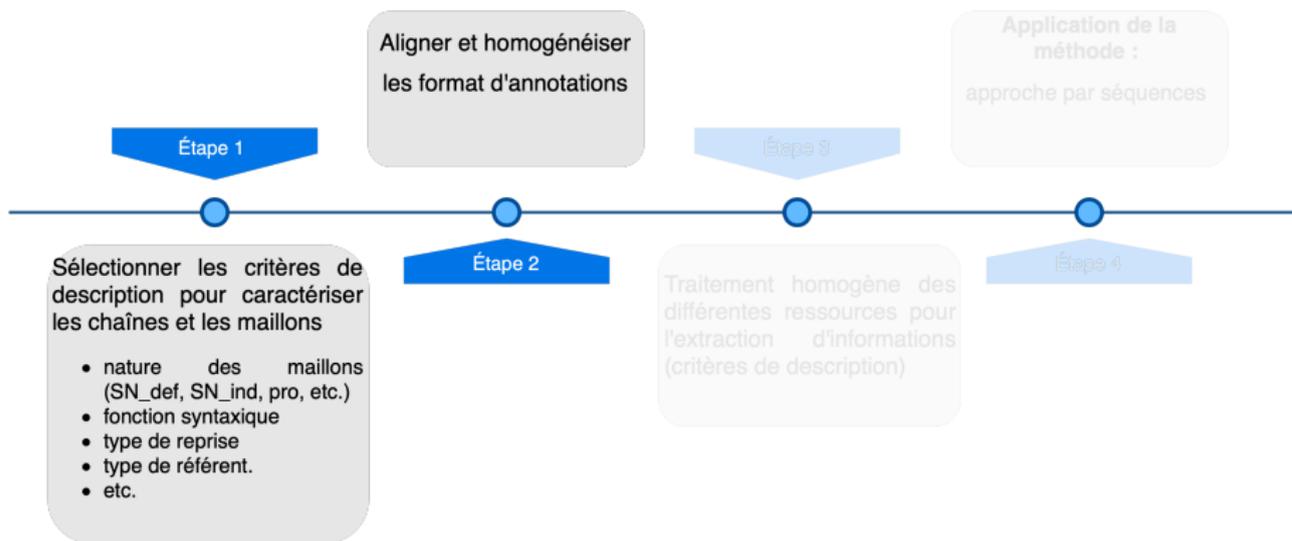
## Typologie des chaînes de référence

Description **systematique** et **exhaustive** de la **variété** et de la **complexité** des chaînes de référence à la lumière de **corpus diversifiés**

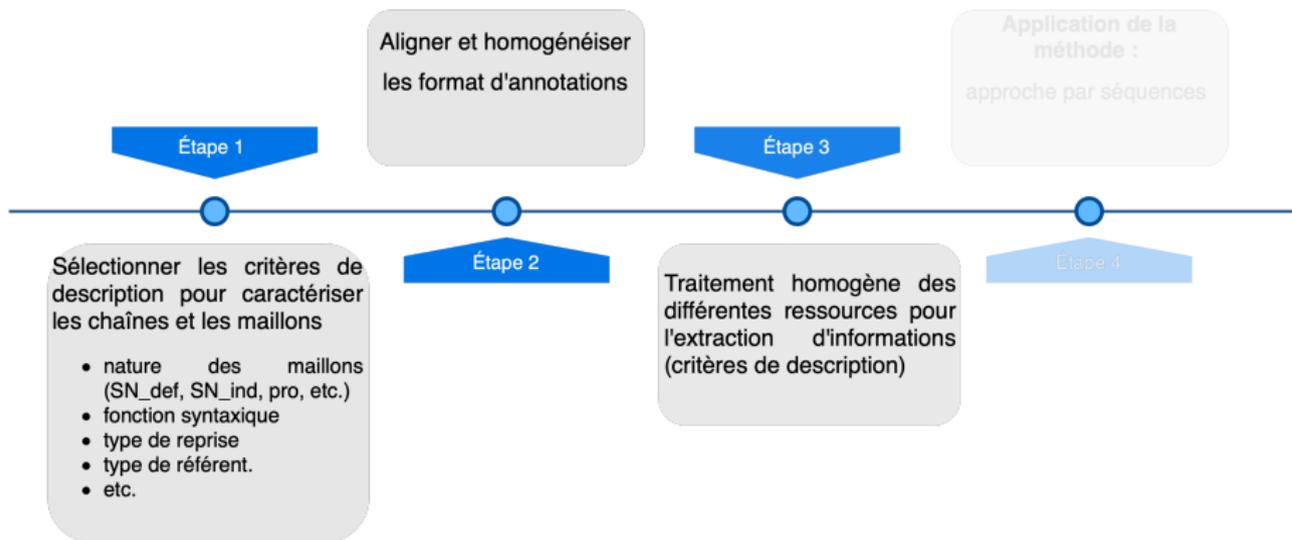
# Plusieurs étapes



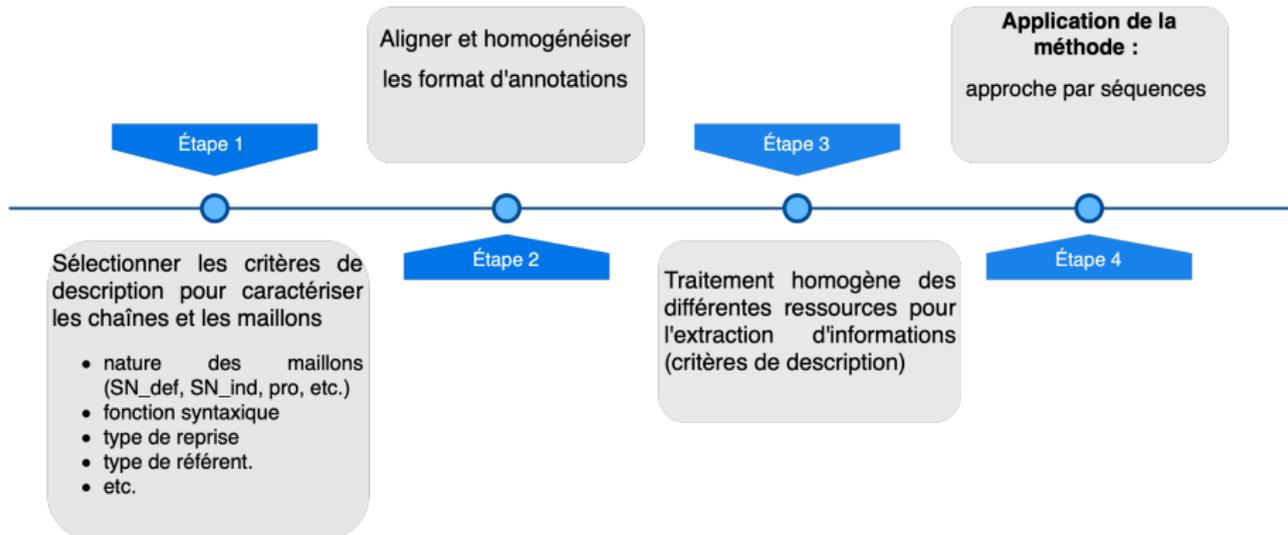
# Plusieurs étapes



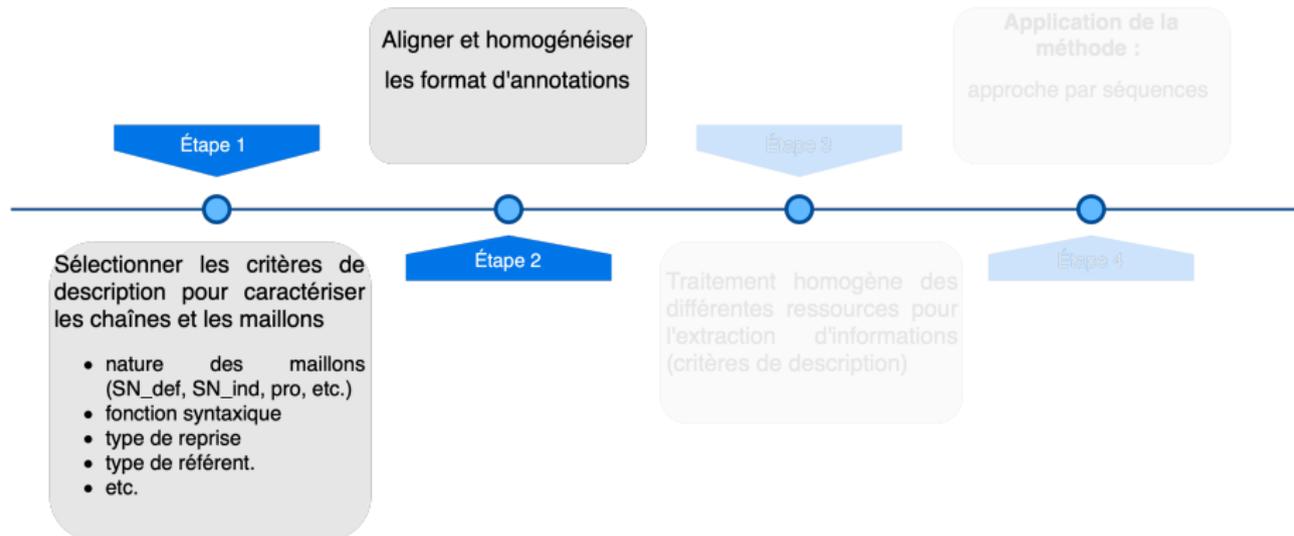
# Plusieurs étapes



# Plusieurs étapes



# Aligner et homogénéiser les formats d'annotation



## Objectifs différents : annotations différentes

MUC (MUC consortium 1995c)

AnCora (Taulé et al., 2008)

ACE (Doddington et al., 2004)

Différents modèles d'annotation :

- ① types de référent
- ② critères de délimitation
- ③ annotation des singletons
- ④ informations annotées

AnnoDis (Péry-Woodley et al., 2011)

Ontonotes (Pradhan et al., 2012)

Différents formats :

- ① SGML (in-line/stand-off)
- ② MMAX2
- ③ Glozz
- ④ XML-TEI-URS

AnCor (Muzurelle et al., 2014)

ARRAU (Poesio et al., 2013)

Democrat (Landragin et al., 2019)

WikiCoref (Ghaddar & Langlais, 2016)

E-Calm (Garcia-Debanc et al., 2017)

## Objectifs différents : annotations différentes

MUC (MUC consortium 1995c)

AnCora (Taulé et al., 2008)

ACE (Doddington et al., 2004)

### Différents modèles d'annotation :

- ① types de référent
- ② critères de délimitation
- ③ annotation des singletons
- ④ informations annotées

AnnoDis (Péry-Woodley et al., 2011)

Ontonotes (Pradhan et al., 2012)

### Différents formats :

- ① SGML (in-line/stand-off)
- ② MMAX2
- ③ Glozz
- ④ XML-TEI-URS

AnCor (Muzerelle et al., 2014)

ARRAU (Poesio et al., 2013)

Democrat (Landragin et al., 2019)

WikiCoref (Ghaddar & Langlais, 2016)

E-Calm (Garcia-Debanc et al., 2017)

- AnnoDis (2011) : textes longs entièrement annotés, non narratifs, 3 genres textuels (rapports et articles scientifiques, textes encyclopédiques)
  - Sonder l'intuition des locuteurs
  - Annotateurs naïfs

- AnnoDis (2011) : textes longs entièrement annotés, non narratifs, 3 genres textuels (rapports et articles scientifiques, textes encyclopédiques)
  - Sonder l'intuition des locuteurs
  - Annotateurs naïfs
- DEMOCRAT (2019) : échantillons des textes (10 000 mots), narratifs (en prose et en vers) et non narratifs, 18 genres textuels, du 12<sup>ème</sup> au 21<sup>ème</sup> siècle
  - Annoter toutes les expressions référentielles
  - Annotateurs experts

- AnnoDis (2011) : textes longs entièrement annotés, non narratifs, 3 genres textuels (rapports et articles scientifiques, textes encyclopédiques)
  - Sonder l'intuition des locuteurs
  - Annotateurs naïfs
- DEMOCRAT (2019) : échantillons des textes (10 000 mots), narratifs (en prose et en vers) et non narratifs, 18 genres textuels, du 12<sup>ème</sup> au 21<sup>ème</sup> siècle
  - Annoter toutes les expressions référentielles
  - Annotateurs experts
- RésolCo (E-CALM, 2019-2021) : écrits scolaires à différents niveaux de littératie (du primaire à l'Université)
  - Annoter les référents humains principaux introduit dans la consigne
  - Annotateurs experts

Corpus	Textes	Mots	CR	Maillons	Singletons	
					#	%
<b>AnnoDis</b>	87	666 000	581	3456	0	0
<b>Democrat</b>	58	688 851	20 410	124 529	72 394	36,8
<b>RésolCo</b>	383	70 257	1 979	12 737	138	2

Corpus	Textes	Mots	CR	Maillons	Singletons	
					#	%
<b>AnnoDis</b>	87	666 000	581	3456	0	0
<b>Democrat</b>	58	688 851	20 410	124 529	72 394	36,8
<b>RésolCo</b>	383	70 257	1 979	12 737	138	2

- Différents choix d'annotation

Un jour, **[le père de Léonard]**<sub>Léo</sub>, ser Piero, « prit plusieurs de **[ses]**<sub>Léo</sub> dessins et les soumit à son ami Andrea del Verrocchio qu'il pria instamment de lui dire si **[Léonard]**<sub>Léo</sub> devait se consacrer à l'art du dessin et s'**[il]**<sub>Léo</sub> pourrait parvenir à quelque chose en cette matière. Andrea s'étonna fort des débuts extraordinaires **[de Léonard]**<sub>Léo</sub> et exhorta ser Piero à **[lui]**<sub>Léo</sub> permettre de choisir ce métier, sur quoi, ser Piero résolut que **[Léonard]**<sub>Léo</sub> entrerait à l'atelier d'Andrea.

AnnoDis

Guidée par un prémarquage automatique

DEMOCRAT

Volonté d'annoter toutes les expressions référentielles

Un jour, **[le père de Léonard]**<sub>Léo</sub>, ser Piero, « prit plusieurs de **[ses]**<sub>Léo</sub> dessins et les soumit à son ami Andrea del Verrocchio qu'il pria instamment de lui dire si **[Léonard]**<sub>Léo</sub> devait se consacrer à l'art du dessin et s'**[il]**<sub>Léo</sub> pourrait parvenir à quelque chose en cette matière. Andrea s'étonna fort des débuts extraordinaires **[de Léonard]**<sub>Léo</sub> et exhorta ser Piero à **[lui]**<sub>Léo</sub> permettre de choisir ce métier, sur quoi, ser Piero résolut que **[Léonard]**<sub>Léo</sub> entrerait à l'atelier d'Andrea.

AnnoDis

Guidée par un prémarquage automatique

Un jour, **[le père de [Léonard]**<sub>Léo</sub> ]<sub>P</sub>, ser Piero, « prit plusieurs de **[ses]**<sub>Léo</sub> **dessins]**<sub>d</sub> et **[les]**<sub>d</sub> soumit à **[son]**<sub>P</sub> **ami Andrea del Verrocchio]**<sub>AV</sub> qu'**[il]**<sub>P</sub> pria instamment de **[lui]**<sub>P</sub> dire si **[Léonard]**<sub>Léo</sub> devait se consacrer à **[l'art du dessin]**<sub>Ad</sub> et s'**[il]**<sub>Léo</sub> pourrait parvenir à quelque chose en **[cette matière]**<sub>Ad</sub>. **[Andrea]**<sub>AV</sub> s'étonna fort des débuts extraordinaires de **[Léonard]**<sub>Léo</sub> et **[exhorta]**<sub>AV</sub> **[ser Piero]**<sub>P</sub> à **[lui]**<sub>Léo</sub> permettre de choisir **[ce métier]**<sub>Ad</sub>, sur quoi, **[ser Piero]**<sub>P</sub> résolut que **[Léonard]**<sub>Léo</sub> entrerait à **[l'atelier d'[Andrea]**<sub>AV</sub> ]<sub>y</sub>

DEMOCRAT

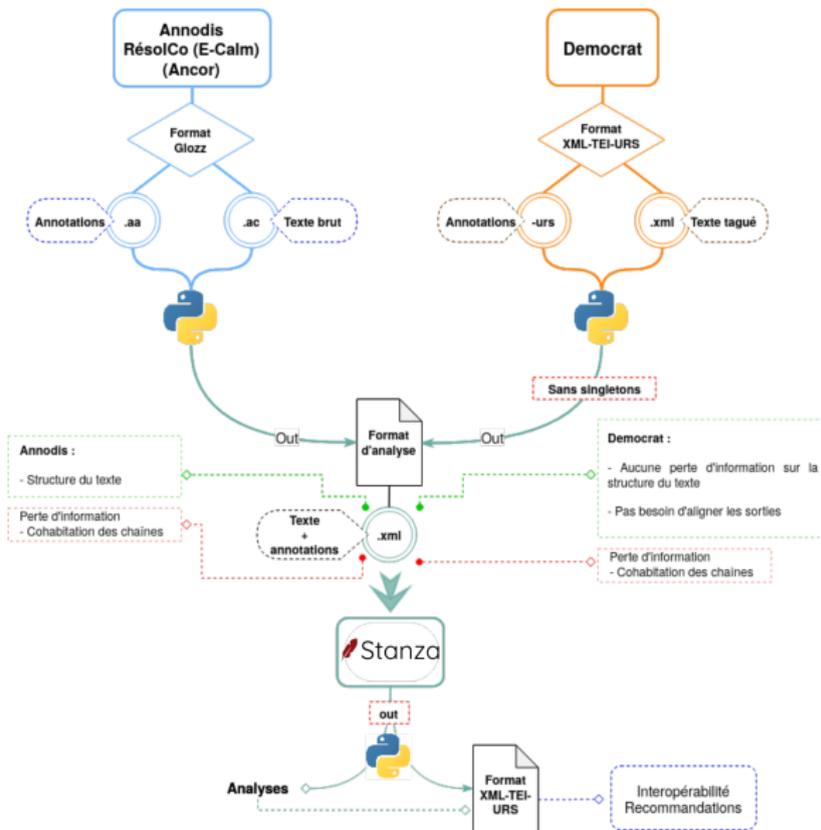
Volonté d'annoter toutes les expressions référentielles

	AnnoDis	Democrat	RésolCo
<b>Type d'approche</b>	Descendante	Ascendante	Ascendante
<b>Conception</b>	Charolles	Landragin	Landragin (adaptée)
<b>Pré-marquage</b>	✓	✗	✓
<b>Structure de l'annotation</b>	Schéma	Schéma	Relations
<b>Type de référent</b>	Tout type	Pas événements	Humain
<b>Délimitation CR</b>	2 phrases min / 1 section max	Aucune	Aucune
<b>Délimitation maillons</b>	Pas guidée	Guidée	Guidée
<b>Singletons</b>	✗	✓	✓
<b>Propriétés annotées</b>	Nature maillon	Aucune	Aucune

	AnnoDis	Democrat	RésolCo
<b>Type d'approche</b>	Descendante	Ascendante	Ascendante
<b>Conception</b>	Charolles	Landragin	Landragin (adaptée)
<b>Pré-marquage</b>	✓	✗	✓
<b>Structure de l'annotation</b>	Schéma	Schéma	Relations
<b>Type de référent</b>	Tout type	Pas événements	Humain
<b>Délimitation CR</b>	2 phrases min / 1 section max	Aucune	Aucune
<b>Délimitation maillons</b>	Pas guidée	Guidée	Guidée
<b>Singletons</b>	✗	✓	✓
<b>Propriétés annotées</b>	Nature maillon	Aucune	Aucune

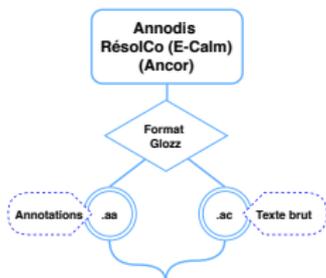
**Ressources hétérogènes : comment les traiter de façon homogène ?**

# Aligner les formats d'annotation



# Aligner les formats d'annotation

## Point en commun : annotationn déportée



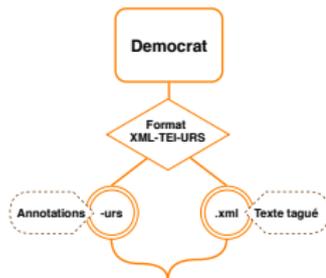
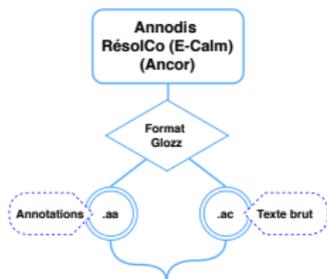
Renvoi au texte : nombre de caractères

```
</unit>
<unit id="markmacr_1252528266">
  <metadata>
    <author>markmacr</author>
    <creation-date>1252528266</creation-date>
  </metadata>
  <characterisation>
    <type>COREFpropos</type>
    <featureSet/>
  </characterisation>
  <positioning>
    <start>
      <singlePosition index="9146"/>
    </start>
    <end>
      <singlePosition index="9151"/>
    </end>
  </positioning>
```

e le terrorisme : essai de bilan institutionnel François Vergnolle de  
is Vergnolle de Chantal, Docteur en Sciences Politiques de l'IEP de Paris,  
Conférences en civilisation américaine à l'université de Bourgogne. Depuis  
e d'années, les républicains se sont fait fort de réduire le poids de l'Etat  
actuelle lutte contre le terrorisme, menée par une équipe républicaine qui,  
re totalement aux critiques contre le Big Government, renouveau en cause  
conservateur en faveur de la décentralisation. Les différentes mesures  
is septembre 2001 vont toutes dans le même sens, un considérable renforcement  
de l'Etat fédéral. Comme toutes les guerres menées par les Etats-Unis, celle  
le terrorisme risquerait-elle aussi, de renforcer la centralisation. Quels  
ts de ce retour de l'Etat central, comment s'opère la recentralisation, et  
onsequentes dans l'Etat fédéral. Finalement, quelles sont les  
tirés de cette évolution? En particulier, comment s'articule la lutte contre  
avec l'engagement conservateur en faveur des Etats fédérés? Selon nous, la  
e terrorisme ne serait pas similaire aux évolutions entraînés par les autres  
débouche en fait sur un activisme top-down, qui concerne aussi bien  
que les Etats fédérés et les autorités locales (villes, comtés). Plutôt que  
entralisation, il faudrait évoquer un renforcement des fonctions légitimes de  
eaux du gouvernement : la défense et la protection des citoyens pour le  
; les autorités locales, elles, perdront les moyens de réponse immédiats aux  
roristes (police, pompier, santé). L'essentiel des problèmes suscités par la  
territoire contre le terrorisme réside dans la coordination entre les  
anes. L'administration actuelle s'engage résolument dans cette voie, et  
rganisation massive des administrations nationales. Face à l'urgence : les

# Aligner les formats d'annotation

## Point en commun : annotationn déportée



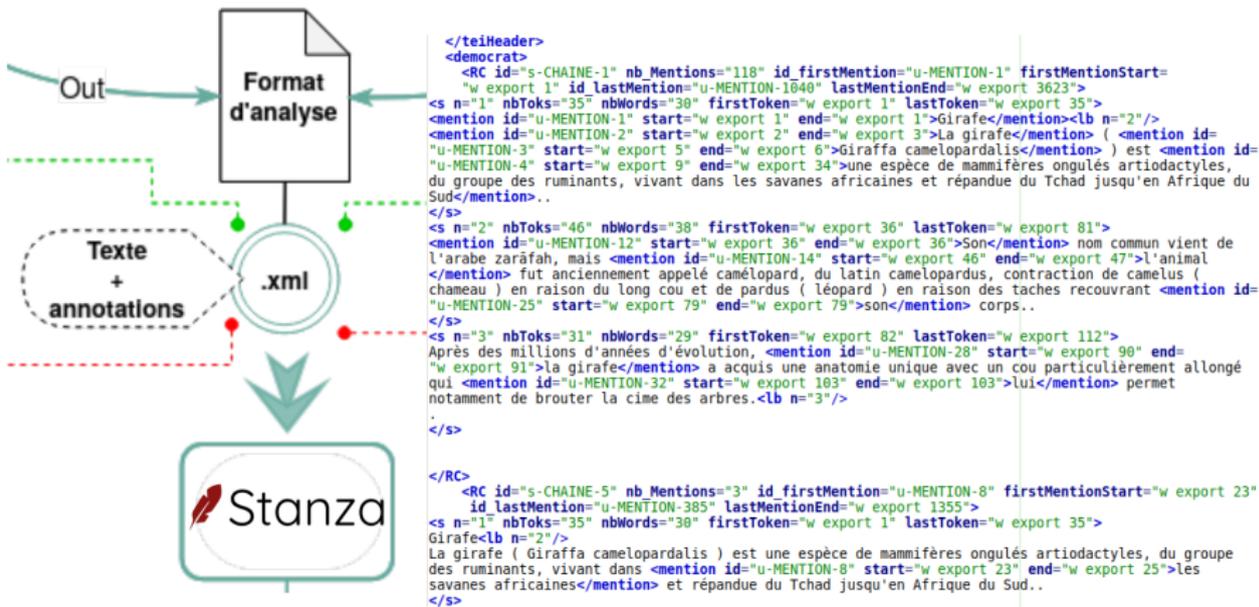
Renvoi au texte : nombre de tokens

```
</unit>
<unit id="markmacr_1252528"
  <metadata>
    <author>markmacr/
    <creation-date>125
  </metadata>
  <characterisation>
    <type>COREFpropos
    <featureSet>
  </characterisation>
  <positioning>
    <start>
      <singlePosition
    </start>
    <end>
      <singlePosition
    </end>
  </positioning>
  <annotations type="coreference">
    <annotationGrp type="unit" subtype="MENTION">
      <span id="u-MENTION-1" from="text:w.export_1" to="text:w.export_1" ana="#u-MENTION-1-fs"/>
      <span id="u-MENTION-2" from="text:w.export_2" to="text:w.export_3" ana="#u-MENTION-2-fs"/>
      <span id="u-MENTION-3" from="text:w.export_5" to="text:w.export_6" ana="#u-MENTION-3-fs"/>
    </annotationGrp>
    <link id="s-CHAINE-1" target="#u-MENTION-1040 #u-MENTION-202 #u-MENTION-686 #u-MENTION-687
      #u-MENTION-841 #u-MENTION-221 #u-MENTION-689 #u-MENTION-104 #u-MENTION-285 #u-MENTION-849
      #u-MENTION-289 #u-MENTION-14 #u-MENTION-844 #u-MENTION-842 #u-MENTION-689 #u-MENTION-326
      #u-MENTION-839 #u-MENTION-281 #u-MENTION-213 #u-MENTION-576 #u-MENTION-455 #u-MENTION-335
      #u-MENTION-852 #u-MENTION-574 ana="#s-CHAINE-1-fs"/>
    </link>
  </annotations>
  <div type="unit-fs">
    <fs id="u-MENTION-1-fs">
      <f name="REF">
        <string>G/string
      </f>
    </div>
    <div type="schema-fs">
      <fs id="s-CHAINE-1-fs">
        <f name="REF">
          <string>G/string
        </f>
        <f name="NB MAILLONS">
          <string>118G/string
        </f>
      </div>
  </div>
  <div id="export">
    <lb n="1"/>
    <w id="w.export_1" n="1">
      <tax:form>G/rafec/tax:form
      <tax:ana resp="#txm" type="#frpos">NDRK/tax:ana
      <tax:ana resp="#txm" type="#frlemma">G/rafec/tax:ana
    </w>
    <lb n="2"/>
    <w id="w.export_2" n="2">
      <tax:form>Lc/tax:form
      <tax:ana resp="#txm" type="#frpos">DET:ART/tax:ana
      <tax:ana resp="#txm" type="#frlemma">Lc/tax:ana
    </w>
    <w id="w.export_3" n="3">
      <tax:form>G/rafec/tax:form
      <tax:ana resp="#txm" type="#frpos">NDRK/tax:ana
      <tax:ana resp="#txm" type="#frlemma">G/rafec/tax:ana
    </w>
    <w id="w.export_4" n="4">
      <tax:form>Lc/tax:form
      <tax:ana resp="#txm" type="#frpos">PUN/tax:ana
      <tax:ana resp="#txm" type="#frlemma">Lc/tax:ana
    </w>
    <w id="w.export_5" n="5">
      <tax:form>G/rafec/tax:form
      <tax:ana resp="#txm" type="#frpos">NDRK/tax:ana
      <tax:ana resp="#txm" type="#frlemma">G/rafec/tax:ana
    </w>
  </div>
  <div ont="1">
    cette vole, et
  </div>
  <div ont="2">
    à l'urgence : les
  </div>
```

# Aligner les formats d'annotation

**Besoin** : ne pas perdre d'informations et effectuer des traitements homogènes

**Proposition** : format intermédiaire

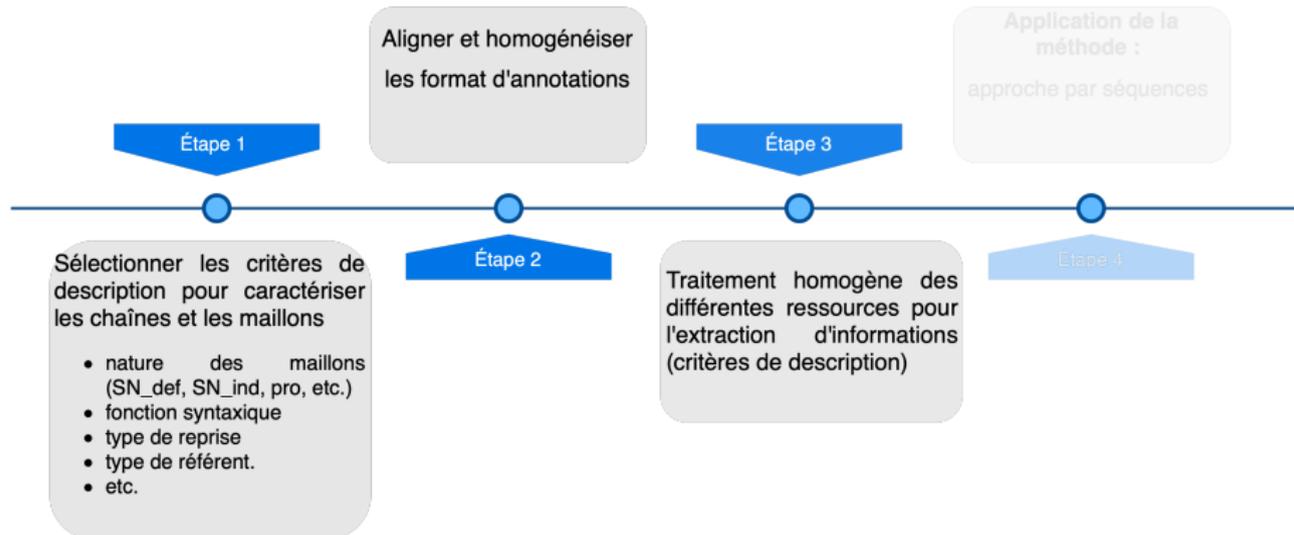


Informations morpho-syntaxiques

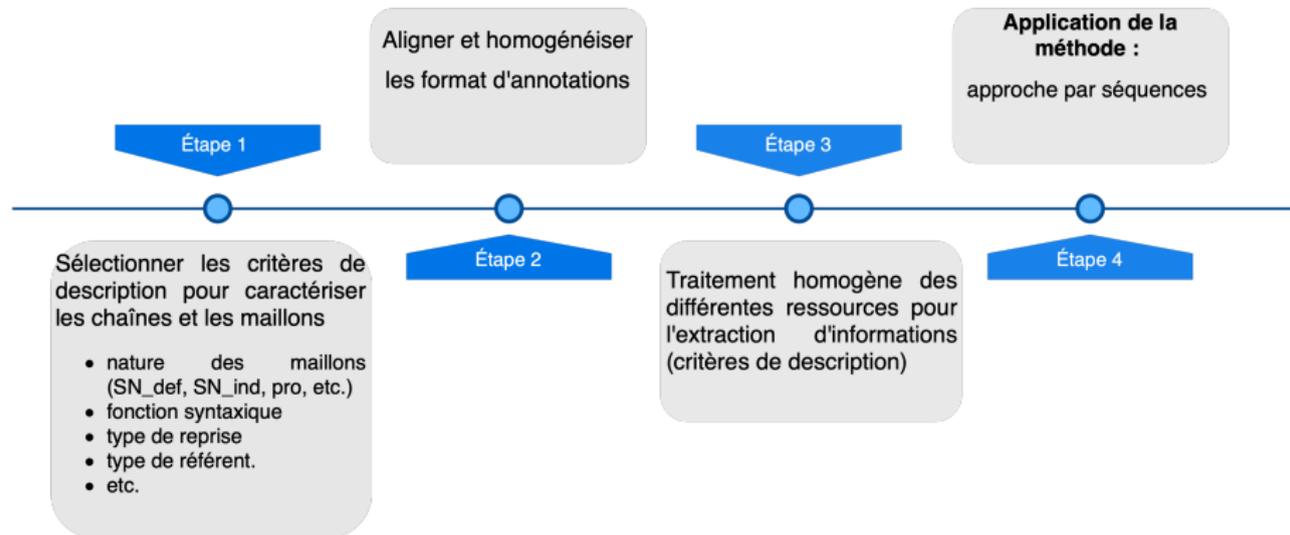
# Aligner les formats d'annotation

1 Girafe	Girafe	PROPN	-	0/root	root	B-PER	0	6	S-mention	u-MENTION-1
1 La	le	DET	Definite=Def Gender=Fem Number=Sing PronType=Art	2/girafe	det	O	8	10	B-mention	u-MENTION-2
2 girafe	girafe	NOUN	Gender=Fem Number=Sing	9/espèce	nsubj	O	11	17	E-mention	u-MENTION-2
3 (	(	PUNCT	-	4/Giraffa	punct	O	18	19	-	-
4 Giraffa	Giraffa	PROPN	-	2/girafe	appos	B-LOC	20	27	B-mention	u-MENTION-3
5 camelopardalis	camelopardalis	PROPN	-	4/Giraffa	flat_name	E-LOC	28	42	E-mention	u-MENTION-3
6 )	)	PUNCT	-	4/Giraffa	punct	O	43	44	-	-
7 est	être	AUX	Mood=Ind Number=Sing Person=3 Tense=Pres VerbForm=Fin	9/espèce	cop	O	45	48	-	-
8 une	un	DET	Definite=Ind Gender=Fem Number=Sing PronType=Art	9/espèce	det	O	49	52	B-mention	u-MENTION-4
9 espèce	espèce	NOUN	Gender=Fem Number=Sing	0/root	root	O	53	59	I-mention	u-MENTION-4
10 de	de	ADP	-	11/mammifères	case	O	60	62	I-mention	u-MENTION-4
11 mammifères	mammifère	NOUN	Gender=Masc Number=Plur	9/espèce	rmod	O	63	73	I-mention	u-MENTION-4
12 ongulés	ongulé	ADJ	Gender=Masc Number=Plur	11/mammifères	amod	O	74	81	I-mention	u-MENTION-4
13 artiodactyles	artiodactyl	ADJ	Gender=Masc Number=Plur	11/mammifères	amod	O	82	95	I-mention	u-MENTION-4
14 .	.	PUNCT	-	17/groupe	punct	O	95	96	I-mention	u-MENTION-4
15 de	de	ADP	-	17/groupe	case	O	97	99	I-mention	u-MENTION-4
16 le	le	DET	Definite=Def Gender=Masc Number=Sing PronType=Art	17/groupe	det	O	97	99	I-mention	u-MENTION-4
17 groupe	groupe	NOUN	Gender=Masc Number=Sing	11/mammifères	conj	O	100	106	I-mention	u-MENTION-4
18 de	de	ADP	-	20/nominants	case	O	107	110	I-mention	u-MENTION-4
19 les	le	DET	Definite=Def Number=Plur PronType=Art	20/nominants	det	O	107	110	I-mention	u-MENTION-4
20 ruminants	ruminant	NOUN	Gender=Masc Number=Plur	17/groupe	rmod	O	111	120	I-mention	u-MENTION-4
21 .	.	PUNCT	-	22/vivant	punct	O	120	121	I-mention	u-MENTION-4
22 vivant	vivre	VERB	Tense=Pres VerbForm=Part	9/espèce	acl	O	122	128	I-mention	u-MENTION-4
23 dans	dans	ADP	-	25/savanes	case	O	129	133	I-mention	u-MENTION-4
24 les	le	DET	Definite=Def Number=Plur PronType=Art	25/savanes	det	O	134	137	I-mention	u-MENTION-4
25 savanes	savane	NOUN	Gender=Fem Number=Plur	22/vivant	obl_arg	O	138	145	I-mention	u-MENTION-4
26 africaines	africain	ADJ	Gender=Fem Number=Plur	25/savanes	amod	O	146	156	I-mention	u-MENTION-4
27 et	et	CCONJ	-	28/répandue	cc	O	157	159	I-mention	u-MENTION-4
28 répandue	répandu	VERB	Gender=Fem Number=Sing Tense=Past VerbForm=Part	26/africaines	conj	O	160	168	I-mention	u-MENTION-4
29 de	de	ADP	-	31/Tchad	case	O	169	171	I-mention	u-MENTION-4
30 le	le	DET	Definite=Def Gender=Masc Number=Sing PronType=Art	31/Tchad	det	O	169	171	I-mention	u-MENTION-4
31 Tchad	Tchad	PROPN	Gender=Masc Number=Sing	25/savanes	rmod	S-LOC	172	177	I-mention	u-MENTION-4
32 jusqu'	jusque	ADP	-	34/Afrique	case	O	178	184	I-mention	u-MENTION-4
33 en	en	ADP	-	34/Afrique	case	O	184	186	I-mention	u-MENTION-4
34 Afrique	Afrique	PROPN	-	22/vivant	obl_mod	B-LOC	187	194	I-mention	u-MENTION-4
35 de	de	ADP	-	37/Sud	case	I-LOC	195	197	I-mention	u-MENTION-4
36 le	le	DET	Definite=Def Gender=Masc Number=Sing PronType=Art	37/Sud	det	I-LOC	195	197	I-mention	u-MENTION-4
37 Sud	sud	NOUN	Gender=Masc Number=Sing	34/Afrique	rmod	E-LOC	198	201	E-mention	u-MENTION-4
38 ..	..	PUNCT	-	9/espèce	punct	O	202	204	-	-
1 Son	son	DET	Number=Sing PossNumber=Sing PossPerson=3 PronType=Prs	2/nom	det	O	205	208	S-mention	u-MENTION-12
2 nom	nom	NOUN	Gender=Masc Number=Sing	4/vient	nsubj	O	209	212	-	-
3 commun	commun	ADJ	Gender=Masc Number=Sing	2/nom	amod	O	213	219	-	-
4 vient	venir	VERB	Mood=Ind Number=Sing Person=3 Tense=Pres VerbForm=Fin	0/root	root	O	220	225	-	-
5 de	de	ADP	-	7/arabe	case	O	226	228	-	-
6 l'	le	DET	Definite=Def Number=Sing PronType=Art	7/arabe	det	O	229	231	-	-

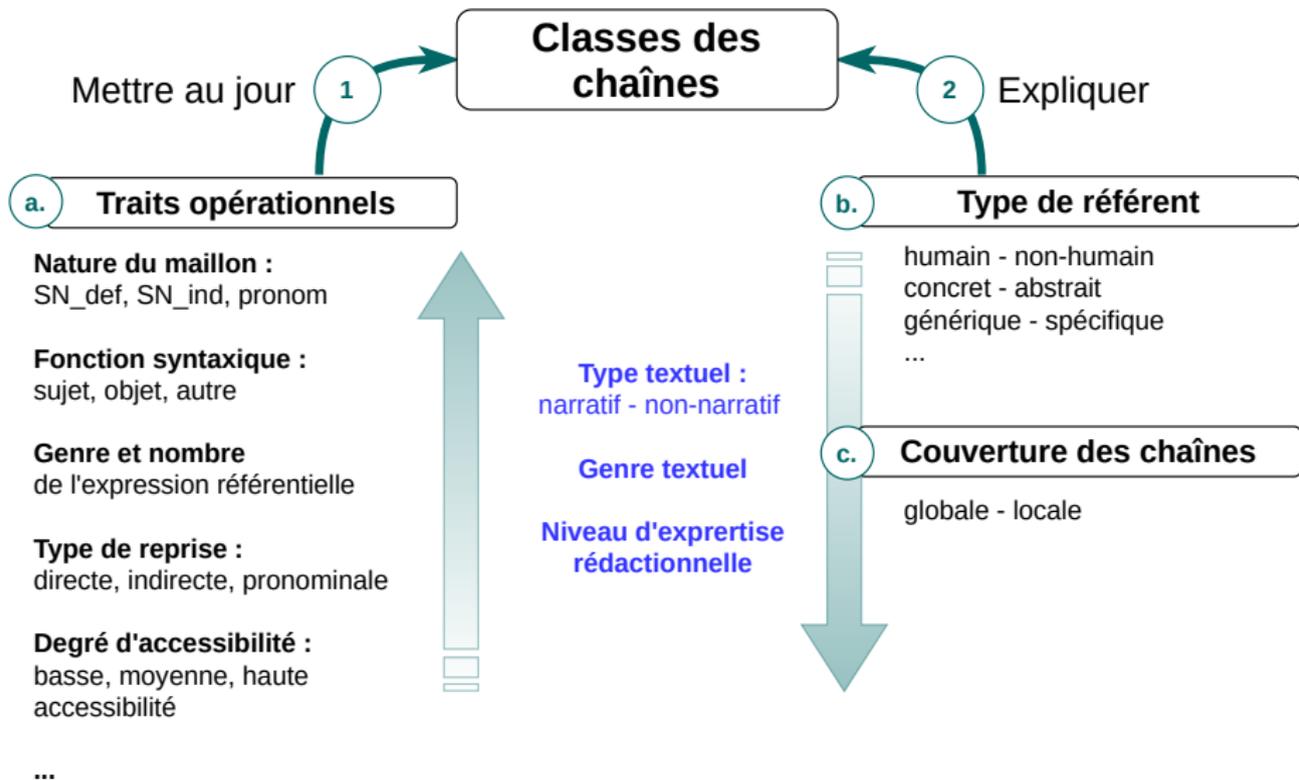
# Aligner les formats d'annotation



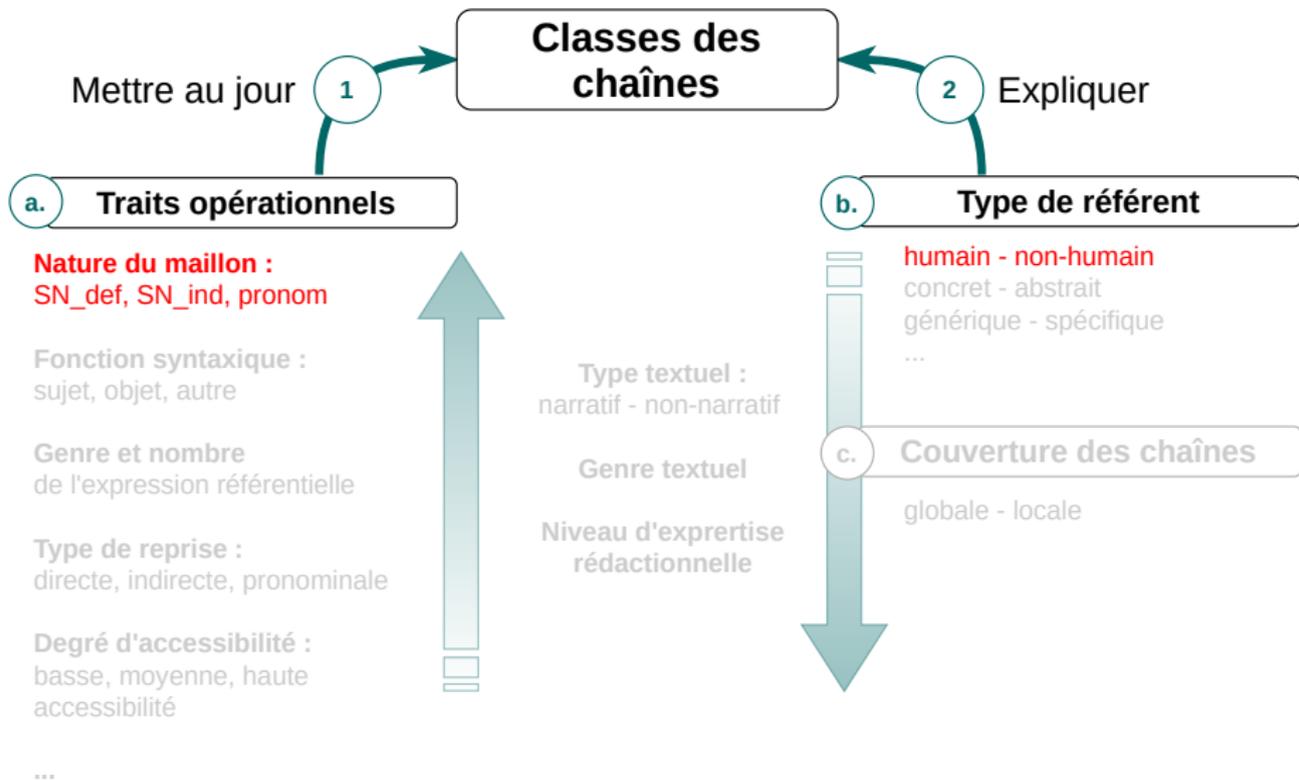
# Application de la méthode



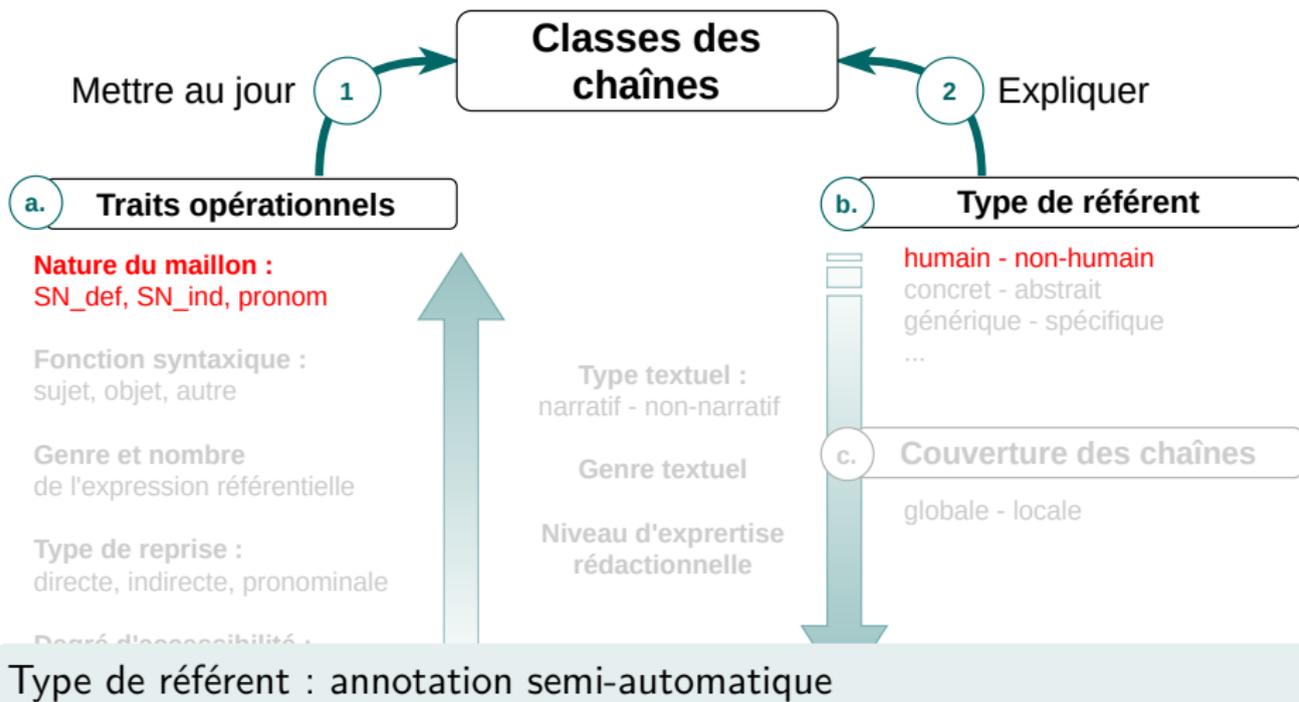
# Méthode de description des chaînes



# Méthode de description des chaînes



# Méthode de description des chaînes



**Méthode : analyse des séquences** (Quiniou, Cellier, Charnois, & Legallois, 2012)

Identifier les régularités, les ressemblances, construire des typologies de « séquences-types » (Robette, 2011)

Méthode : analyse des séquences (Quiniou et al., 2012)

Identifier les régularités, les ressemblances, construire des typologies de « séquences-types » (Robette, 2011) ⇒ Stratégies référentielles

# Sequence Analysis

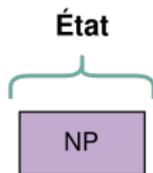
Méthode : analyse des séquences (Quiniou et al., 2012)

Identifier les régularités, les ressemblances, construire des typologies de « séquences-types » (Robette, 2011) ⇒ Stratégies référentielles

**Séquence** : liste ordonnée d'**états** ou d'événements (Brzinsky-Fay et al., 2006)

**Mme. Seton** est-elle aussi connue chez-nous qu'elle le devrait être ?

Je ne le crois pas : et, m'aidant de ses historiens, je vais donner aux lecteurs de la Revue Canadienne une biographie de cette femme, l'une des plus accomplies, plus aimantes et des plus aimées qui aient jamais foulé la terre. Elle a beaucoup écrit;



Séquence



# Sequence Analysis

Méthode : analyse des séquences (Quiniou et al., 2012)

Identifier les régularités, les ressemblances, construire des typologies de « séquences-types » (Robette, 2011) ⇒ Stratégies référentielles

**Séquence** : liste ordonnée d'états ou d'événements (Brzinsky-Fay et al., 2006)

Méthode : analyse des séquences (Quiniou et al., 2012)

Identifier les régularités, les ressemblances, construire des typologies de « séquences-types » (Robette, 2011) ⇒ Stratégies référentielles

Nature maillons	#	%
SN_def	1198	34,66
SN_dem	272	7,87
SN_ind	126	3,65
SN_sansDET	64	1,85
NP	442	12,79
Poss	182	5,27
Pro	1026	29,69
Autre	146	4,22
<b>Total</b>	<b>3456</b>	<b>100</b>

**Trait utilisé** : nature des maillons

- La fréquence des catégories grammaticales : variations les plus fortes
- Informations les plus riches sur la typologie des chaînes : degré d'homogénéité

(Obry, Glikman, Guillot-Barbance, & Pincemin, 2017)

Méthode : analyse des séquences (Quiniou et al., 2012)

Identifier les régularités, les ressemblances, construire des typologies de « séquences-types » (Robette, 2011) ⇒ Stratégies référentielles

**TraMineR** (Gabadinho, Ritschard, Studer, & Müller, 2009) :

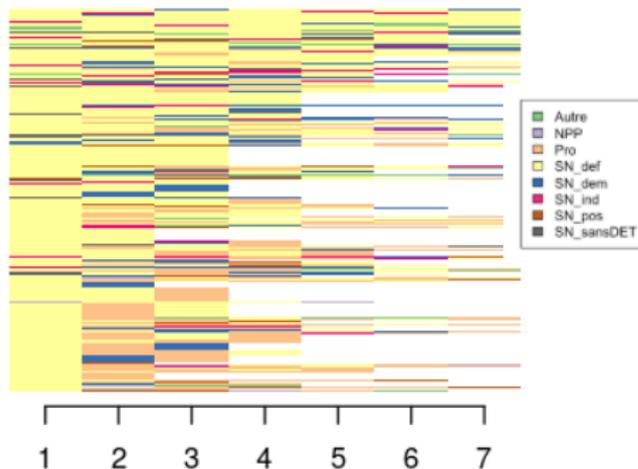
- Visualisation spécifique pour l'analyse des séquences
- Méthodes de classification automatique
- Méthodes statistiques

**Paramètres :**

- ① Taille des séquences  $\leq 7$  mentions (21% ont été coupées)
- ② Distance : optimal matching (basée sur la distance de Levenshtein)
- ③ Clustering : clustering hiérarchique (méthode de 'Ward')
- ④ Nombre de clusters : 3 classes (observation de 2,3, 4 et 5 classes)

# Vers une typologie des chaînes

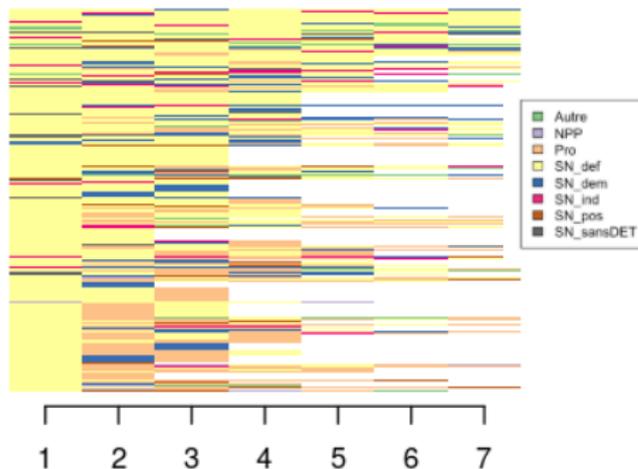
Classe TraMineR 1



*le vignoble champenois s'étendait sur quelques [...] le vignoble connaît [...] et de la Grande guerre, le vignoble s'est réduit à 12 000 hectares. Aujourd'hui, en 2007, le vignoble champenois s'étend sur 32 341 hectares. (encyclopedic text)*

# Vers une typologie des chaînes

Classe TraMineR 1

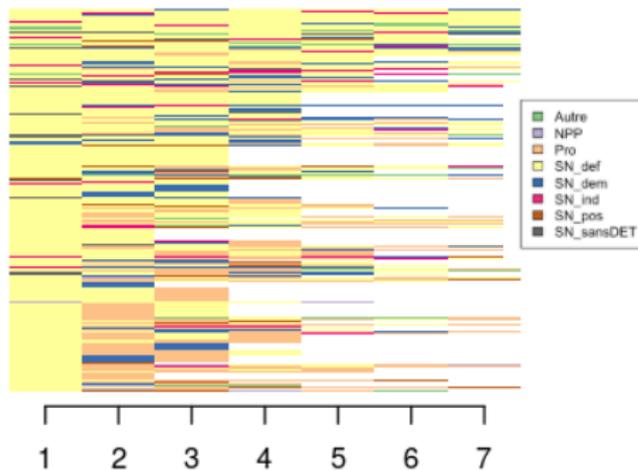


- aperçu global de la composition des chaînes
- chaque couleur correspond à une catégorie grammaticale
- les chaînes sont triées par leur similarité

*le vignoble champenois s'étendait sur quelques [...] le vignoble connaît [...] et de la Grande guerre, le vignoble s'est réduit à 12 000 hectares. Aujourd'hui, en 2007, le vignoble champenois s'étend sur 32 341 hectares. (encyclopedic text)*

# Vers une typologie des chaînes

Classe TraMineR 1

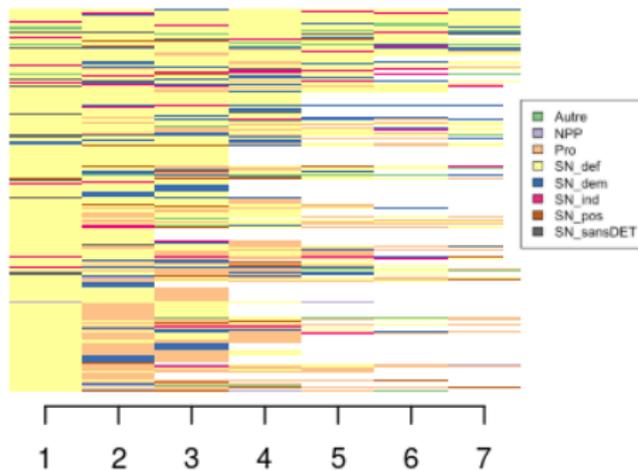


- aperçu global de la composition des chaînes
- chaque couleur correspond à une catégorie grammaticale
- les chaînes sont triées par leur similarité

*le vignoble champenois s'étendait sur quelques [...] le vignoble connaît [...] et de la Grande guerre, le vignoble s'est réduit à 12 000 hectares. Aujourd'hui, en 2007, le vignoble champenois s'étend sur 32 341 hectares. (encyclopedic text)*

# Vers une typologie des chaînes

Classe TraMineR 1

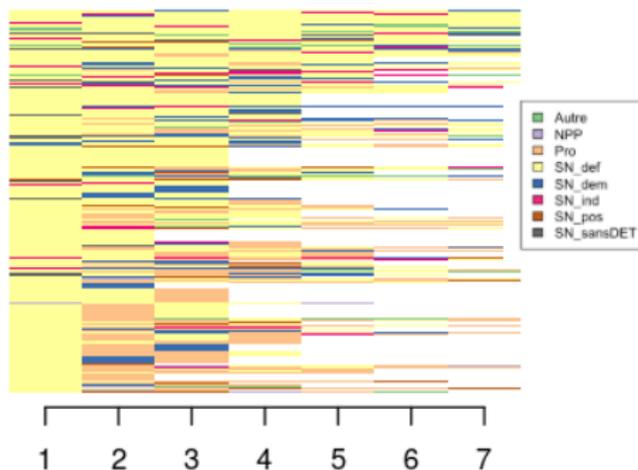


- aperçu global de la composition des chaînes
- chaque couleur correspond à une catégorie grammaticale
- les chaînes sont triées par leur similarité

*le vignoble champenois s'étendait sur quelques [...] le vignoble connaît [...] et de la Grande guerre, le vignoble s'est réduit à 12 000 hectares. Aujourd'hui, en 2007, le vignoble champenois s'étend sur 32 341 hectares. (encyclopedic text)*

# Vers une typologie des chaînes

Classe TraMineR 1

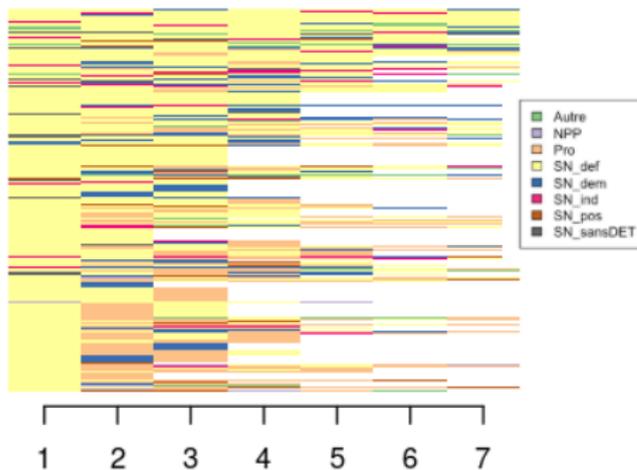


- 371 séquences :
  - 89.6% commence avec un SN\_def ■
  - suivis d'autre SN\_def ■

*le vignoble champenois s'étendait sur quelques [...] le vignoble connaît [...] et de la Grande guerre, le vignoble s'est réduit à 12 000 hectares. Aujourd'hui, en 2007, le vignoble champenois s'étend sur 32 341 hectares. (encyclopedic text)*

# Vers une typologie des chaînes

Classe TraMineR 1

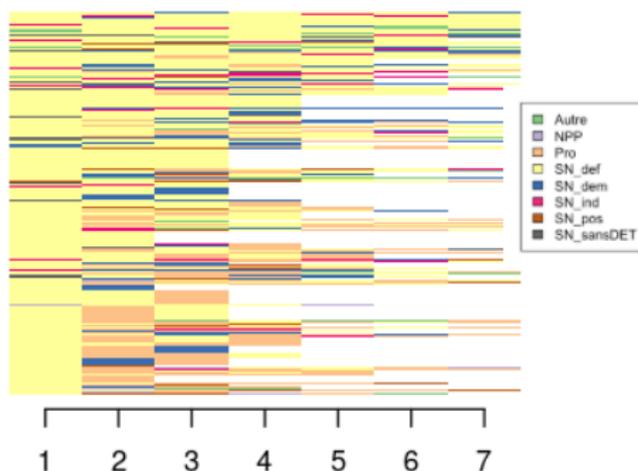


- Stratégies :
  - Introduction et maintien des référents non humains
  - Introduction et maintien des référents humains

*le vignoble champenois s'étendait sur quelques [...] le vignoble connaît [...] et de la Grande guerre, le vignoble s'est réduit à 12 000 hectares. Aujourd'hui, en 2007, le vignoble champenois s'étend sur 32 341 hectares. (encyclopedic text)*

# Vers une typologie des chaînes

Classe TraMineR 1

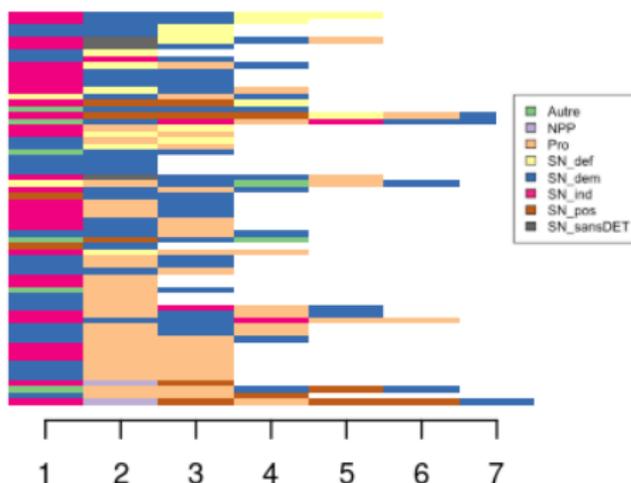


- Stratégies :
  - Introduction et maintien des référents non humains
  - Introduction et maintien des référents humains

*L'Agence de Protection de l'Environnement* était en charge [...] Plusieurs mois après, *l'EPA* enregistrait [...] Le 21 août 2003 cependant, *l'Agence* [...] où *l'Agence* se voit reprocher de [...] *L'EPA* envisagea [...]. (encyclopedic text)

# Vers une typologie des chaînes

Classe TraMineR 2

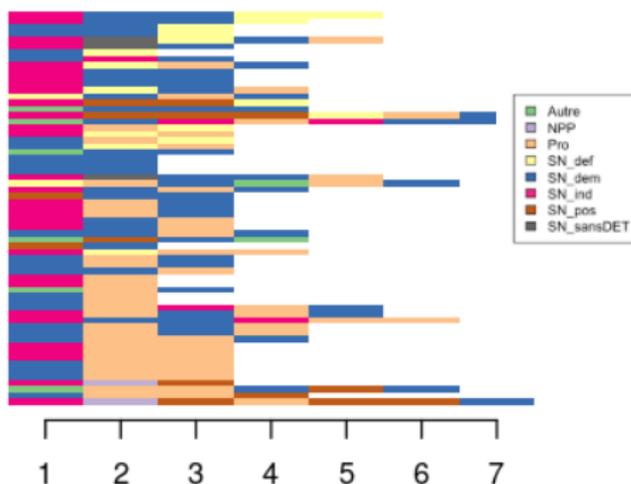


- 63 séquences :
  - 47.6% commence avec un SN\_ind
  - 36.5% commence avec un SN\_dem
  - suivi par d'autres SN\_dem ou pronoms

[...] « le " **communicatif** " présente ainsi fréquemment **une connotation oppositionnelle** (sinon contradictoire) avec le linguistique. **Ceci** est particulièrement crucial lorsqu'on traite de la " **compétence de communication** " [...] **Cette dichotomie** pose problème au psychologue [...]. (scientific article)

# Vers une typologie des chaînes

Classe TraMineR 2



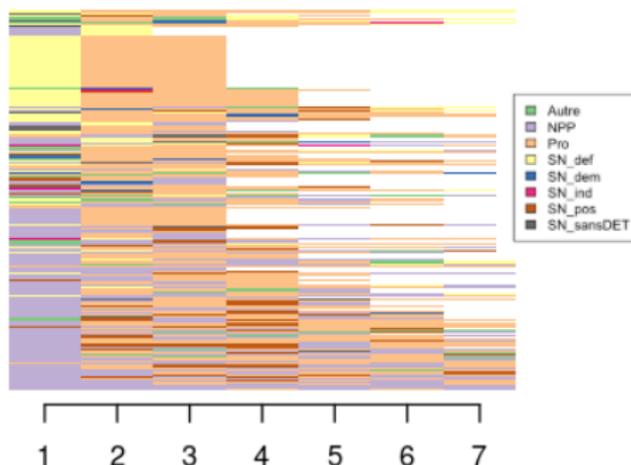
- Stratégies :

- Référencer à des idées ou concepts
- Encapsuler des portions de texte

[...] « le " *communicatif* " présente ainsi fréquemment **une connotation oppositionnelle** (sinon contradictoire) avec le linguistique. **Ceci** est particulièrement crucial lorsqu'on traite de la " *compétence de communication* " [...] **Cette dichotomie** pose problème au psychologue [...]. (scientific article)

# Vers une typologie des chaînes

Classe TraMineR 3

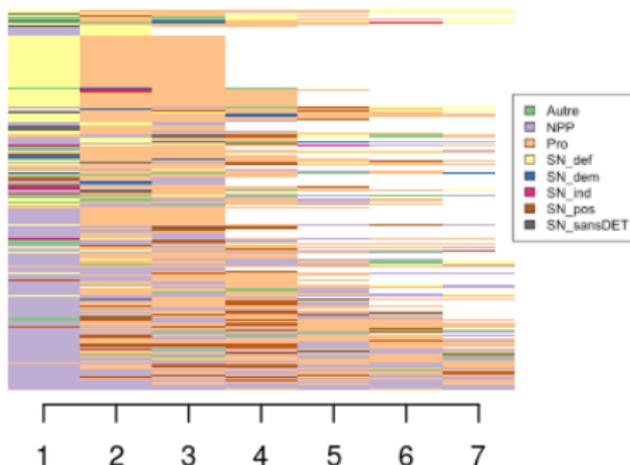


- 201 séquences :
  - 50.7% commence avec un NP ■
  - 28.8% commence avec un SN\_def ■
  - suivi d'autres NP ■ ou pronoms ■

Chez F. de Saussure, l'analogie [...], il pose que les facteurs de trouble [...]. Pour lui, cette tendance à l'irrégularité est heureusement contrebalancée par l'analogie [...]. Comme H. Paul, il ramène le concept au calcul de l'équation de la quatrième proportionnelle. (scientific article)

# Vers une typologie des chaînes

Classe TraMineR 3



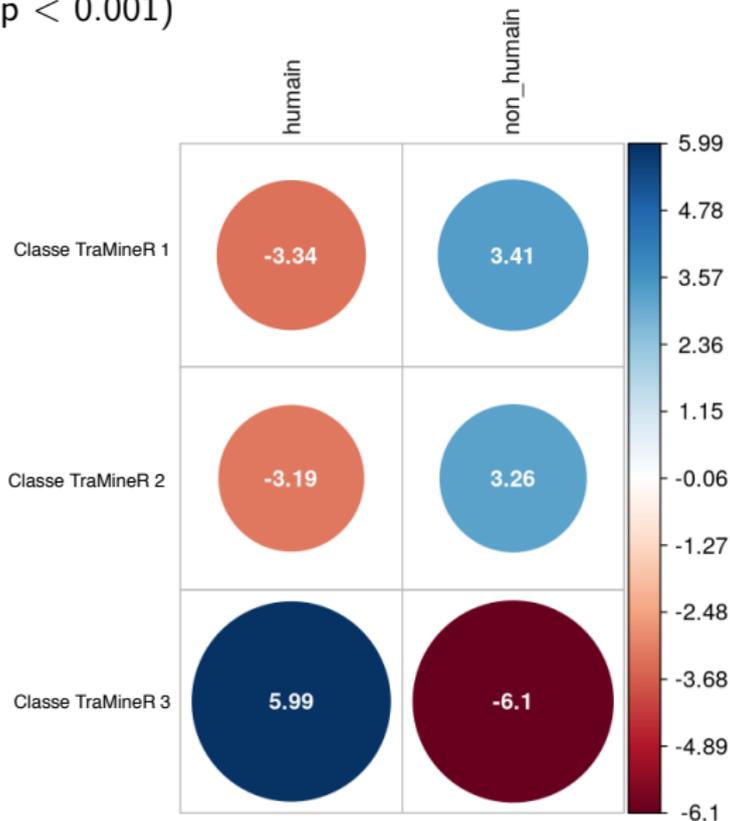
- Stratégies :

- Introduire et maintenir des référents humains

Chez **F. de Saussure** , l'analogie [...], **il** pose que les facteurs de trouble [...]. Pour **lui** , cette tendance à l'irrégularité est heureusement contrebalancée par l'analogie [...]. Comme H. Paul, **il** ramène le concept au calcul de l'équation de la quatrième proportionnelle. (scientific article)

# Projeter le type de référent

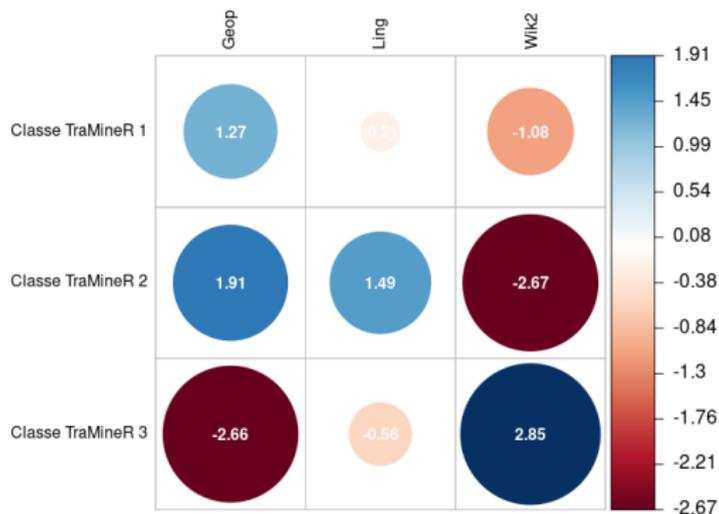
Liaison significative entre les classes et le type de référent  
(test du khi-deux  $p < 0.001$ )



# Projeter le genre textuel

Liaison significative entre les classes et le genre textuel  
(test du khi-deux  $p < 0.001$ )

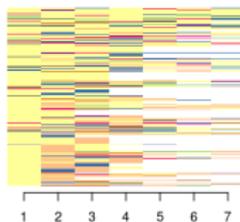
	Geop		Ling		Wiki	
	#	%	#	%	#	%
<b>Humain</b>	<b>136</b>	<b>58,12</b>	22	25,29	<b>138</b>	<b>53,08</b>
<b>Non humain</b>	98	41,88	65	74,71	122	46,92



# Approche par séquences

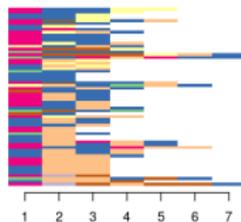
Classe 1

317 seq. (n=317), sorted



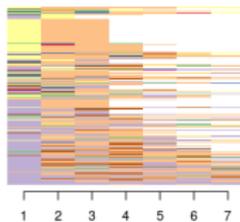
Classe 2

63 seq. (n=63), sorted



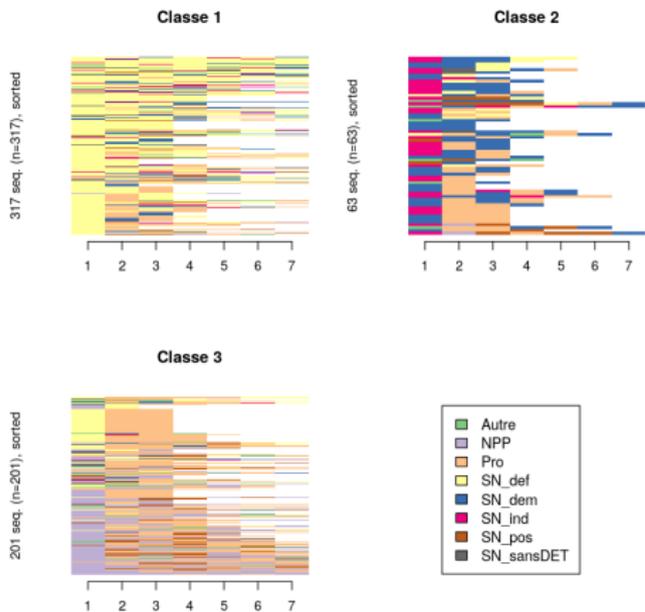
Classe 3

201 seq. (n=201), sorted



- Résultats obtenus directement interprétables

# Approche par séquences

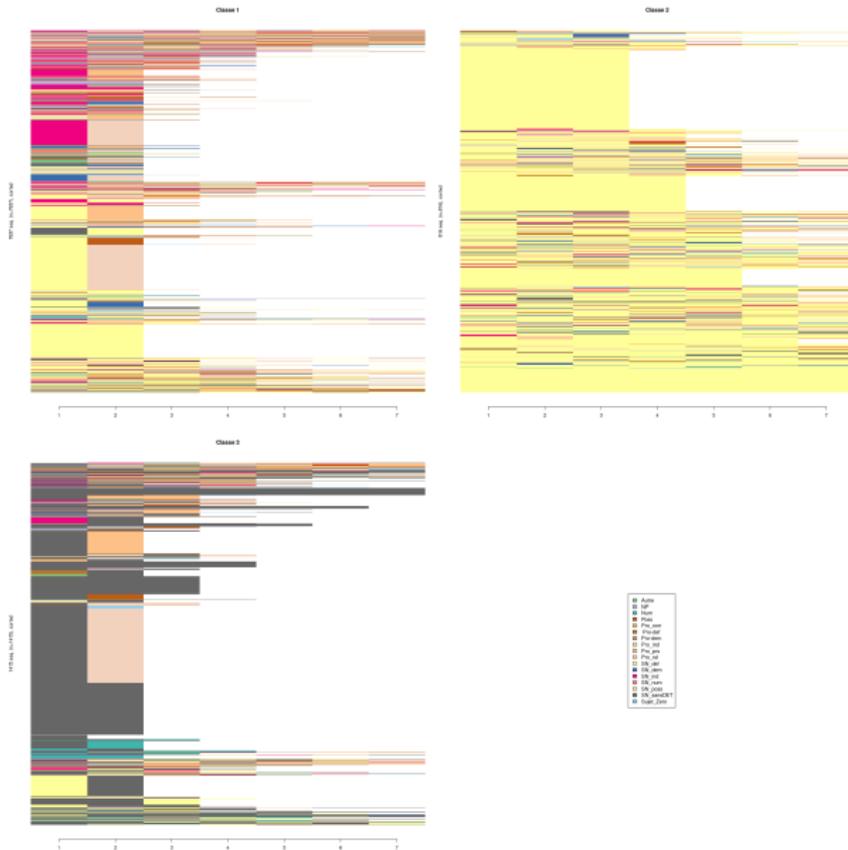


- Résultats obtenus directement interprétables
- Résultats obtenus pertinents d'un point de vu linguistique
- Résultats suggèrent l'existence de 3 stratégies différentes pour introduire et maintenir les référents

- Prise en compte des textes du 19<sup>ème</sup>, 20<sup>ème</sup> et 21<sup>ème</sup> siècle

Textes	Mots	CR	CR=2		CR <sub>≥</sub> 3		Maillons
			#	%	#	%	
29	347 155	9 869	5633	57,1	4236	42,9	59 478

# Première exploration sur Democrat





## Prochaines étapes :

- ① Prendre en compte d'autres traits opérationnels :
  - Fonction syntaxique : sujet, objet, autre
  - Degré d'informativité du maillon
  - Type de reprise : directe, indirecte, pronominale
  - Interdistance (Rousier-Vercruyssen & Landragin, 2019)
  - Instabilité des chaînes (Rousier-Vercruyssen & Landragin, 2019)
  
- ② Combiner plusieurs traits opérationnels
  
- ③ Projeter d'autres types de référent :
  - PERS vs LOC vs ORG vs MISC ⇒ comment traiter les entités nomées ?
  - concret - abstrait, générique - spécifique
  - distinctions plus fines : individus particuliers - humain collectifs
  
- Faire varier les paramètres pour le clustering
  
- Compléter les analyses avec une description globale

## Prochaines étapes :

- ① Prendre en compte d'autres traits opérationnels :
  - Fonction syntaxique : sujet, objet, autre
  - Degré d'informativité du maillon
  - Type de reprise : directe, indirecte, pronominale
  - Interdistance (Rousier-Vercauysen & Landragin, 2019)
  - Instabilité des chaînes (Rousier-Vercauysen & Landragin, 2019)
  
- ② Combiner plusieurs traits opérationnels
  
- ③ Projeter d'autres types de référent :
  - PERS vs LOC vs ORG vs MISC ⇒ comment traiter les entités nomées ?
  - concret - abstrait, générique - spécifique
  - distinctions plus fines : individus particuliers - humain collectifs
  
- Faire varier les paramètres pour le clustering
- Compléter les analyses avec une description globale

## Prochaines étapes :

- ① Prendre en compte d'autres traits opérationnels :
  - Fonction syntaxique : sujet, objet, autre
  - Degré d'informativité du maillon
  - Type de reprise : directe, indirecte, pronominale
  - Interdistance (Rousier-Vercruyssen & Landragin, 2019)
  - Instabilité des chaînes (Rousier-Vercruyssen & Landragin, 2019)
  
- ② Combiner plusieurs traits opérationnels
  
- ③ Projeter d'autres types de référent :
  - PERS vs LOC vs ORG vs MISC ⇒ comment traiter les entités nomées ?
  - concret - abstrait, générique - spécifique
  - distinctions plus fines : individus particuliers - humain collectifs
  - Faire varier les paramètres pour le clustering
  - Compléter les analyses avec une description globale

## Prochaines étapes :

- ① Prendre en compte d'autres traits opérationnels :
  - Fonction syntaxique : sujet, objet, autre
  - Degré d'informativité du maillon
  - Type de reprise : directe, indirecte, pronominale
  - Interdistance (Rousier-Vercreuysen & Landragin, 2019)
  - Instabilité des chaînes (Rousier-Vercreuysen & Landragin, 2019)
  
- ② Combiner plusieurs traits opérationnels
  
- ③ Projeter d'autres types de référent :
  - PERS vs LOC vs ORG vs MISC ⇒ comment traiter les entités nomées ?
  - concret - abstrait, générique - spécifique
  - distinctions plus fines : individus particuliers - humain collectifs
  
- Faire varier les paramètres pour le clustering
  
- Compléter les analyses avec une description globale

## Prochaines étapes :

- ① Prendre en compte d'autres traits opérationnels :
  - Fonction syntaxique : sujet, objet, autre
  - Degré d'informativité du maillon
  - Type de reprise : directe, indirecte, pronominale
  - Interdistance (Rousier-Veracruz & Landragin, 2019)
  - Instabilité des chaînes (Rousier-Veracruz & Landragin, 2019)
  
- ② Combiner plusieurs traits opérationnels
  
- ③ Projeter d'autres types de référent :
  - PERS vs LOC vs ORG vs MISC ⇒ comment traiter les entités nomées ?
  - concret - abstrait, générique - spécifique
  - distinctions plus fines : individus particuliers - humain collectifs
  
- Faire varier les paramètres pour le clustering
  
- Compléter les analyses avec une description globale

- Ariel, M. (2001). Accessibility theory : An overview. *Text representation : Linguistic and psycholinguistic aspects*, 8, 29–87.
- Brzinsky-Fay, C., Kohler, U., & Luniak, M. (2006). Sequence analysis with Stata. *The Stata Journal*, 6(4), 435–460.
- Charolles, M. (2002). *La référence et les expressions référentielles en français*. Editions Ophrys.
- Corblin, F. (1995). *Les formes de reprise dans le discours. anaphores et chaînes de référence*. Presses Universitaires de Rennes.
- Cornish, F. (2000). L'accessibilité cognitive des référents, le centrage d'attention, et la structuration du discours : une vue d'ensemble. *Verbum*, 22(1), 7–30.
- De Marneffe, M.-C., Recasens, M., & Potts, C. (2015). Modeling the lifespan of discourse entities with application to coreference resolution. *Journal of Artificial Intelligence Research*, 52, 445–475.
- Fossard, M., Achim, A. M., Rousier-Vercruyssen, L., Gonzalez, S., Bureau, A., & Champagne-Lavau, M. (2018). Referential choices in a collaborative storytelling task : Discourse stages and referential complexity matter. *Frontiers in psychology*, 9, 176.
- Gabardinho, A., Ritschard, G., Studer, M., & Müller, N. S. (2009). Mining

sequence data in R with the TraMineR package : A user's guide.  
*Geneva : Department of Econometrics and Laboratory of  
Demography, University of Geneva.*

- Grishina, Y., & Stede, M. (2015). Knowledge-lean projection of coreference chains across languages. In *Proceedings of the eighth workshop on building and using comparable corpora*. Beijing, China.
- Gundel, J. K., Hedberg, N., & Zacharski, R. (2019). Cognitive status and the form of referring expressions in discourse. In J. Gundel & B. Abbott (Eds.), *The oxford handbook of reference* (pp. 67–99). New York : Oxford University Press.
- Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in english*. Routledge.
- Kaiser, E., & Fedele, E. (2019). Reference resolution : A psycholinguistic perspective. In J. Gundel & B. Abbott (Eds.), *The oxford handbook of reference* (pp. 309–336). New York : Oxford University Press.
- Kleiber, G. (1994). Anaphores et pronoms. *Louvain-la-Neuve, Duculot*.
- Kleiber, G. (2002). Marqueurs référentiels et théorie du centrage. *Linx. Revue des linguistes de l'université Paris X Nanterre*(47), 107–119.
- Kunz, K., Lapshinova-Koltunski, E., & Martinez, J. M. (2016). Beyond identity coreference : Contrasting indicators of textual coherence in

- English and German. In *Proceedings of the workshop on coreference resolution beyond ontonotes (corbon 2016)* (pp. 23–31).
- Landragin, F. (2018, juillet). Étude de la référence et de la coréférence : rôle des petits corpus et observations à partir du corpus mc4. *Corpus*, 18.
- Longo, L. (2013). *Vers des moteurs de recherche "intelligents" : un outil de détection automatique de thèmes*. (PhD Thesis). Université de Strasbourg.
- Longo, L., & Todirascu, A. (2014). Vers une typologie des chaînes de référence dans des textes administratifs et juridiques. *Langages*(195/3), 79-98.
- Mitkov, R. (2014). *Anaphora resolution*. Routledge.
- Obry, V., Glikman, J., Guillot-Barbance, C., & Pincemin, B. (2017). Les chaînes de référence dans les récits brefs en français : étude diachronique (xiii<sup>e</sup>-xv<sup>e</sup> s.). *Langue française*, 3, 91–110.
- Poesio, M., Pradhan, S., Recasens, M., Rodriguez, K., & Versley, Y. (2016). Annotated corpora and annotation tools. In M. Poesio, R. Stuckardt, & Y. Versley (Eds.), *Anaphora resolution : Algorithms, resources, and applications* (pp. 97–140). Berlin, Heidelberg :

Springer Berlin Heidelberg. Consulté sur

[https://doi.org/10.1007/978-3-662-47909-4\\_4](https://doi.org/10.1007/978-3-662-47909-4_4) doi:  
10.1007/978-3-662-47909-4\_4

- Quiniou, S., Cellier, P., Charnois, T., & Legallois, D. (2012, juin). Fouille de données pour la stylistique : cas des motifs séquentiels émergents. *Journées Internationales d'Analyse Statistique des Données Textuelles (JADT'12)*, 821–833. Consulté sur <https://hal.archives-ouvertes.fr/hal-00675586>
- Recasens, M., & Hovy, E. (2010, juillet). Coreference resolution across corpora : Languages, coding schemes, and preprocessing information. In *Proceedings of the 48th annual meeting of the association for computational linguistics* (pp. 1423–1432). Uppsala, Sweden : Association for Computational Linguistics. Consulté sur <https://aclanthology.org/P10-1144>
- Roberts, C. (2019). Contextual influences on reference. In J. Gundel & B. Abbott (Eds.), *The oxford handbook of reference* (pp. 260–280). New York : Oxford University Press.
- Robette, N. (2011). *Explorer et décrire les parcours de vie : les typologies de trajectoires*. CEPED. Consulté sur

- Rousier-Vercruyssen, L., & Landragin, F. (2019). Interdistance et instabilité au sein des chaînes de référence : indices textuels ? *Discours. Revue de linguistique, psycholinguistique et informatique. A journal of linguistics, psycholinguistics and computational linguistics*, 25.
- Salazar Orvig, A. (2019). Reference and referring expressions in first language acquisition. In J. Gundel & B. Abbott (Eds.), *The oxford handbook of reference* (pp. 283–308). New York : Oxford University Press.
- Schnedecker, C. (2005). Les chaînes de référence dans les portraits journalistiques : éléments de description. *Travaux de linguistique*(2), 85–133.
- Schnedecker, C. (2021). *Les chaînes de référence en français*. Paris : Éditions Ophrys.
- Schnedecker, C., & Longo, L. (2012, juillet). Impact des genres sur la composition des chaînes de référence : le cas des faits divers. In *3ième congrès mondial de linguistique française* (p. 1957-1972). Lyon, France.

- Vogels, J., Krahmer, E., & Maes, A. (2013). When a stone tries to climb up a slope : the interplay between lexical and perceptual animacy in referential choices. *Frontiers in psychology*, 4, 154.
- Walker, M. A., Joshi, A. K., & Prince, E. F. (1998). *Centering theory in discourse*. Oxford University Press.
- Wilkins, R., Oberle, B., Landragin, F., & Todirascu, A. (2020). French coreference for spoken and written language. In *Language resources and evaluation conference (Irec 2020)* (pp. 80–89). Marseille, France.