

Chaînes et maillons dans AnnoDis : référents humains versus référents non humains

Silvia Federzoni
CLLE (UMR 5263)

Journée d'étude
Annotation de la coréférence

Organisée par :

Lydia-Mai Ho-Dac, Josette Rebeyrolle, Claudine Garcia-Debanc et Marie-Paule Jacques



19 novembre 2020

① Contexte

Objectif

Méthode

② Chaînes et maillons dans la ressource AnnoDis

La ressource AnnoDis

Vers une typologie des chaînes

Méthode : fouille des séquences

Résultats

③ Perspectives

① Contexte

Objectif

Méthode

② Chaînes et maillons dans la ressource AnnoDis

La ressource AnnoDis

Vers une typologie des chaînes

Méthode : fouille des séquences

Résultats

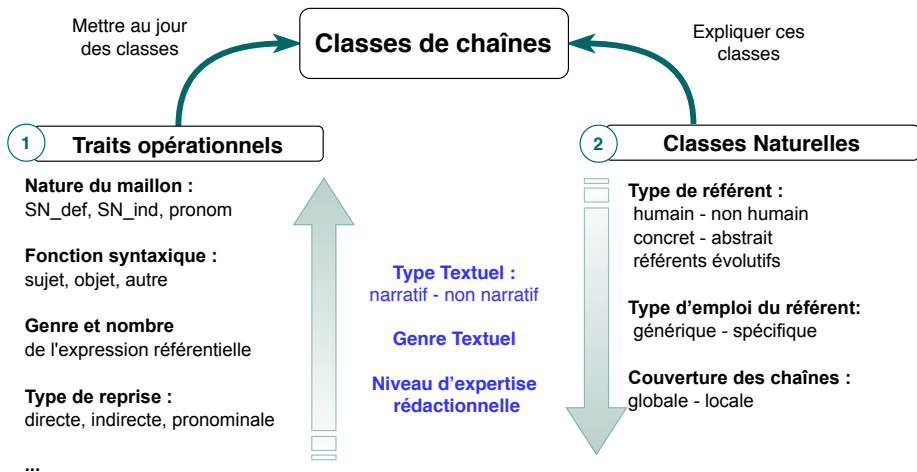
③ Perspectives

Typologie des chaînes de coréférence

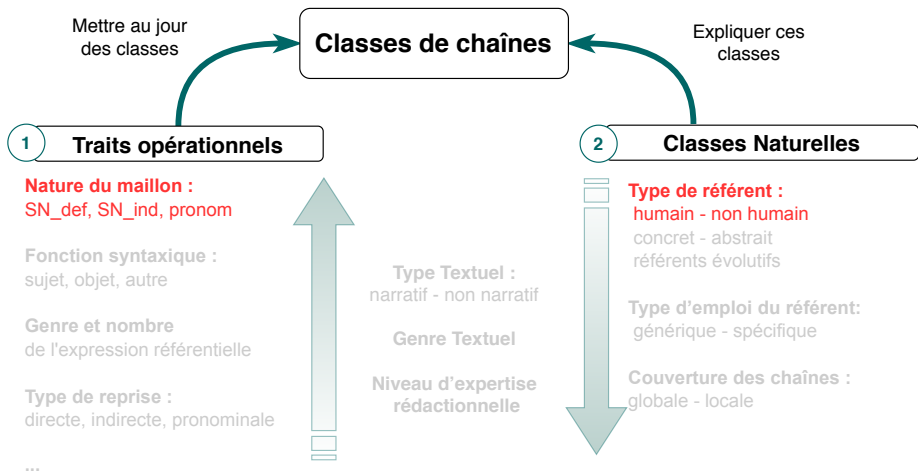
Mettre au jour des **types** de chaînes en fonction de leur **rôle discursif**.

Proposer une description **systematique** et **exhaustive** de la **composition**, de la **variété** et de la **complexité** des chaînes de référence à la lumière de **ressources diversifiées** :

- corpus de **grande taille**
- différents **types** et **genres textuels**
- différents **niveaux d'expertise rédactionnelle**
- plusieurs **types de référents**



Méthode de description



① Contexte

Objectif

Méthode

② Chaînes et maillons dans la ressource AnnoDis

La ressource AnnoDis

Vers une typologie des chaînes

Méthode : fouille des séquences

Résultats

③ Perspectives

- Ressource annotée en **chaînes topicales** \Rightarrow assimilées aux chaînes de référence (Federzoni, Ho-Dac, & Rebeyrolle, 2020)
 - textes longs structurés
 - entièrement annotés
 - non narratifs (expositifs)
 - 3 genres textuels (rapports, articles scientifiques, textes encyclopédiques)

Sous-corpus	Nb de mots	Nb CR	Nb maillons	Long. moyenne CR
GEOP	266 000	234	1125	4,81
LING	169 000	87	478	5,49
WIK2	231 000	260	1853	7,13
AnnoDis	666 000	581	3456	5,95

Vers une typologie des chaînes

Objectif : décrire la composition des chaînes

Mettre au jour des classes de chaînes de référence en observant les enchaînements les plus fréquents.

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

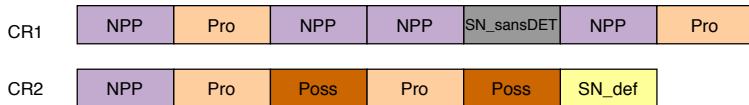
Trait utilisé : nature des maillons

- 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)

Enchaînements : description de la composition des chaînes

- aller au delà d'une description globale



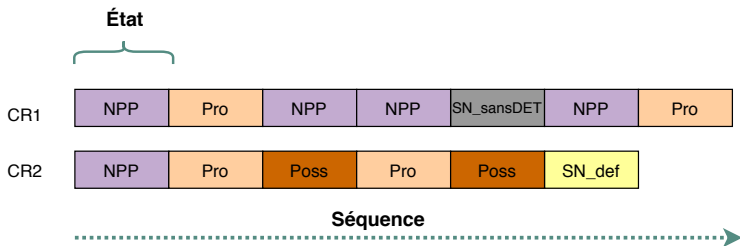
Vers une typologie des chaînes

Enchaînements : description de la composition des chaînes

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

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

Séquence : liste ordonnée d'**états** ou d'événements (Brzinsky-Fay, Kohler, & Luniak, 2006)



Considérer l'étude des chaînes de référence comme une analyse de séquences

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

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

Décisions à prendre :

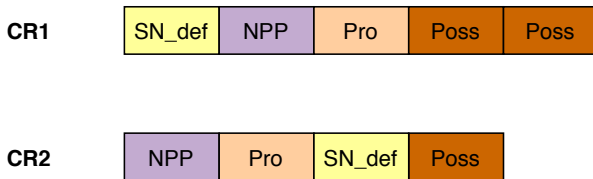
- ① Méthode de classification
- ② Taille des séquences
- ③ Nombre de classes

Décision 1 - méthode de classification

Classification : regrouper les chaînes voisines \Rightarrow **distance faible entre les séquences**

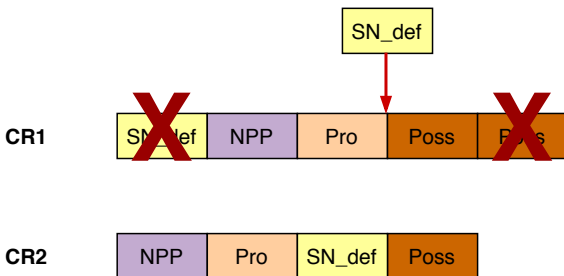
Distance entre les séquences : calculée au moyen d'opérations de transformation (l'insertion, la suppression ou la substitution)

Appariement optimal : trouver pour chaque couple de séquences comment transformer l'une en l'autre le plus facilement possible \Rightarrow **coût minimum**



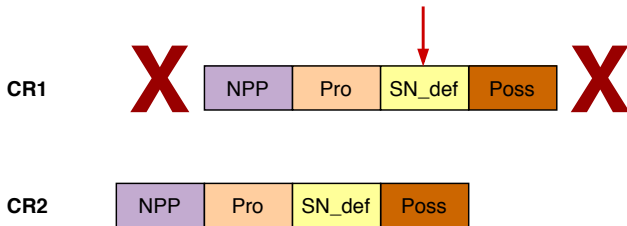
Décision 1 - méthode de classification

Distance entre les séquences : calculée au moyen d'opérations de transformation (l'insertion, la suppression ou la substitution)



Décision 1 - méthode de classification

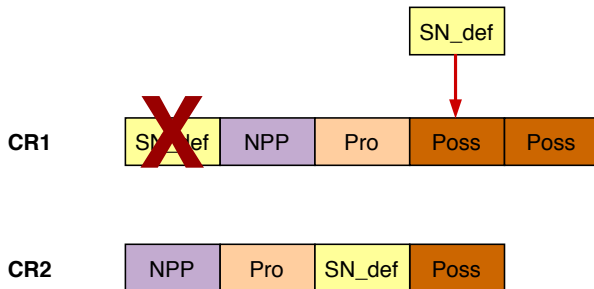
Distance entre les séquences : calculée au moyen d'opérations de transformation (l'insertion, la suppression ou la substitution)



Coût = 3 : la somme de deux suppressions plus une insertion

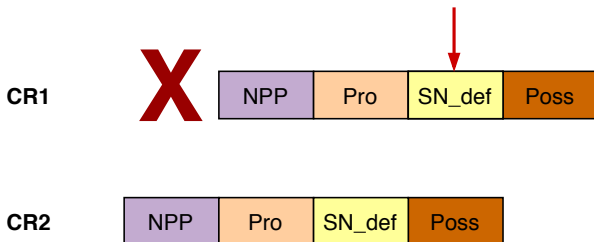
Décision 1 - méthode de classification

Distance entre les séquences : calculée au moyen d'opérations de transformation (l'insertion, la suppression ou la substitution)



Décision 1 - méthode de classification

Distance entre les séquences : calculée au moyen d'opérations de transformation (l'insertion, la suppression ou la substitution)

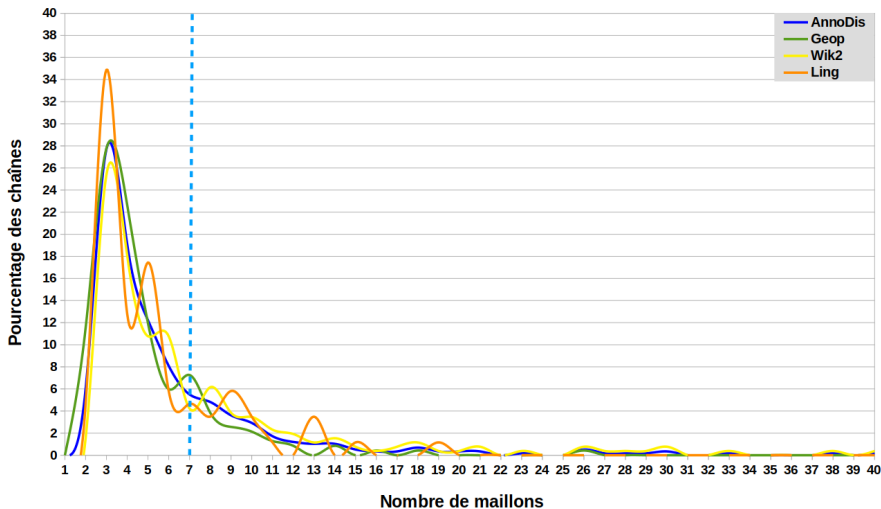


Coût = 2 : une suppression plus une substitution

Décision 2 - taille des séquences

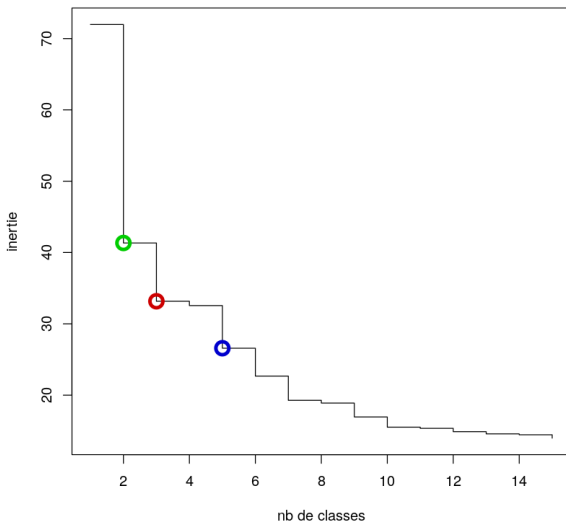
Constat : 73,15% moins de 7 maillons

Décision ⇒ considérer les 7 premiers maillons



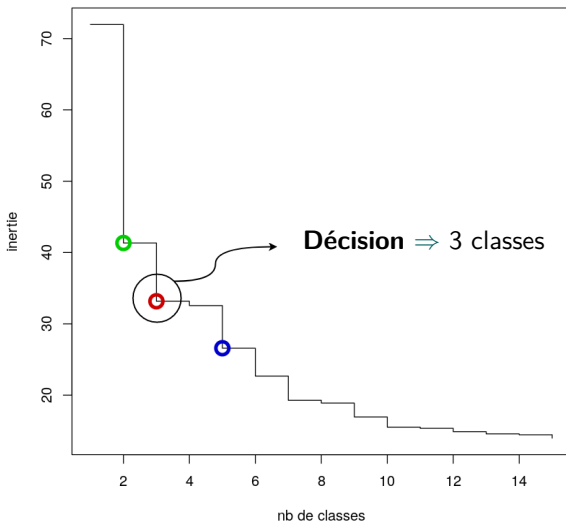
Décision 3 - nombre de classes

Établir le nombre de classes : sauts d'inertie (2,3,5) et observations des classes

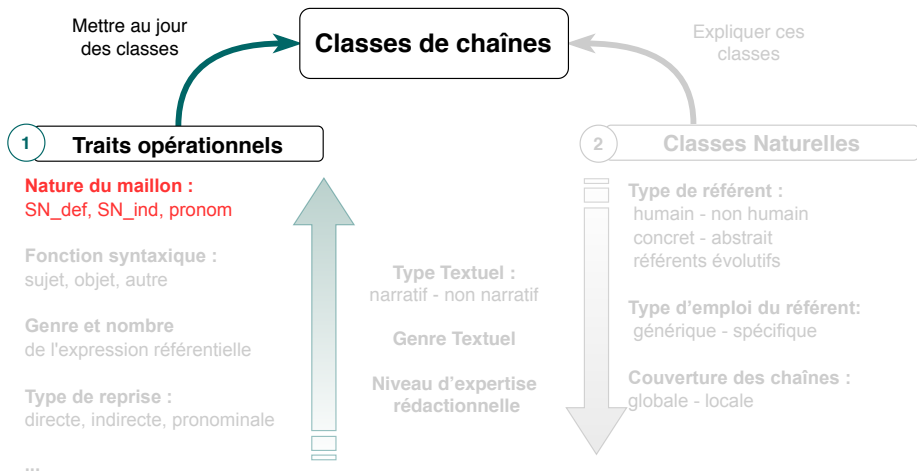


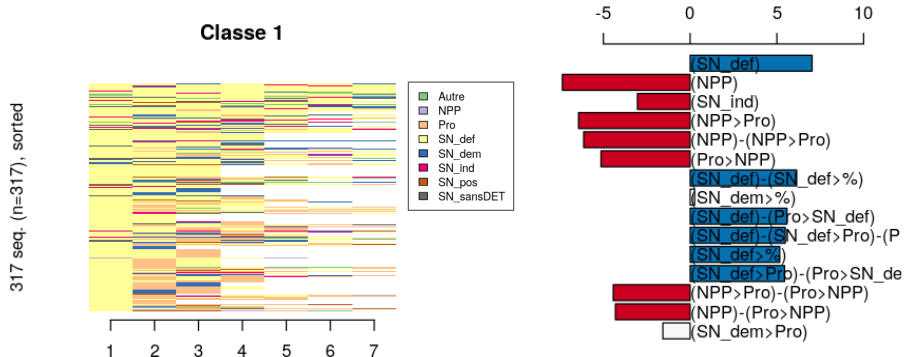
Décision 3 - nombre de classes

Établir le nombre de classes : sauts d'inertie (2,3,5) et observations des classes

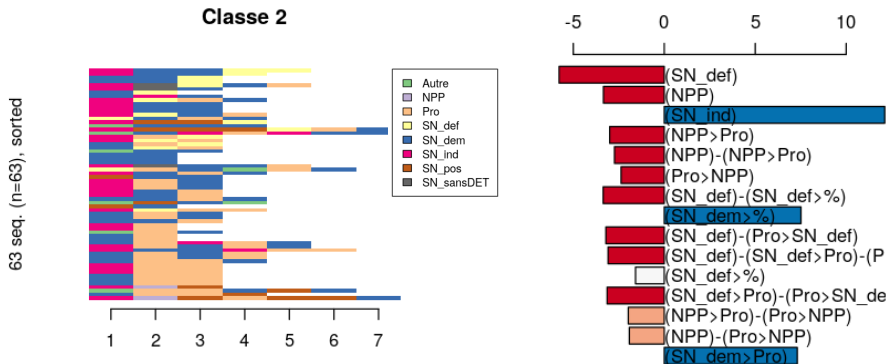


Mettre au jour des classes





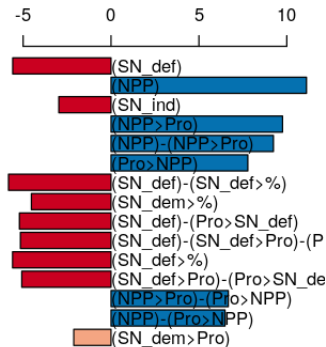
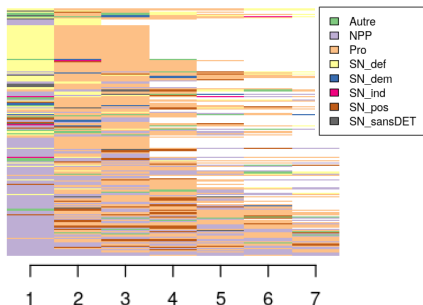
le vignoble champenois s'étendait sur quelques [...] le vignoble connaît [...] Après les fléau du phylloxéra et de la Grande guerre, le vignoble s'est réduit à 12 000 hectares. Aujourd'hui, en 2007, le vignoble champenois s'étend sur 32 341 hectares.



[...] « 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 [...].

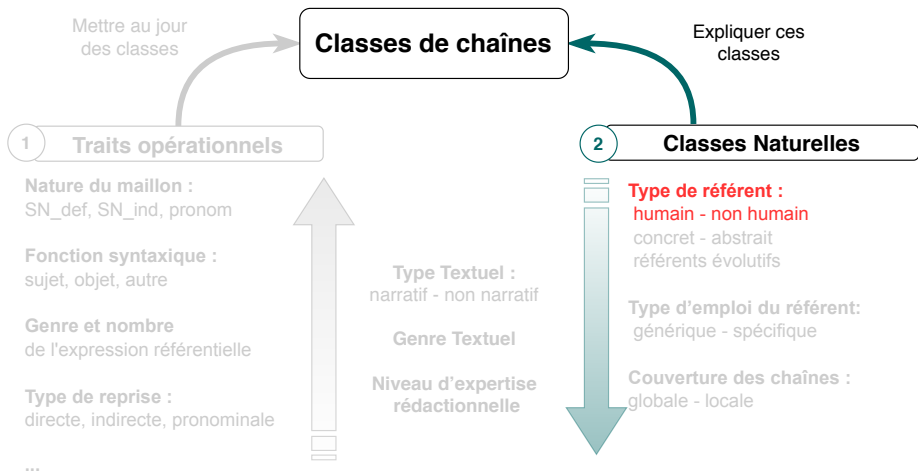
Classe 3

201 seq. (n=201), sorted



Chez **F. de Saussure**, l'analogie [...], **il** pose que les facteurs de trouble [...]. Pour **lui**, cette tendance à l'irrégularité est heureusement contre-balanç e par l'analogie [...]. Comme H. Paul, **il** ram ne le concept au calcul de l' quation de la quatri me proportionnelle.

Expliquer les Classes TraMineR



Projeter les classes naturelles

Conclusion : liaison significative entre les classes et le type de référent (test du khi-deux)



Analyser les classes TraMineR

	Composition	Humain		Non humain	
		#	%	#	%
Classe TraMineR 1	SN_def++	119	37,54	198	62,46
Classe TraMineR 2	SN_ind SN_dem	14	22,22	49	77,78
Classe TraMineR 3	NPP Pro	163	81,09	38	18,91

	Composition	Humain		Non humain	
		#	%	#	%
Classe TraMineR 1	SN_def++	119	37,54	198	62,46
Classe TraMineR 2	SN_ind SN_dem	14	22,22	49	77,78
Classe TraMineR 3	NPP Pro	163	81,09	38	18,91

Référents non humains

le vignoble champenois
s'étendait sur quelques
[...] *le vignoble* connaît
[...] Après les fléau du
phylloxéra 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.

Référents humains

Le Japon connaît une pé-
riode d'avant guerre très dif-
férente des démocraties [...]
La guerre avec les Etats-
Unis sera largement pro-
voquée par l'embargo dé-
cidé par ce pays sur les
exportations pétrolières vers
le Japon. *Le pays* sortira
ruiné par la guerre.

Analyser les classes TraMineR

	Composition	Humain		Non humain	
		#	%	#	%
Classe TraMineR 1	SN_def++	119	37,54	198	62,46
Classe TraMineR 2	SN_ind SN_dem	14	22,22	49	77,78
Classe TraMineR 3	NPP Pro	163	81,09	38	18,91

Référents humains

Le prisonnier n'est en rien au courant des événements qui se déroulent à des milliers de kilomètres de lui. Ni des complots ourdis pour que jamais il ne puisse revenir, [...]. À la fin de l'année 1898, il apprend avec stupéfaction la dimension réelle de l'Affaire, dont il ne sait rien [...].

Référents non humains

[...] le navire, nommé Titan, [...] est présenté comme insubmersible grâce à ses 19 compartiments étanches. De fait, il ne dispose que du nombre minimum de canots de sauvetage requis par la loi. Il heurte un iceberg, coule et la majorité des passagers sont victimes du naufrage.

- Méthode proposée valide pour mettre au jour des classes de chaînes

① **Classes naturelles plus fines** (Longo & Todirascu, 2014)

référents humains : individus particuliers

référents humains : humains collectifs

② **Classes TraMineR**

comprendre les erreurs

affiner la classification

Classe TraMineR 1

Le colonel Picquart est [...] Il est même ministre de la Guerre de 1906 à 1909 dans le premier gouvernement Clemenceau. Il meurt en 1914 d'un accident de cheval.

Classe TraMineR 3

Le prisonnier n'est en rien au courant des événements [...] lui. Ni des complots ourdis pour que jamais il ne puisse revenir, [...]. il apprend avec stupéfaction [...] il ne sait rien [...].

- Méthode proposée valide pour mettre au jour des classes de chaînes
 - ① **Classes naturelles plus fines** (Longo & Todirascu, 2014)
 - référents humains : individus particuliers
 - référents humains : humains collectifs
 - ② **Classes TraMineR**
 - comprendre les erreurs
 - affiner la classification

Classe TraMineR 1

Le colonel Picquart est [...] Il est même ministre de la Guerre de 1906 à 1909 dans le premier gouvernement Clemenceau. Il meurt en 1914 d'un accident de cheval.

Classe TraMineR 3

Le prisonnier n'est en rien au courant des événements [...] lui. Ni des complots ourdis pour que jamais il ne puisse revenir, [...]. il apprend avec stupéfaction [...] il ne sait rien [...].

① Contexte

Objectif

Méthode

② Chaînes et maillons dans la ressource AnnoDis

La ressource AnnoDis

Vers une typologie des chaînes

Méthode : fouille des séquences

Résultats

③ Perspectives

- ① Prendre en compte d'autres traits opérationnels :
 - Fonction syntaxique : sujet, objet, autre
 - Genre et nombre
 - Degré d'informativité du maillon
 - Type de reprise : directe, indirecte, pronominale
 - Interdistance (Rousier-Vercruyssen & Landragin, 2019) : distance en nombre de mots
 - Distance en nombre de phrases
 - Distance en nombre de paragraphes
- ② Combiner plusieurs traits opérationnels
- ③ Projeter d'autres classes naturelles :
 - type de référent : concret - abstrait
 - type d'emploi du référent : générique - spécifique
- ④ Reproduire la même étude sur Democrat (Landragin, 2015) et E-Calm (Garcia-Debanc, Ho-Dac, Bras, & Rebeyrolle, 2017) :
 - type textuel
 - niveaux d'expertise rédactionnelle

- ① Prendre en compte d'autres traits opérationnels :
 - Fonction syntaxique : sujet, objet, autre
 - Genre et nombre
 - Degré d'informativité du maillon
 - Type de reprise : directe, indirecte, pronominale
 - Interdistance (Rousier-Vercruyssen & Landragin, 2019) : distance en nombre de mots
 - Distance en nombre de phrases
 - Distance en nombre de paragraphes
- ② Combiner plusieurs traits opérationnels
- ③ Projeter d'autres classes naturelles :
 - type de référent : concret - abstrait
 - type d'emploi du référent : générique - spécifique
- ④ Reproduire la même étude sur Democrat (Landragin, 2015) et E-Calm (Garcia-Debanc et al., 2017) :
 - type textuel
 - niveaux d'expertise rédactionnelle

- ① Prendre en compte d'autres traits opérationnels :
 - Fonction syntaxique : sujet, objet, autre
 - Genre et nombre
 - Degré d'informativité du maillon
 - Type de reprise : directe, indirecte, pronominale
 - Interdistance (Rousier-Vercreuysen & Landragin, 2019) : distance en nombre de mots
 - Distance en nombre de phrases
 - Distance en nombre de paragraphes
- ② Combiner plusieurs traits opérationnels
- ③ Projeter d'autres classes naturelles :
 - type de référent : concret - abstrait
 - type d'emploi du référent : générique - spécifique
- ④ Reproduire la même étude sur Democrat (Landragin, 2015) et E-Calm (Garcia-Debanc et al., 2017) :
 - type textuel
 - niveaux d'expertise rédactionnelle

- ① Prendre en compte d'autres traits opérationnels :
 - Fonction syntaxique : sujet, objet, autre
 - Genre et nombre
 - Degré d'informativité du maillon
 - Type de reprise : directe, indirecte, pronominale
 - Interdistance (Rousier-Vercruyssen & Landragin, 2019) : distance en nombre de mots
 - Distance en nombre de phrases
 - Distance en nombre de paragraphes
- ② Combiner plusieurs traits opérationnels
- ③ Projeter d'autres classes naturelles :
 - type de référent : concret - abstrait
 - type d'emploi du référent : générique - spécifique
- ④ Reproduire la même étude sur Democrat (Landragin, 2015) et E-Calm (Garcia-Debanc et al., 2017) :
 - type textuel
 - niveaux d'expertise rédactionnelle

- Brzinsky-Fay, C., Kohler, U., & Luniak, M. (2006). Sequence analysis with stata. *The Stata Journal*, 6(4), 435–460.
- Federzoni, S., Ho-Dac, L.-M., & Rebeyrolle, J. (2020, juillet). Les chaînes topicales dans la ressource ANNODIS. In *CMLF2020 : 7e Congrès Mondial de Linguistique Française*. Montpellier, France. Consulté sur <https://hal.archives-ouvertes.fr/hal-02890989>
- Gabadinho, 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*.
- Garcia-Debanc, C., Ho-Dac, L.-M., Bras, M., & Rebeyrolle, J. (2017). Vers l'annotation discursive de textes d'élèves. *Corpus*(16).
- Landragin, F. (2015). Description, modélisation et détection automatique des chaînes de référence (democrat). *Bulletin de l'AFIA*(92), 11–15.
- Longo, L., & Todirascu, A. (2014). Vers une typologie des chaînes de référence dans des textes administratifs et juridiques. *Langages*(195/3), 79-98.
- Obry, V., Glikman, J., Guillot-Barbance, C., & Pincemin, B. (2017). Les chaînes de référence dans les récits brefs en français : étude

- diachronique (xiii^e-xv^e s.). *Langue française*(3), 91–110.
- Péry-Woodley, M.-P., Afantenos, S., Ho-Dac, L.-M., & Asher, N. (2011). La ressource annodis, un corpus enrichi d'annotations discursives. *Traitement Automatique des Langues*, 52(3), 71-101.
- 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>
- Robette, N. (2011). *Explorer et décrire les parcours de vie : les typologies de trajectoires*. CEPED. Consulté sur <https://halshs.archives-ouvertes.fr/halshs-01016125>
- 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).

Autres références

- Antoine, J.-Y., Lefeuve, A., & Schang, E. (2016, juillet). Codage enchaîné ou en première mention de la coréférence : approcher la structure des chaînes de référence par comparaison des deux annotations . In 5ème Congrès Mondial de Linguistique Française (CMLF'2016).Tours, France.
- Ariel, M. (2001). Accessibility theory : An overview. Text representation : Linguistic and psycholinguistic aspects, 8, 29–87.
- Bagga, A., & Baldwin, B. (1998). Algorithms for scoring coreference chains. In The first international conference on language resources and evaluation workshop on linguistics coreference (Vol. 1, pp. 563–566).
- Charolles, M. (2002). La référence et les expressions référentielles en français. Editions Ophrys.
- Corblin, F. (1987). Indéfini, défini et démonstratif. Droz. Genève.
- Corblin, F. (1995). Les formes de reprise dans le discours. anaphores et chaînes de référence. Presses Universitaires de Rennes.
- Cornish, F. (1998). Les chaînes topicales : leur rôle dans la gestion et la structuration du discours. Cahiers de grammaire, 23(1), 9–40.
- 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.
- Denis, P., & Baldrige, J. (2009). Global joint models for coreference resolution and named entity classification. Procesamiento del Lenguaje Natural, 42.
- Désoyer, A., Landragin, F., & Tellier, I. (2015, juin). Machine Learning for Coreference Resolution of Transcribed Oral French Data : the CROC System. Vingt-deuxième Conférence sur le Traitement Automatique des Langues 439-445.
- Federzoni, S., Ho-Dac, L.M., Rebeyrolle, J., (2020) Les chaînes topicales dans la ressource ANNODIS. CMLF2020 : 7e Congrès Mondial de Linguistique Française, Jul 2020, Montpellier, France.
- Gabado, 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.
- Guillot-Barbance, C., & Quignard, M. (2019). Chaînes de référence et structure textuelle dans les essais sur la peinture de diderot. Discours. Revue de linguistique, psycholinguistique et informatique. A journal of linguistics, psycholinguistics and computational linguistics(25).
- 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.
- 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.

Autres références

- Landragin, F. (2015). Description, modélisation et détection automatique des chaînes de référence (democrat). *Bulletin de l'AFIA*(92), 11–15.
- Longo, L. (2013). Vers des moteurs de recherche "intelligents" : un outil de détection automatique de thèmes. méthode basée sur l'identification automatique des chaînes de référence (Thèse de doctorat).
- Luo, X. (2005, octobre). On coreference resolution performance metrics. In *Proceedings of human language technology conference and conference on empirical methods in natural language processing* (pp. 25–32). Vancouver, British Columbia, Canada : Association for Computational Linguistics.
- Manuélian, H. (2003). Descriptions définies et démonstratives : analyses de corpus pour la génération de textes (Thèse de doctorat non publiée). (Thèse de doctorat dirigée par Riley, Philip et Pierrel, Jean-Marie Sciences du langage Nancy 2 2003)
- Mélanie-Becquet, F., & Landragin, F. (2014, septembre). Linguistique outillée pour l'étude des chaînes de référence : questions méthodologiques et solutions techniques. *Langages*(195), 117-137.
- Mitkov, R. (2014). *Anaphora resolution*. Routledge.
- Oberle, B. (2019). Détection automatique de chaînes de coréférence pour le français écrit. *Conférence sur le Traitement Automatique des Langues Naturelles (TALN)*
- Oby, V., Glikman, J., Guillot-Barbance, C., & Pincemin, B. (2017). Les chaînes de référence dans les récits brefs en français : étude diachronique (xiii^e-xv^e s.). *Langue française*(3), 91–110.
- Quiniou, S., Cellier, P., Charnois, T., & Legallois, D. (2012). Fouille de données pour la stylistique : cas des motifs séquentiels émergents. *Journées Internationales d'Analyse Statistique des Données Textuelles* (821-833).
- Recasens, M., & Hovy, E. (2010). 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).
- Recasens, M., & Hovy, E. (2011). Blanc : Implementing the rand index for coreference evaluation. *Natural Language Engineering*, 17(4), 485–510.
- 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.
- 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.
- Salles, M. (2015). Chaînes de référence : la deuxième mention. l'exemple des entités inanimées dans les narrations littéraires. *Travaux de linguistique*(2), 111–133.
- Schnedecker, C. (2005). Les chaînes de référence dans les portraits journalistiques : éléments de description. *Travaux de linguistique*(2), 85–133.
- Schnedecker, C. (2014). Chaînes de référence et variations selon le genre. *Langages*(3), 23–42.
- Schnedecker, C., & Landragin, F. (2014). Les chaînes de référence : présentation. *Langages*(195), 3–22.

Autres références

- Schnedecker, C., & Longo, L. (2012). Impact des genres sur la composition des chaînes de référence : le cas des faits divers. In 3^{ième} congrès mondial de linguistique française (p. 1957-1972). Lyon, France.
- Todirascu, A., François, T., Bernhard, D., Gala, N., Ligozat, A.-L., & Khobzi, R. (2017, septembre). Chaînes de référence et lisibilité des textes : Le projet allusif. *Langue française*, 195(3), 35-52.
- Vilain, M., Burger, J., Aberdeen, J., Connolly, D., & Hirschman, L. (1995). A model-theoretic coreference scoring scheme. In Sixth message understanding conference (MUC-6) : Proceedings of a conference held in Columbia, Maryland, november 6-8, 1995.
- Walker, M. A., Joshi, A. K., & Prince, E. F. (1998). *Centering theory in discourse*. Oxford University Press.
- Wilkens, R., Oberle, B., Landragin, F., & Todirascu, A. (2020). French coreference for spoken and written language. In *Language Resources and Evaluation Conference (LREC 2020)* (p. 80-89). Marseille, France.

① Maillons

- Nature du maillon : SN_def, SN_dem, pronom, noms propres
- Fonction syntaxique : sujet, objet, autre
- Genre et nombre
- Position dans la chaîne
- Degré d'informativité du maillon : présence ou absence de modifieurs
- Type de référent : humain, non-humain, entité abstraite ou concrète, événement
- Type de reprise :
 - directe** : *La Standard Oil... ...la Standard Oil...*
 - indirecte** : *George H.B. Bush... ...le président...*
 - pronominale** : *Un jeune homme de 25 ans... ...il...*
- Type d'emploi : générique ou spécifique

① Maillons

- Interdistance (Rousier-Vercruyssen & Landragin, 2019) : distance en nombre de mots
- Distance en nombre de mentions
- Distance en nombre de phrases
- Distance en nombre de paragraphes

② Chaînes

- Longueur des chaînes en nombre de maillons
- Couverture des chaînes
- Cohabitation des chaînes

Ex. : **Un professeur d'histoire** vit une vie de famille sans histoire, entouré de l'affection des **siens** et aveugle aux problèmes politiques et sociaux engendrés par l'"apartheid", régime ségrégationniste qui sévit en Union sud-africaine. **Cet homme paisible** a **un jardinier noir**, paisible **lui** aussi, et soumis à la fatalité. **Ce jardinier** a **un fils** et **ce fils** [...]

Ex. : **Deux hommes** parurent. **L'un** venait de la Bastille, **l'autre** du Jardin des Plantes. **Le plus grand**, vêtu de toile, marchait le chapeau en arrière, le gilet déboutonné et sa cravate à la main. **Le plus petit**, dont le corps disparaissait dans une redingote marron baissait la tête sous une casquette à visière jaune. Quand **ils** furent arrivés au milieu du boulevard, ils s'assirent à la même minute, sur le même banc.

- Instabilité des chaînes : pourcentage de désignations différentes à l'intérieur d'une même chaîne (Rousier-Vercruyssen & Landragin, 2019)

③ Texte

- Structure textuelle
 - **Chapitres, sections, paragraphes** : chaînes locales ou globales
- Type textuel
- Genre textuel