



L'annotation du pronom " nous " dans un corpus de rapports éducatifs. Objectifs, méthodes, résultats

Emilie Née, Frédérique Sitri, Serge Fleury

► To cite this version:

Emilie Née, Frédérique Sitri, Serge Fleury. L'annotation du pronom " nous " dans un corpus de rapports éducatifs. Objectifs, méthodes, résultats. JADT 2014, Jun 2014, Paris, France. pp.495-506. hal-01070623

HAL Id: hal-01070623

<https://hal.science/hal-01070623>

Submitted on 3 Oct 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

L'annotation du pronom « nous » dans un corpus de rapports éducatifs. *Objectifs, méthodes, résultats*

Emilie Née¹, Frédérique Sitri², Serge Fleury³

¹ Université Paris-Est-Créteil – CEDITEC – emilie.nee@u-pec.fr

² Université Paris-Ouest Nanterre – MoDyCo – fsitri@u-paris10.fr

³ Université Sorbonne Nouvelle – CLESTHIA – serge.fleury@univ-paris3.fr

Abstract

This contribution describes the objectives, method and initial results of ongoing research on the annotation of the pronoun “nous” (“we/us”) in a corpus of reports on children “at risk” written by social workers (ANR Ecritures project). In a discourse analytical framework, this study uses textometric methods and tools in an attempt to uncover discursive “routines” from the social workers' final texts.

Résumé

Cet article vise à présenter les objectifs, la méthode et les premiers résultats d'un travail en cours sur l'annotation de la forme *nous* dans un corpus de rapports éducatifs produits par des travailleurs sociaux dans le cadre de la protection de l'enfance, corpus constitué au sein de l'ANR Ecritures. Ce travail se situe dans le cadre d'une analyse de discours outillée par la textométrie, qui vise à mettre en évidence des « routines » discursives.

Mots-clés : annotation, énonciation, analyse de discours, discours professionnel, phraséologie

1. Introduction

Le repérage des marques renvoyant à l'énonciateur, et tout particulièrement des pronoms personnels de 1^{ère} personne, constitue une « entrée » fréquente en analyse de discours, ouvrant entre autres à des analyses en termes de « positionnement du scripteur » (dans les écrits scientifiques), de « représentation de soi » (dans les discours politiques) ou encore d'« éthos » du locuteur, défini par Dominique Maingueneau comme « l'image de soi que le locuteur construit dans son discours pour exercer une influence sur son allocutaire » (Maingueneau, 2002 :238). D'un point de vue textométrique, ce type d'approche, qui peut passer par une annotation morphosyntaxique, recourt à des concordanciers ou des calculs de cooccurrences permettant d'identifier et de catégoriser le contexte et plus particulièrement le prédicat associé à la marque de personne (Poudat, 2006 ; Poudat et Follette, 2012 ; Leblanc et Fiala, 2004 ; Mayaffre, 2004, 2012).

Le travail que nous présentons se situe dans le cadre d'une recherche sur un genre de discours professionnel à visée évaluative – des rapports éducatifs produits dans la sphère de la protection de l'enfance – recherche visant à mettre au jour des *routines discursives* (Née et al., 2014) définies comme des blocs plus ou moins figés (patrons lexico-syntaxiques) appréhendés en relation avec les contraintes énonciatives et pragmatiques complexes de ce genre d'écrit. Nous nous intéressons ici à la façon dont le scripteur s'inscrit dans le texte de façon différentielle, selon les tâches ou les activités représentées et selon les « parties » du texte. Pour cela, nous nous sommes focalisés dans un premier temps sur les usages et les emplois de « nous » (« nos », « notre ») en procédant à une annotation sémantico-pragmatique des prédicats associés à « nous » à l'aide des fonctionnalités du logiciel *Le Trameur* (Fleury,

2013). Nous présentons ici les principes méthodologiques et les premiers résultats de cette recherche exploratoire.

1. Contexte de la recherche

1.1. Cadre de la recherche

Cette recherche s'inscrit dans le cadre de l'ANR *Écritures*¹, dont l'objectif est d'étudier, en conjuguant une approche discursive, génétique et textométrique, le processus d'écriture de rapports éducatifs produits dans le cadre de la protection de l'enfance et recueillis auprès d'un service d'accueil. Il s'agit plus précisément de rapports de synthèse ou de rapports d'échéance adressés au juge, qui visent à évaluer la situation du jeune confié au service – en internat ou en famille d'accueil – et à préconiser le maintien ou non dans le dispositif, ou une simple adaptation (changement de famille d'accueil, etc.).

Le corpus numérisé comprend différents états des rapports, enregistrés par les travailleurs sociaux. L'analyse textométrique présentée dans cet article a porté sur les états finaux des rapports, lesquels correspondent grosso modo au texte définitif envoyé au juge².

Pour le traitement lexicométrique, nous avons procédé à un double partitionnement du corpus : un partitionnement en dossiers (*i.e* les parties résultantes correspondant aux états finaux de tous les rapports) et un partitionnement en rubriques – dont nous allons voir plus loin l'intérêt. Les différents états de rédaction sont en effet découpés par les scripteurs en rubriques³ dont les libellés et les contenus se superposent plus ou moins. Afin de procéder à des comparaisons statistiques d'une portion de texte à une autre, nous avons superposé aux rubriques « endogènes » des scripteurs des « méta-rubriques » obtenues à l'aide d'une AFC (analyse factorielle des correspondances) sur le vocabulaire des rubriques « endogènes ».

Type d'écrits	Partitionnement	Nombre d'occurrences	Nombre de formes
états finaux de rapports éducatifs produits par un service de protection de l'enfance (SAFE de Caen)	dossiers (22) rubriques (12)	44 178	4 876

Tableau 1. Caractéristiques du Corpus ANR-Brouillons-Etats finaux

1.2. Hypothèses et questions de recherche

1.2.1. Genèse

Ce travail centré sur les emplois de « nous » s'inscrit dans la continuité des observations menées depuis 2003 sur un premier corpus de rapports éducatifs analysés dans le cadre d'un groupe de recherche mis en place à l'initiative d'André Collinot (Sorbonne Nouvelle,

¹ <http://www.univ-paris3.fr/anr-ecritures>

² Voir (Née et al., 2012) pour une présentation des différents corpus et une analyse des différents états de rédaction.

³ Ce découpage en rubrique évolue progressivement tout au long du processus de rédaction d'un rapport pour se stabiliser dans sa version finale.

Paris 3)⁴, rapports possédant des caractéristiques génériques communes avec ceux recueillis dans le cadre de l'ANR.

Du point de vue énonciatif, des contraintes complexes et paradoxales peuvent en effet être identifiées :

- le texte est rédigé par un ou plusieurs éducateurs et systématiquement relu par le chef de service qui le valide ; il est adressé au juge qui a ordonné la mesure mais est accessible à la famille et reste dans le dossier de l'enfant ;
- le scripteur est pris entre la relation d'aide établie avec la famille et la nécessité de « signaler » toute situation de danger encourue par l'enfant (positionnement paradoxal) ;
- il doit composer avec des contraintes rédactionnelles prônant l'objectivité qui peuvent être interprétées de différentes façons selon les services.

Ainsi, du point de vue du fonctionnement des marques renvoyant aux scripteurs, les analyses menées sur le premier corpus ont permis plusieurs constats. Tout d'abord, on note que « nous », seule marque de personne à apparaître dans ces textes, fait l'objet d'un emploi différencié selon les services : l'un des services se caractérise par un évitement systématique de « nous » (9 occurrences sur 27 887 contre 551 occurrences sur 63 540 dans un autre service → 0,3‰ vs 8 ‰) qui donne lieu à des tournures impersonnelles ou passives parfois surprenantes :

- *Madame DUVAL n'est rencontrée qu'en début de soirée*

- *Au regard de ces éléments, Benjamin pourrait bénéficier de la mise en œuvre d'une mesure d'AEMO*

Ensuite et de façon plus générale, dans les rapports où cette forme est employée, l'observation des occurrences de « nous » et de ses contextes d'emploi (en particulier les prédicats associés) nous a permis de faire apparaître une première catégorisation des activités ou des tâches des éducateurs, telles qu'elles sont représentées dans les rapports.

1.2.2. Questions de recherche

Cette première catégorisation a pu être mise à l'épreuve des données recueillies dans le cadre de l'ANR Écritures. Plus précisément, l'observation des corrections affectant l'emploi de « nous » d'un état de texte à l'autre nous a permis de formuler l'hypothèse d'une utilisation différenciée du pronom « nous » selon les activités représentées. On a pu en effet constater une tendance à l'ajout de « nous » dans le cas d'un locuteur « interprétant », déchiffrant un comportement ou un discours, et inversement une tendance à la suppression ou au remplacement de « nous » par « l'équipe » quand le locuteur se trouvait en position de prendre une décision relativement à la situation de l'enfant (changement de famille d'accueil par exemple) :

*il **paraît** nous **paraît également important** de relever que toute cette énergie et ce temps passé à s'occuper des autres lui permet de ne pas réfléchir à sa propre situation et surtout de ne pas avoir à affronter une réalité certainement difficile à vivre. (Pamela, doc manuscrit 1)*

⁴ Ce premier corpus était constitué de rapports correspondant à différentes mesures d'évaluation et d'accompagnement éducatif, issus de deux types de services (service d'accompagnement éducatif et service de recueil de signalement du conseil général) situés dans deux départements différents.

Notons d'ailleurs la variation référentielle de « nous ». Une des propriétés de « nous », on le sait, est sa malléabilité référentielle. Dans les textes que nous étudions, « nous » inclut toujours le scripteur mais son référent précis peut varier en fonction du prédicat associé, c'est-à-dire du type d'activité représentée. Globalement il peut s'agir des éducateurs chargés de suivre le jeune, de rencontrer la famille, etc. ou de l'équipe éducative en général ; les brouillons montrent d'ailleurs des substitutions entre « nous » et « l'équipe » :

~~Cette relation peut alors poser question au sein de l'équipe ...~~ → Cette relation complexe **nous** interroge car au-delà de leur camaraderie leur problématique XXX (à savoir addiction au tabac et consommation de produits illicites) peut s'avérer délétère. (Manon Dore, Manon sur le groupe de l'Orangerie)

En outre, les premiers relevés textométriques ont confirmé l'importance de cette forme, en raison de sa fréquence (elle se situe au 19^e rang des formes les plus employées, cf tableau 2) et de sa présence dans tous les dossiers (figure 1) – même si on note des variations importantes qui ne sont pas forcément liées à la taille différente des dossiers mais plutôt aux modalités de prise en charge, la forme « nous » étant logiquement plus présente dans le cas où l'enfant est placé au foyer, au contact quotidien des éducateurs.

Fréquence	Forme	Fréquence	Forme
2320	de	534	les
963	et	509	sa
909	à	477	est
868	la	452	une
815	le	446	elle
739	en	432	un
730	l	401	nous
704	d	395	du
607	a	391	avec
554	son	385	au
548	des	384	madame
541	il		

Tableau 2. Liste des formes les plus employées (Corpus ANR-Brouillons-Etats finaux)

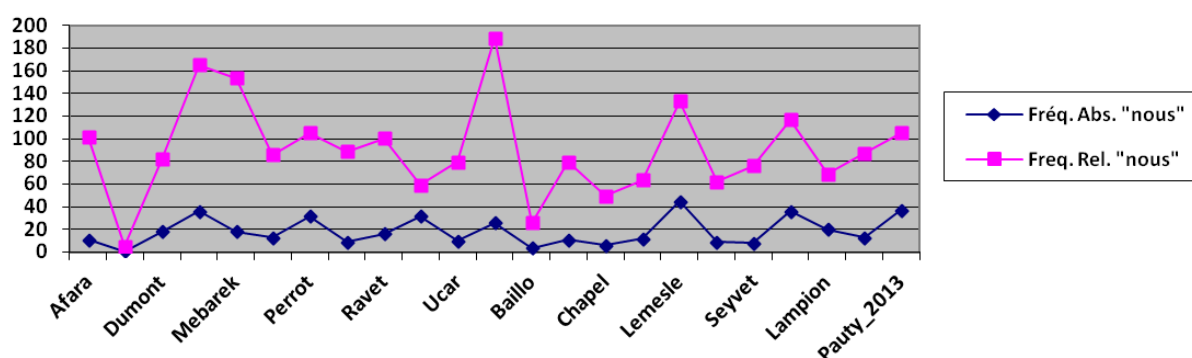


Figure 1. Fréquences de la forme « nous » dans l'ensemble des dossiers (Corpus ANR-Brouillons-Etats finaux)

Parallèlement, on observe que les segments répétés avec « nous » sont relativement nombreux, ce qui peut indiquer l'existence de patrons lexico-syntaxiques plus ou moins figés et de routines associées à « nous » :

Fréquence	SR <i>nous</i> + prédicat	Fréquence	SR <i>nous</i> + prédicat
65	nous avons	4	nous devons
15	nous constatons	4	nous restons
13	nous sommes	4	nous avons proposé
6	nous constatons que	4	nous constatons une
6	nous avons constaté	4	nous avions
5	nous avons été	4	nous avons organisé
5	nous observons	3	nous avons donc
5	nous lui avons	3	nous avons tenté
5	nous allons	3	nous rencontrer
5	nous ne	3	nous rencontrons
5	nous avons été	3	nous semblerait
5	nous observons	3	nous avons effectué
5	nous lui avons	3	nous avons rencontré
5	nous allons		

Tableau 3. Segments répétés avec « nous » (fréq. > 3, Corpus ANR-Brouillons-Etats finaux)

L'objectif de la recherche que nous présentons aujourd'hui est donc de valider ces premières observations, en nous demandant si l'on peut mettre en évidence l'existence de « postures » énonciatives différentes selon l'activité du scripteur, et si l'on peut associer ces « postures » différenciées à des routines discursives. Pour répondre à ces questions, nous avons décidé d'analyser plus finement les occurrences de « nous », c'est-à-dire leur contexte et le prédicat associé.

2. Méthodes et principes de l'annotation

Etant donné les questions de recherche formulées ci-dessus, nous avons fait l'hypothèse que l'annotation des prédicats associés à « nous » nous permettrait d'observer :

- leur fréquence respective ;
- leur répartition selon les rubriques ;
- les réseaux de cooccurents dans lesquels ils s'inscrivent, permettant de mettre au jour des patrons lexico-syntaxiques et donnant accès le cas échéant à des routines discursives.

Du point de vue méthodologique, nous avons pris la décision de faire porter l'annotation non pas directement sur le prédicat mais sur la forme « nous » associée au prédicat. Cette méthode permettait tout d'abord de distinguer « nous sujet » de « nous objet » (voir *infra*), distinction indispensable pour observer la façon dont le scripteur représentait son activité professionnelle. En outre, l'annotation des prédicats associés à « nous » supposait l'annotation de l'ensemble des prédicats du corpus, quel que soit le pronom dont ils sont précédés, entreprise beaucoup plus coûteuse en temps.

L'annotation a été réalisée *via* l'outil d'annotation intégré dans le logiciel *Le Trameur*⁵ (Fleury, 2013). Nous avons procédé en deux temps, en produisant tout d'abord une annotation syntaxique puis une annotation sémantique. Précisons déjà ici que les annotations ont porté sur deux niveaux, le niveau de la catégorie syntaxique mais aussi le niveau du lemme (figure 2), ce qui facilite la recherche d'associations de nature multidimensionnelles (*Nous_Sujet* + ADV ou *nous_sujet_sentiment* + *toutefois*, par exemple) et, par là, la mise en évidence de patrons lexico-syntaxiques.

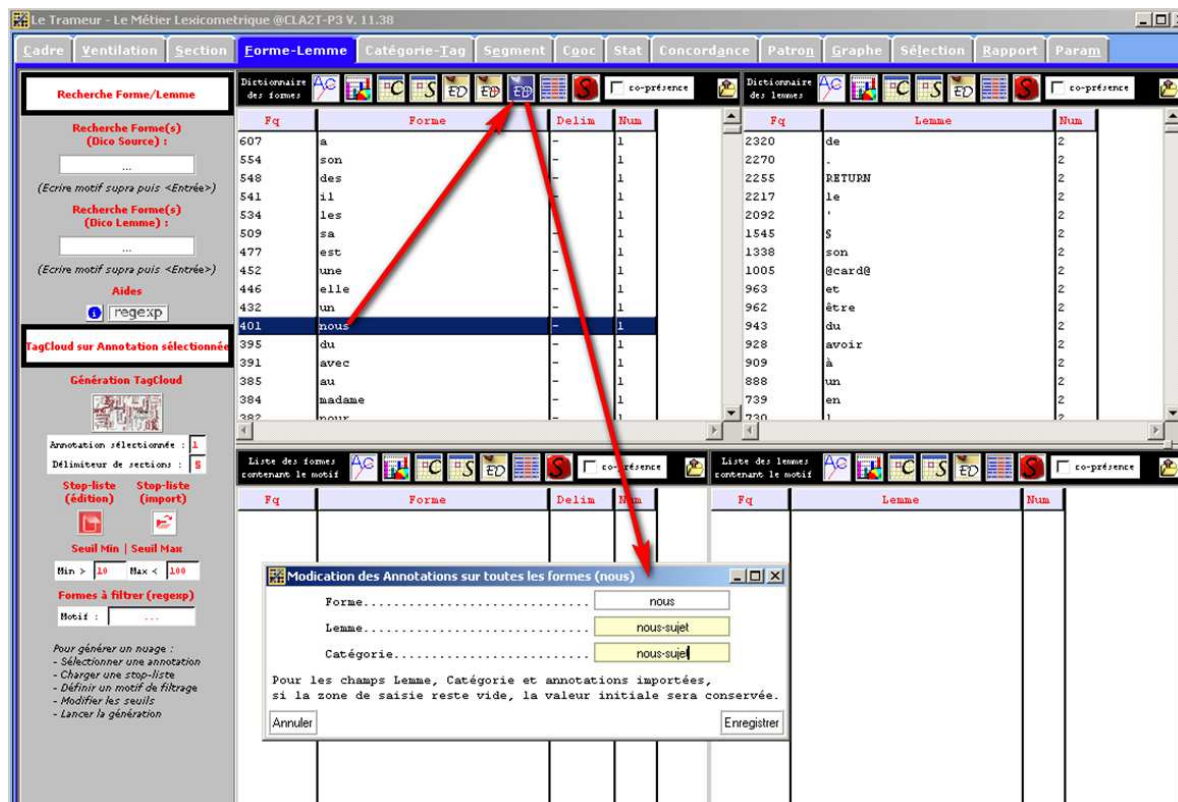


Figure 2. Annotation de « nous » avec Le Trameur

2.1. L'annotation syntaxique

L'annotation syntaxique a conduit à distinguer les *nous* « sujets » des *nous* « objets ». Nous avons tout d'abord attribué à toutes les occurrences de « nous » l'annotation « sujet », ressentie intuitivement comme la plus fréquente, puis nous avons procédé à une correction manuelle. L'annotation sur ce premier niveau confirme une distribution asymétrique :

*Nous_Obj*et : 70 occurrences - *Nous_Suj*et : 357 occurrences

Les occurrences de *Nous_Obj*et correspondent pour l'essentiel à des constructions du type :

x + nous + verbe de parole

Didier nous informe que... Elle a pu nous dire que...Etc.

⁵ Le logiciel *Le Trameur* met en effet à disposition du chercheur un outil permettant d'associer à un item (unité textuelle issue de la segmentation du texte) autant d'annotation qu'il le souhaite. De plus, il intègre l'annotateur morphosyntaxique *Treetagger*.

Dans la catégorie *Nous_Sujet*, nous avons pris en compte bien évidemment les « nous » en fonction sujet, mais également les modalisations avec constructions impersonnelles *Il nous semble que...* / *Il nous est apparu indispensable de...* ainsi que les constructions du type *cela nous a permis de mesurer la capacité de*. Avec « sembler », « le datif indique l'actant animé aux yeux de qui se produit l'impression de semblance » (Le Goffic, 1993:§125). Dans les constructions avec « nous » datif et infinitive, « nous » peut être analysé comme sujet de l'infinitif : « si le verbe impersonnel a un complément indirect, c'est naturellement celui-là qui permet de restituer le sujet de l'infinitif » (Riegel et al., 1994 : 498).

2.2. L'annotation sémantique

L'annotation sémantique s'est appuyée sur des catégories sémantico-pragmatiques correspondant aux différentes activités des scripteurs. Procédant de façon inductive à partir de notre connaissance du corpus, nous avons dégagé les catégories suivantes :

Catégories	Exemples
<i>nous_sujet_faire</i>	nous avons organisé des rencontres à chaque période de vacances scolaires.
<i>nous_sujet_dire</i>	nous lui avons précisé qu'elle ne pourrait avoir des photos de didier
<i>nous_sujet_sentiment</i>	en revanche nous restons inquiets en ce qui concerne les relations amoureuses de jade
<i>nous_sujet_preco</i> (=préconisation)	nous proposons la poursuite de la mesure de placement familial séquentiel
<i>nous_sujet_constat</i>	nous ne constatons pas une alimentation excessive ni pendant les repas ou par de bonbons ou des gâteaux par exemple.
<i>nous_sujet_pensée</i>	ne nous semble pas en capacité d'élaborer le lien entre ses attitudes ou propos avec les réponses très agressives des autres garçons.

Tableau 4. Catégories sémantico-pragmatiques pour l'annotation des prédicats avec « nous »⁶

Notre démarche ici est proche de celle adoptée par les chercheurs travaillant sur les discours scientifiques, qui dégagent des catégories de prédicats en relation avec les activités ou les tâches du chercheur (voir par exemple (Fløttum et Thue Vold, 2010 ; Tutin, 2010)). (Leblanc et Fiala, 2004), partant eux des catégories sémantiques « a priori », remarquent la nécessité du retour au contexte pour en préciser la valeur dans le discours étudié. D'un point de vue méthodologique, nous avons dans un premier temps associé de manière automatique l'annotation *nous_sujet_faire* à l'ensemble des items comportant l'annotation *Nous_Sujet*, cette catégorie de prédicat paraissant intuitivement la plus fréquente. Nous avons ensuite procédé à une correction manuelle. Par ailleurs, l'observation attentive des données a montré la nécessité d'annoter également les adjectifs possessifs dans des expressions du type « *notre service* », « *de notre part* », « *nos observations* », « *nos remarques* » comme dans :

- *cette consultation a été effective le 10/11/09 et a nécessité une contention longue de notre part*

⁶ La catégorie *nous_sujet_preco* formellement hétérogène comprend des verbes de dire en emploi performatif, comme le montre l'exemple.

- madame mercier qui souhaitait maintenir son accueil du jeune au moment de son orientation au abcd *avait été recrutée par notre service.*
- il peut formuler des excuses et *se montrer sensible à nos observations.*

3. Premiers résultats

3.1. Répartition selon les rubriques

Un premier résultat attendu est la mise au jour du lien de certains prédicats avec certaines rubriques. Un exemple particulièrement significatif est l'emploi privilégié du prédicat *nous_sujet_preco* dans la rubrique *conclusion* des rapports (figure 3).

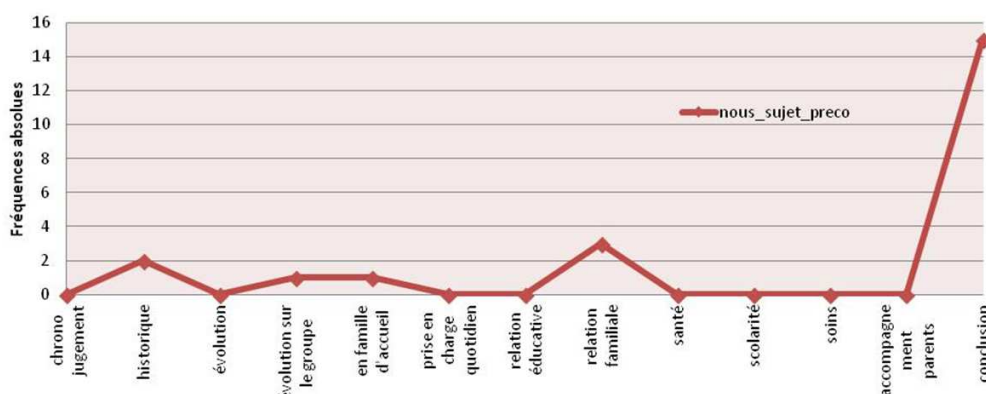


Figure 3. Ventilation des *nous_sujet_preco* dans les rubriques du corpus ANR-Brouillons-Etats finaux

De plus, la courbe de la ventilation des formes « nous » + prédicats sur la partition *rubriques* met en évidence des profils parallèles ou au contraire contrastés en fonction des prédicats. Par exemple, les prédicats *nous_sujet_faire* et *nous_sujet_pensée* ont une distribution analogue, avec néanmoins des écarts de fréquence substantiels, tandis que le prédicat *nous_sujet_constat* investit de manière plus importante des rubriques (*évolution* et *santé*) où les autres prédicats sont bien moins représentés (figure 4).

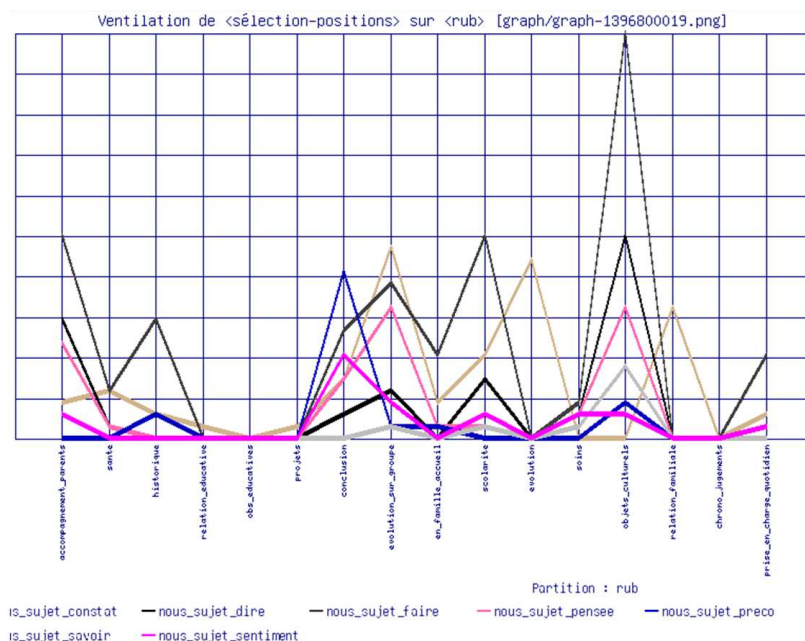


Figure 4. Ventilation des différentes catégories de prédicats dans les rubriques du corpus ANR-Brouillons-Etats finaux (sortie Le Trameur)

3.2. Distribution différentielle des modalisations selon les types de prédicats

L'approche quantitative par l'annotation de « nous » ouvre sur une description plus qualitative des caractéristiques des prédicats. Un certain nombre d'observations peuvent être faites, qui peuvent à leur tour ouvrir sur d'autres explorations textométriques (calculs de cooccurrences par exemple).

3.2.1. Diversité des prédicats

On constate tout d'abord que la diversité des prédicats varie de façon notable selon la catégorie. La plus grande diversité s'observe avec « faire », catégorie également la plus importante quantitativement (130 occurrences) : cette diversité correspond à la variété des activités et des tâches réalisées par les travailleurs sociaux dans le quotidien de la prise en charge d'un enfant, activités qui sont de l'ordre de l'établissement d'un lien et de l'organisation du quotidien : *organiser (des rencontres), rencontrer, recevoir, effectuer (visite), reporter (visite), être en contact avec, accompagner, mettre en place (accueils relais, échanges téléphoniques), maintenir le placement, effectuer un changement d'école, ...* À l'inverse on note la faible variété des prédicats « constat » dont les 72 occurrences se réduisent à 5 verbes différents : *constater, noter, s'apercevoir, observer, remarquer* et *montrer* dans la tournure « *X nous montre que* ». Ainsi comprend-on mieux la fréquence du segment répété « nous constatons » repérée dans la liste des segments répétés. Un examen attentif du contexte met par ailleurs en évidence que « nous constatons » est fréquemment suivi d'éléments de diagnostic, de type psychologique ou portant sur un comportement ou une situation : « nous avons pu constater que les défenses psychiques de R étaient très installées », « nous avons constaté que les efforts fournis par Damien concernant son comportement n'ont pas été pris en compte », « nous constatons une dégradation de son état physique et psychique », « nous constatons que Damien n'utilise plus de grossièretés dans son langage ». On voit apparaître de la sorte des patrons déjà identifiés : *nous constatons + difficultés/est en/dans ; nous constatons + évolution* (voir (Née et al., 2014)).

3.2.2. Un nombre important d'auxiliations/modalisations

Le classement des prédicats permis par l'annotation met en évidence un autre trait de variation : certaines catégories de prédicats se caractérisent par la présence d'un nombre important d'auxiliaires, correspondant le plus souvent à l'expression d'une modalisation, et que l'on peut répartir en 5 ensembles :

Possibilité/permission : *nous pouvons/X nous permet de ;*

Obligation/contrainte : *nous devons/sommes contraints/obligés de/X nous fait⁷/nous amène/conduit à/nécessite de notre part ;*

Décision : *nous faisons le choix/choisissons/décidons de/voulons ;*

Tentative : *nous tentons/nous nous efforçons de ;*

Atténuation : *il nous semble, nous paraît ;*

La catégorie « faire » est celle qui comprend le plus grand nombre de prédicats modalisés selon ces 4 axes, avec une prédominance de la modalisation de type « obligation » : *nous a amené à conduire Léa aux urgences, nous ont conduit à stopper les rencontres ...* et « décision » : *nous avons fait le choix d'un changement de famille d'accueil, nous avons donc*

⁷ On a là affaire formellement à un factitif. Il ne nous est pas possible dans les limites de cet article de développer cet aspect de nos analyses.

sentiment (peur/inquiétude) » (tableau 5) que nous avons associé à un fonctionnement discursif relevant de la prise de position par le scripteur dans ces écrits – l'association du patron et du fonctionnement discursif constituant une routine discursive, selon notre définition (Née et al., 2014).

PATRON	EXEMPLES	FONCTIONNEMENT DISCURSIF
Patron « marqueur concessif + nous + prédicat sentiment » CONN CONC + nous + prédicat sentiment (peur/inquiétude)	<i>en revanche nous restons inquiets en ce qui concerne les relations amoureuses de Jade</i> <i>le collègue parle de possibilité d'intégrer un lycée professionnel standard mais nous craignons que adèle ne se retrouve en décalage important sur le plan scolaire</i> <i>nous restons toutefois attentifs à l'évolution d'Elise dans la mesure où elle présente un retard de langage important</i>	PRENDRE POSITION

Tableau 5. Le patron « CONN CONS + nous +prédicat sentiment »

4. Conclusion

Un certain nombre de difficultés méthodologiques doivent être soulevées. Ainsi, comme nous l'avons vu à propos de l'un des premiers corpus étudiés, les activités du scripteur peuvent-elles être représentées par des expressions qui ne contiennent pas « nous » comme *il y a lieu de* :

§ actuellement, il y a lieu de garantir ce lieu de vie et la prise en charge en famille d'accueil qui répond à ses besoins

Par ailleurs, on peut relever des exemples où les marques du scripteur restent implicites, comme le s du pluriel dans l'exemple suivant :

*cependant les attitudes de monsieur roux observées lors des rencontres avec evelyne et à l'occasion de la visite avec odile laissent très **interrogatifs** quant à sa capacité à porter de l'attention à ses enfants pendant un droit de visite.*

Néanmoins nous pensons avoir montré que le type d'annotation proposée ici, qui affine l'exploration textométrique, permet à la fois de préciser l'analyse énonciative (différentes catégories de prédicats et type de modalisations) et de mettre au jour des patrons plus complexes que les simples segments répétés.

Ce premier travail ouvre d'autres perspectives : tout d'abord procéder à l'annotation du premier corpus afin de valider nos résultats sur un nombre d'occurrences plus important. Nous nous proposons ensuite de mettre à l'épreuve la méthodologie présentée sur d'autres corpus relevant de genres évaluatifs (rapports d'inspection par exemple).

Références

Fleury S. (2013). Le Trameur. Propositions de description et d'implémentation des objets textométriques. Publication sur le site de l'Université Paris 3 <<http://www.tal.univ-paris3.fr/trameur/trameur-propositions-definitions-objets-textometriques.pdf>>

- Fløttum K. et Thue Vold E. (2010). L'éthos auto-attribué d'auteurs doctorants dans le discours scientifique. *Lidil*, 41 : 41-58.
- Leblanc J.-M. et Fiala P. (2004). Autour du Je présidentiel. Développer la méthode expérimentale en analyse textuelle. In Purnelle G., Fairon C., Dister A., eds, *Actes des 7^{èmes} Journées Internationales d'Analyse des Données Textuelles*, pp. 716-725.
- Maingueneau D. (2002). Ethos. In Charaudeau, P. et Maingueneau, D., *Dictionnaire d'analyse du discours*. Seuil.
- Mayaffre D. (2012). *Mesure et démesure du discours. Nicolas Sarkozy (2007-2012)*, Paris, Les Presses de Sciences Po.
- Née E., Sitri F. et Veniard M. (2014). Pour une approche des routines discursives dans les écrits professionnels. *Actes du 4^{ème} Congrès Mondial de Linguistique Française* (en cours de publication).
- Née E., MacMurray E. et Fleury S. (2012). Textometric Explorations of writing processes: a discursive and genetic approach to the study of drafts. In Dister A., Longrée D., Purnelle G., eds, *Actes des 11^e Journées internationales d'Analyses statistiques des Données Textuelles*, pp. 767-778.
- Poudat C. (2006). *Etude contrastive de l'article scientifique de revue linguistique dans une perspective d'analyse des genres*. Thèse pour le doctorat en Sciences du Langage, Université d'Orléans.
- Poudat C. et Follette P. (2012). Corpora and Academic writing. A contrastive analysis of research articles in biology and linguistics. In Boulton, A., Carter-Thomas, S., Rowlet-Jolivet, E. *Corpus-Informed Research and learning in ESP. Issues and applications*. John Benjamins Publishing Company.
- Tutin A. (2010). Dans cet article, nous souhaitons montrer que... Lexique verbal et positionnement de l'auteur dans les articles en sciences humaines. *Lidil*, 41: 15-40.