

Des dynamiques entre un genre de discours et un agencement textuel : le cas de la liste dans les rapports d'activité de laboratoire (1970-2018)

Virginie Lethier^{1*}, Émilie Née², Timothée Premat^{2,3}

¹Université Marie & Louis Pasteur, *ELLIADD*, 30-32 rue Mégevand, 25030 Besançon, France

²Université Paris-Est Créteil, *Céditec*, 61 avenue du Générale de Gaule, 94010 Créteil, France

³Université de Neuchâtel, *chaire de dialectologie*, Pierre-à-Mazel n°7, 2000 Neuchâtel, Suisse

Résumé. Cet article interroge dans une perspective discursive les formes, les usages et les fonctions de la liste dans un genre de discours où elle foisonne : le rapport d'activité de laboratoire, de 1971 à 2018. Après avoir justifié notre définition de l'objet liste, nous exposons la méthode par laquelle ont été annotées les listes dans le corpus étudié. Nous présentons une première analyse des caractéristiques majeures de l'emploi des listes et pointons comment cet agencement textuel engage des effets témoignant de son rôle d'interface entre l'écriture gestionnaire et scientifique.

Forme textuelle aux apparences minimalistes, la liste a longtemps été reléguée à un procédé auxiliaire de mise en page et par suite, soustraite à un examen théorique approfondi, tout comme l'a été un autre palier mésotextuel de structuration du texte qu'est le paragraphe [1]. L'éclatement disciplinaire des travaux ayant pris cette forme langagière pour objetⁱ, tout en contribuant à l'absence d'une définition stabilisée pour un objet extrêmement protéiforme et plastique, nous semble témoigner de la difficulté de saisir la liste en dehors de ses actualisations dans des configurations et des pratiques sociohistoriques données. Pour autant, à l'exception notable de quelques travaux [2, 3, 4, 5], la liste n'a que rarement été interrogée au prisme de ses rapports aux genres de discours [6] et aux formations discursives, et l'analyse du discours ne s'est que rarement risqué à prendre cet observable pour objet [5, 7].

S'inscrivant dans le projet ArchivUⁱⁱ, cette communication se propose d'interroger dans une perspective discursive les formes, les usages et les fonctions de la liste dans un genre de discours où elle foisonne : le rapport d'activité de laboratoire, de 1971 à 2018.

Dans un premier mouvement, nous exposerons les postulats et les résultats antérieurs de nos travaux sur le genre du rapport qui invitent l'analyse du discours à s'ouvrir aux composantes sémiotiques et textuelles et à intégrer comme niveau d'analyse, non seulement le texte, mais également les images du texte [8]. Envisageant la liste comme un lieu textuel [9] reflétant et alimentant les conditions de production et de circulation d'un genre de discours destiné à en alimenter d'autres, nous formulons également l'hypothèse que cet agencement participe des visées et dimensions argumentatives [10] du genre du rapport. Il

* Auteur correspondant : virginie.lethier@univ-fcomte.fr

s'agit donc de montrer comment le texte, par les images qu'il donne à voir de lui-même [8], permet de construire l'image d'un laboratoire [11, 12].

Après avoir justifié notre définition de l'objet liste, élaborée à partir de l'observation des opérations de segmentation et de liaisons opérées dans le genre du rapport d'activité, nous exposerons brièvement la méthode par laquelle ont été annotées les listes dans le corpus étudié, avant de présenter une première analyse des caractéristiques majeures de l'emploi des listes. En particulier, nous pointerons comment cet agencement textuel engage des effets témoignant de son rôle d'interface entre l'écriture gestionnaire et scientifique. Sur le plan méthodologique, il s'agit ainsi de montrer les apports des solutions d'exploration et de visualisation permises par la textométrie à partir de corpus finement annotés et tenant compte de la structure compositionnelle et de la segmentation graphique des textes.

1 Du genre de discours *rapport de laboratoire* à la liste

Afin d'éclairer les usages textuels et discursifs de la liste dans les rapports de laboratoire, il convient tout d'abord de préciser quelques-unes des caractéristiques fonctionnelles, communicationnelles et matérielles du genre discursif considéré.

1.1 Un genre de discours interface entre les sphères gestionnaires et scientifiques

S'inscrivant dans un circuit de communication institutionnelle, le rapport de laboratoire correspond à un écrit de commande adressé, destiné à faire interface entre deux sphères d'activité, et par suite, deux types de discours : le discours scientifique d'une part, le discours gestionnaire de l'autre. Que le commanditaire soit une tutelle (Conseil scientifique de l'université, CNRS) ou une agence évaluatrice (AÉRES, HCÉRES), celui-ci contraint la production du rapport, en imposant une trame, des thématiques à traiter, des séquences textuelles (narrative, descriptive, explicative, argumentative) à respecter. C'est ce qu'explique le scripteur de l'extrait suivant, critiquant précautionneusement la pertinence du cadrage imposé et pointant les difficultés de « faire avec » la double visée informative et argumentative du genre :

Nous suivrons dans la présentation de ce rapport le plan proposé par la commission, tout en soulignant qu'il est souvent difficile de scinder l'ensemble des arguments et des faits et de les répartir [sic] entre chaque rubrique (Rapport CNL 1974, p. 2)

Le rapport de laboratoire est en effet marqué par une tension constitutive entre ses visées informatives et argumentatives, puisque dans son espace, le laboratoire doit simultanément « rendre compte » d'une activité scientifique et s'autolégitimer en valorisant son activité [13]. Dans un contexte concurrentiel que l'exercice participe à entretenir, le rapport est un lieu de construction de l'ethos mélioratif et distinctif d'un laboratoire, intégrant, dans un intérêt stratégique, les normes et critères du commanditaire.

Enfin, le rapport nécessite d'être éclairé à la lumière des réseaux de genres de discours dans lesquels il est destiné à s'insérer, en amont et en aval. Une des caractéristiques du rapport est en effet sa forte *valence générique externe* [14], c'est-à-dire que le rapport entretient de façon consubstantielle des liens de séquentialité qui l'unissent à d'autres genres de discours : en aval, le rapport d'activité donne lieu à un « rapport de rapport » [15]. Du côté du scripteur, l'anticipation du devenir générique du rapport d'activité et du type de lecture évaluative qu'il induit transparaît nettement dans l'extrait suivant, adressé à un « rapporteur » :

Le rapporteur à qui sera confié le dossier aura intérêt à se reporter la brochure que nous avons rédigée pour l'année 1972-1973, afin de mieux situer l'état de notre recherche. Il pourra prendre connaissance aussi de la photocopie du bref rapport manuscrit déposé à l'appui de notre demande [sic] d'un de technicien auprès du CNRS. Ce rapport expose les arguments que nous ne pouvons

que répéter dans ce rapport-ci. Il illustre aussi les conditions du fonctionnement de l'équipe. (Rapport CNL, 1974, p. 2-3)

En amont, comme l'explique cet extrait, le rapport d'activité est relié à d'autres genres (des notes, des « rapports », des lettres, des échanges oraux) antérieurement produits, souvent joints en annexe. Engagée sous la contrainte d'obligations légales et/ou statutaires, la rédaction du rapport se surajoute aux activités principales du laboratoire : de fait, la production d'un rapport est soumise à un impératif d'économie et fréquemment opérée par le recyclage de matériaux textuels préexistants.

1.2 De la dimension argumentative du plan verbo-visuel dans le genre du rapport à la liste

Dans le cadre de nos précédentes recherches sur le rapport de laboratoire en diachronie [16, 11], nous avons montré la façon dont la dimension verbo-visuelle (iconographie, police de caractères, segmentation du texte) agit, dans le rapport, comme un ressort argumentatif de premier plan et participe à construire l'ethos distinctif du laboratoire [12]. Dès lors, nous visons, dans notre pratique d'analyse du discours, à enrichir notre approche traditionnelle centrée sur la dimension verbale par des observables liés au format, tel que la liste, sur laquelle sera centrée la présente contribution.

Envisageant la liste comme un agencement textuel reflétant et alimentant les conditions de production et de circulation d'un genre discursif destiné à en alimenter d'autres, nous visons en premier lieu à mesurer la place occupée par la liste dans les rapports en diachronie, pour interroger ses liens aux conditions concrètes de production du rapport, notamment dans ses dimensions technologiques et temporelles.

Par ailleurs, le rapport est un genre discursif dont les fonctions ont profondément évolué dans un contexte institutionnel en constante transformation depuis la loi Savary (1984). La « contractualisation » (1989) a entraîné le passage d'un commanditaire interne (le conseil scientifique de l'université) à un commanditaire externe (le ministère), remplacé en 2006 par l'Agence d'Évaluation de la Recherche et de l'Enseignement Supérieur (AÉRES), elle-même devenue en 2013 le Haut Conseil de l'Évaluation de la Recherche et de l'Enseignement Supérieur (HCÉRES). Ce faisant, d'un instrument de reconnaissance et de contrôle (interne puis externe) de l'activité scientifique, le rapport devient avec l'AÉRES un outil d'évaluation externe extrêmement standardisé afin de permettre des comparaisons nationales de l'activité scientifique. Dans quelle mesure la liste apparaît-elle sensible à ces jalons et témoigne-t-elle d'une évolution des stratégies discursives et argumentatives des unités de recherche vis-à-vis de leur commanditaire ? Au-delà du profil quantitatif de la liste en diachronie, nous visons à observer les caractéristiques lexico-thématiques et syntaxiques de cet agencement textuel. Dans le rapport, en investissant la forme textuelle de la liste, laquelle anticipe une lecture non linéaire et orientée vers la sélection et la reprise d'éléments de contenu dans un « rapport de rapport », que tend-on à dire, à mettre en avant, ou au contraire à passer sous silence ?

1.3 Corpus d'étude

Pour tenter de répondre à ces questions, nous avons constitué un corpus de dix-huit rapports, publiés entre 1971 et 2018 à l'université de Nanterre, par trois laboratoires relevant de la linguistiqueⁱⁱⁱ : le *Centre de néologie lexicale* (CNL) créé en 1971, le *Centre de recherches linguistiques* (CRL), qui succède au CNL à partir de 1979, puis le laboratoire *Modyc*, créé en 2001 et succédant au CRL. Bien qu'il vise à rassembler l'ensemble des documents produits durant la période étudiée, notre corpus d'étude présente des lacunes à partir de 1991. Celles-ci s'expliquent notamment par l'absence d'archives repérées pour certaines années

(1992-2000 ; 2012-2018) et par l’existence de dossiers dont la numérisation ou le prétraitement est encore en cours (rapport Modyco 2005-2008).

Ce corpus représente 148 553 mots, avec en moyenne 8250 mots par rapport. La taille des rapports évolue fortement dans la deuxième partie de notre diachronie (figure 1), avec plus de 32 000 mots dans les deux derniers rapports ; pour le dernier rapport, cela représente 46 pages avec une police de caractère de petite taille et un faible interligne. Cette évolution est peut-être en lien avec le passage à la rédaction sur ordinateur, sensible à partir de 2004 dans notre corpus, et vient nuancer l’importance des lacunes dans la mesure où le corpus reste équilibré en nombre de mots. Enfin, il faut noter que ce corpus dispose de deux textes pour la première année, lors de la création du laboratoire : l’un est adressé à l’université et l’autre au CNRS (ces deux textes sont les plus courts du corpus).

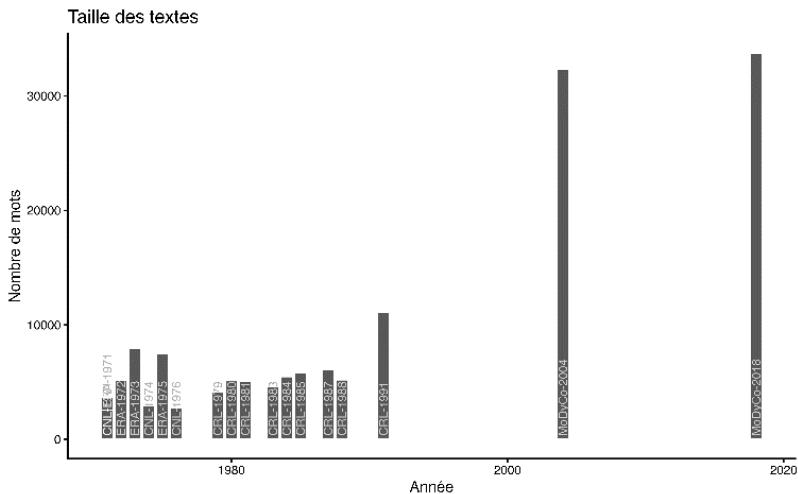


Figure 1. Distribution de la taille des textes du corpus en fonction de leur date.

2 Saisir la liste dans un corpus de rapports d’activités de laboratoire

2.1 Ce dont la liste est le nom : des définitions plus ou moins extensives pour un objet plastique

« [U]ne liste, on sait tous intuitivement ce que c’est, mais, doit-on répondre à l’injonction de la définir, on est à la peine... » [2]. Si de nombreux travaux s’accordent pour identifier la liste comme un agencement textuel (ou un objet textuel) et graphique résultant d’un acte d’énumération [4], la plupart se heurtent à des problèmes de définition ou plus précisément d’identification – cette séquence peut-elle être qualifiée de liste ? – et de délimitation – où s’arrête la liste ? –, en particulier lorsque cet objet fait partie d’un ensemble textuel plus grand. Il est d’ailleurs intéressant de constater que plusieurs travaux préfèrent rejeter le terme de *liste* soit pour ne pas prendre en compte un certain nombre de configurations [5], soit au contraire pour l’intégrer aux diverses configurations énumératives [17].

Dans une conception extensive et souple de la liste [17, 18, 19], où le texte est envisagé comme une structure émanant d’une activité discursive, liste et énumération se recouvrent en partie voire se confondent, la liste pouvant correspondre à des formes non marquées par la verticalité. Dans une conception plus restrictive de la liste, attentive à la matérialité graphique

du texte [4] et à ses effets sur le sens, le critère topographique est discriminant pour distinguer la liste de l'énumération (« on dresse des listes », rappelle Mahrer [4])

Un tel mouvement de balancier définitionnel nous conduit à penser la plasticité de la liste en fonction du genre, ainsi que le proposait avant nous Rabatel :

« Il y a *des* genres de liste et il est difficile d'ériger un agencement (la liste d'items d'un mot en colonne) en un protographe ou en un prototype. L'idée de protographe, pour autant que je puisse en juger, ne résiste pas à l'analyse historique et la prototypie explose devant la diversité des formes. » [2, p. 270]

Notons déjà que si on assimile la liste à toutes les formes d'énumérations, l'identification des listes permet de saisir la prégnance de l'activité énumérative dans un texte ou un genre de discours donné – dans notre cas, elle permet de voir que le rapport de laboratoire contient une activité énumérative importante pour ne pas dire structurante. Restreindre l'identification des listes aux énumérations verticalement disposées sur l'espace graphique, permet d'aiguiller l'analyse vers des moments-clés, propices à une scénographie textuelle [20] visuellement saillante.

Partant de l'observation qu'un certain nombre de configurations énumératives participent à la dimension verbo-visuelle du genre (voir *supra*, 1.2), nous avons pris comme point de départ la définition que proposent Rebeyrolle et Péry-Woodley [17] de l'énumération (incluant la liste), à la suite de Luc *et al.* [21] :

« nous envisageons les structures énumératives comme un procédé textuel de base (une sorte de « brique textuelle ») que l'on peut décrire comme le fait d'agencer du texte de façon que le lecteur soit amené à prendre conscience de cet agencement. Dans les termes de Luc *et al.* [21], cet agencement constitue un acte textuel : **le scripteur dispose du texte de manière que le lecteur reconnaisse à la fois la disposition et l'intention qui lui a présidé**. L'intention est que soit reconnue l'identité des éléments énumérés (item/idem) en relation à un critère spécifique. Ce critère, explicite ou inférable, sera appelé critère de co-énumérabilité. » [17, pp. 3184, nous soulignons],

définition que nous avons été amenés à adapter au genre de discours que nous étudions.

2.2 Des formes variées de la liste dans le corpus de rapport à une nouvelle définition de la liste

La définition ci-dessus nous amène à prendre en compte une grande diversité d'énumération, diluant le phénomène qui avait retenu notre attention au départ, des listes visuellement saillantes, en raison d'un marquage par la ponctuation blanche ou par une ponctuation noire forte. Dans une double logique d'annotation et de conceptualisation qui tienne compte des spécificités du genre du rapport, nous avons retenu deux objets, les listes verticales et les listes horizontales, telles que définies en (1).

(1) Objets retenus comme listes :

- (a) Listes verticales : signalées par un dispositif de ponctuation blanche, avec des retours à la ligne entre les items, avec ou sans puce ou numéros, et avec ou sans dispositif de démarcation par la ponctuation noire,
- (b) Listes horizontales : signalées par un dispositif de démarcation par la ponctuation noire, sans retour à la ligne entre les items (sans 'verticalité'), avec ou sans puce ou numéros entre les items

En pratique, un bloc textuel n'est reconnu comme liste que s'il répond aux critères (2) :

(2) Critère pour l'identification des listes :

- (c) Critère de démarcation : une liste doit être ouverte et fermée par un dispositif de ponctuation blanche (retour à la ligne) ou de ponctuation noire (signe de ponctuation).
- (d) Critère d'énumérabilité : une liste doit être introduite par un énumérable, et les items de la liste doivent correspondre à des sous-types de cet énumérable.

- i. Dans les cas où l'énumérable est sous-entendu (ce qui est courant), sa position syntaxique doit pouvoir être intuitivement restituable avant la ponctuation (noire ou blanche) qui ouvre la liste.

Définie ainsi, la liste se présente comme une sous-catégorie de l'énumération, singulière en ce qu'elle met en scène l'accumulation par un dispositif visuo-spatial ou typographique qui excède la seule syntaxe. Aucune limite maximale de taille ou d'enchâssement des listes n'est prise en compte. Un nombre minimal de deux items est appliqué pour les listes dites horizontales, sauf en cas d'effet de parallélisme dans une série de listes horizontales. Ce critère permet d'exclure les très nombreux cas où une parenthèse suit un nom ou syntagme énumérable en se limitant à le développer ou le compléter.

Notons que les deux types de listes sont pris dans un rapport de continuum avec deux autres modalités de structuration du texte [17] : les listes verticales penchent du côté de la macrostructure du texte (niveaux de titre), tandis que les listes horizontales penchent du côté du plein texte (énumération sans mise en forme spécifique).

2.3 Continuum *listes verticales - macro-structure* : décisions d'annotations

Les listes verticales ne posent le plus souvent pas de problème d'identification, dans la mesure où elles sont fortement séparées du texte qui les environne. Toutefois, certaines configurations peuvent entraîner des hésitations de la part de l'annotateur, comme lorsqu'elles participent de la macro-structure du texte : l'*énumérant* est pris dans un niveau de titre ou commande lui-même un niveau de titre. La structuration hiérarchique (numérotation et traitement typographique) des titres du document ne permet pas toujours de trancher, dans la mesure où elle est peu homogène. L'exemple de figure 2 montre l'une de ces situations de brouillage entre niveaux de titres et listes. Cet exemple a été retravaillé ; dans l'original, seul des niveaux inférieurs (non reproduits ici) sont indentés.

```

2. Activités de l'année universitaire 1985-1986

L'année 1985-1986 comportera une activité répartie entre des [sic] deux direc-
tions suivantes :

2.1. Poursuite, dans le cadre du programme quadriennal, des différents
sous-programmes, l'accent étant porté sur :

2.1.1. Linguistique et informatique :

2.1.1.1. Numéro 14 de LINX, à paraître en mars 1986, intitulé "Langue
et machines". [...]

2.1.1.2. Table ronde sur "Texte et ordinateur", organisé pour
avril 1986.

2.1.2 Problèmes du sens :

Numéro 13 de LINX, à paraître en décembre 1985. [...]

2.1.3 Problèmes du discours :

Numéro 15 de LINX, à paraître en octobre 1986. Le thème de ce numéro
est "Révolution française et fin du 18ème siècle."

2.2. Le Centre de Recherches Linguistiques a proposé à la Commission
Diderot (voir documentation en annexe) de se charger d'une "question
vive" [...]

2.3. Il n'y a rien d'excessif à prétendre que LINX, après son 12ème numéro,
est l'une des revues françaises de linguistique qui jouissent d'une réelle
notoriété [...]
```

Figure 2. Rapport CRL 1985, pp. 3 – 5. Les coupes et l'indentation sont de notre fait.

Ici, l'énumérant explicite [*l*]es deux directions suivantes, suivi de deux points, semble commander les items 2.1. (*Poursuite [...] des différents sous-programmes*) et 2.2. (*[...] se charger d'une "question vive"*), mais l'item 2.1. est un titre qui commande jusqu'à deux autres niveaux de titres numérotés, et l'item 2.2. est un titre qui domine l'équivalent d'une

page de rédaction, organisée en 5 paragraphes et un intertitre non numéroté. Enfin, l'item 2.3. semble s'inscrire dans la structure hiérarchique indépendamment de la commande de l'énumérable, alors même qu'il est un paragraphe de 5 lignes et non un titre.

Contraints de trancher dans le continuum entre macro-structure et listes lors de l'annotation, nous avons considéré que cet exemple ne relevait pas d'une liste, mais plutôt de l'introduction d'une structure de titres relativement libre. Dans de tels cas, aucun critère n'est suffisant ou nécessaire pris isolément ; la catégorisation en liste ou titre est alors le fait de l'annotateur plus que d'une évidence des données.

2.4 Continuum *listes horizontales – plein texte* : décisions d'annotations

Si les listes verticales sont prises dans un continuum avec la macro-structure du document, les listes horizontales semblent quant à elle prises dans un continuum aboutissant au plein texte et à la simple coordination ou juxtaposition syntaxique. Les critères retenus impliquent qu'une liste horizontale doit être détachée du reste de la phrase par des ponctuations noires, tandis que l'énumération ne dispose pas d'une telle démarcation. L'extrait en Figure 3 présente une liste horizontale à la première phrase (*recueil de données, annotation et transformation* etc.), et une énumération hors liste à la seconde phrase (*par leur diversité, leur hétérogénéité* etc.).

[.....] De fait, une part importante des activités a concerné la production de ressources nouvelles : *recueil de données, annotation et transformation des données brutes en observables pour des analyses multifactorielles, mise à disposition et outillage*. Ces ressources, *par leur diversité, leur hétérogénéité, leur multimodalité et leur évolution continue*, interrogent de fait les modèles classiques, et remettent en cause l'autonomie des niveaux de description. [.....]

Figure 3. Rapport Modyco 2018, p. 9. La colorisation est de notre fait.

Contrairement à la liste, l'énumération n'est ici pas détachée immédiatement par une ponctuation (*par* est interposé entre le premier item et la virgule). Elle n'est pas non plus commandée par un *énumérable*, qui ne peut être restitué avant *par* à cause du mouvement syntaxique du groupe prépositionnel placé avant le verbe.

2.5 Outil d'annotation

L'annotation des listes a été effectuée en deux étapes : lors de la phase d'océrisation et de transformation des données en XML, les listes verticales ont été reconnues automatiquement, puis contrôlées par nos soins. Non-automatisable, l'annotation des listes horizontales a pour sa part été effectuée manuellement à l'aide d'une interface HTML locale développée spécifiquement pour autoriser une annotation conviviale. Cette interface permet à l'utilisateur de sélectionner une section de texte et de l'annoter en cliquant sur un bouton correspondant au type de liste, puis de faire de même pour les items.^{iv} Les listes sont catégorisées en quatre types : listes verticales à puce (type « bulleted »), numérotées (« numbered ») ou sans puce ni numéro (« simple »), et listes horizontales (« inline »).

3 Prendre la mesure de la place et des effets de la liste : vers des visualisations textométriques pour l'analyse

Dans cette section, nous présenterons la façon dont les visualisations et analyses statistiques, sur lesquelles s'appuient traditionnellement la textométrie, permettent d'interroger la place et les fonctions ou effets des listes dans notre corpus diachronique.

3.1 Place et distribution des listes, ou rendre compte de la rythmique des listes

Dans notre corpus, les listes représentent 66% du volume textuel (soit près de 100 000 mots) ; seul un tiers du volume textuel est donc rédigé en paragraphes. Le nombre de mots associés au type « background » (correspondant aux paragraphes) ne représente en effet qu'un peu moins de 50 000 mots. Ainsi que représenté en fig. 4, le type de liste le plus fréquent est la liste verticale sans puce ni numéro (annotée « simple ») : ce type représente 43% des mots du corpus, loin devant les listes à puces (« bulleted », 10%), les listes horizontales (« inline », 7%) et les listes numérotées (« numbered », 6%).

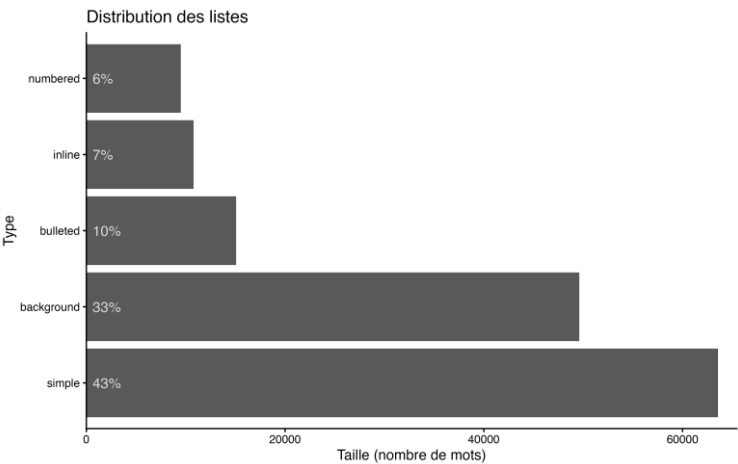


Figure 4. Volume textuel (en nombre de mots) attribué à chaque type de liste et au texte hors liste (background), sur l'ensemble du corpus.

Pour mesurer la variation dans la répartition des types de listes, nous avons calculé le taux de chaque type pour chaque texte et avons projeté ces taux sur un axe diachronique, ce qui produit le graphique en figure 5.

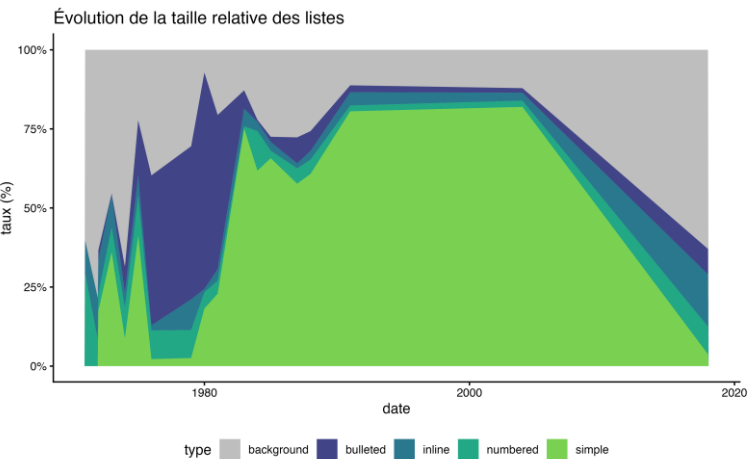


Figure 5. Évolution du volume textuel attribué à chaque type de liste et au texte hors liste, en pourcentage sur le nombre de mots, sur l'ensemble du corpus.

Ce graphique montre que la domination des listes verticales sans puce ni numéro (« simple ») ne vaut que de 1983 à 2004 ; dans les quatre années précédentes, ce sont les listes à puce (« bulleted ») qui dominent, tandis que ce sont les paragraphes qui dominent

dans le reste du corpus. On note donc, dans le dispositif des listes verticales, un net effet de substitution de la liste sans puce ni numéro à la liste à puces, à partir de 1983, tandis que les deux bornes du corpus sont caractérisées par une domination du texte rédigé sans listes. Cette place massive des listes simples est principalement à relier à la présentation des productions des chercheurs dans les rapports, lesquelles sont intégrées aux rapports sous la forme de longues bibliographies de 1983 à 2004.

Aux bornes du corpus, c’est au contraire la rédaction en paragraphes qui domine, notamment parce que les bibliographies sont absentes ou rejetées en annexes ; de telles annexes n’appartiennent pas au rapport lui-même et ne participent pas à son économie. De ce point de vue, le dernier rapport du corpus marque donc un retour à un mode de rédaction qui dominait dans la première moitié des années 1970. Ce retour à la rédaction s’accompagne d’une croissance des listes horizontales, signant la dilution du dispositif d’énumération dans la trame du texte et l’exclusion des listes de productions.

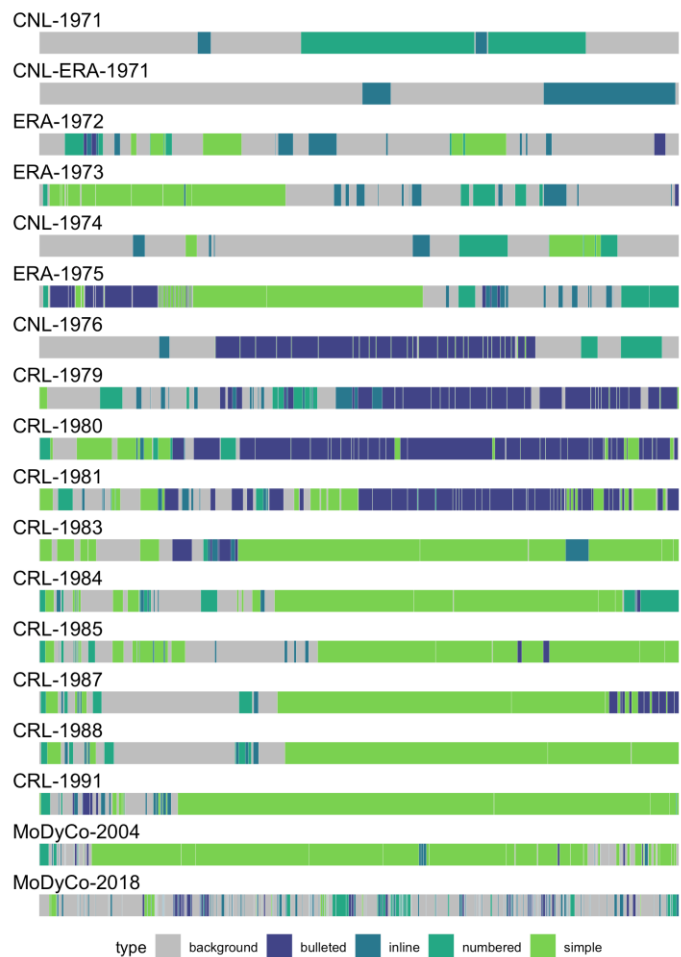


Figure 6. Textogramme montrant la place et la taille des listes dans l’ensemble du corpus.

Une telle saisie quantitative ne rend néanmoins pas compte d’un facteur déterminant pour la fonction des listes : leur place au sein de chaque texte. Pour observer la façon dont les listes viennent rythmer le texte du rapport, nous avons développé ShinyParts,^v une application R permettant la génération de ce que nous avons appelé un *textogramme*, c’est-à-dire un graphique donnant une vue simultanée de la place et la taille des listes au sein du déroulé de

chaque texte du corpus^{vi}. Le textogramme donné en figure 6 permet de saisir non seulement les différences quantitatives discutées ci-dessus, mais aussi la façon dont les listes participent au rythme du texte et à sa structuration micro et macro.

Ainsi, les rapports des années 1983 à 1991 sont caractérisés par la position finale des listes (principalement simples, correspondant notamment aux bibliographies), tandis que les rapports des années 1979 à 1981, où les listes à puce dominent, les répartissent plus équitablement sur l'ensemble du texte. Le dernier rapport du corpus revient à la logique de cette structure, avec des listes réparties sur l'ensemble du texte, mais en réinstaurant la domination d'une trame de plein texte dans lesquelles les listes s'insèrent de manière régulière et discontinue. Notre corpus témoigne donc de deux évolutions, avec le passage en 1975 d'une alternance entre paragraphes et listes à une saturation du texte en listes continues, puis, avec le rapport de 2018, à un retour à une textualité rédigée, où la liste est tissée dans la trame du texte de manière intermittente et non plus continue. Le premier changement, qui est à la fois rythmique (continuité des listes) et quantitatif (saturation du texte) présente des airs de famille avec une textualité administrative, tandis que le second changement, s'il réduit la place brute occupée par les listes, est caractérisé par l'insertion des listes dans un discours à visée et à dimension argumentatives. Cette modalité d'apparition de la liste participe d'une segmentation plus forte de la textualité à mesure qu'on avance dans le temps [22, 16].

3.2 Vers l'analyse des observables lexicaux et grammaticaux spécifiques aux listes

Dans un second mouvement, nous avons souhaité examiner si des marqueurs linguistiques singularisaient la liste (et ses différents sous-types). Après avoir retiré du corpus présenté ci-dessus les bibliographies parfois présentes dans les rapports, nous l'avons partitionné de sorte à mettre en regard les trois agencements textuels que sont les paragraphes, les listes verticales et les listes horizontales, en veillant à garantir l'homogénéité de chaque partition par l'exclusion des cas d'enchaînement.

Nous avons ensuite examiné la façon dont les unités lexicales (segmentées et indexées sous forme de lemmes) et les catégories grammaticales tendent à caractériser chaque agencement textuel de façon contrastive, en nous appuyant sur l'analyse des spécificités [23]. Celle-ci permet de déterminer les unités significativement suremployées (indice positif > +2) ou sous-employées (indice négatif < -2) dans chaque partition.

3.2.1 Dans l'espace du paragraphe : démontrer et justifier l'action passée et projetée du laboratoire

Moins que par les indices attendus d'une complexité (inter)phrastique, le paragraphe du rapport d'activité se singularise en premier lieu par sa forte consistance verbale, que met en évidence un suremploi des observables grammaticaux relevant traditionnellement du pôle verbal (catégorie des verbes et des adverbes), ainsi que de l'ensemble des modes et des temps verbaux. Le paragraphe apparaît ainsi le lieu privilégié de la mise en récit de l'action, au sein duquel les activités sont à la fois rapportées, analysées et projetées. C'est un rapport singulier à la temporalité qui le sépare des listes : le paragraphe rend compte non seulement de ce qui a été fait, mais également de la manière dont l'activité a été pensée, organisée et s'est déroulée dans une trajectoire temporelle. La temporalité a pour repère le temps de l'énonciation, comme le confirment les indices d'un ancrage déictique (*maintenant* et *actuellement* : +3), ainsi que le primat du présent (+53) et le suremploi du passé composé (+27).

Dans l'espace du paragraphe, se déploient les marques de la première personne du pluriel, dont le pronom personnel *nous* (+11), témoignant d'un fort sous-emploi dans les listes (-5 dans les listes verticales ; -6,1 dans les listes horizontales). Dans le paragraphe, le *nous*,

renvoyant principalement à un collectif de chercheurs que l'entité institutionnelle fédère et permet de faire parler à l'unisson, cohabite également avec d'autres noms communs collectifs (*équipe* : +8,9 ; *centre* : +4 ; *laboratoire* : +4) que l'on remarquera être combinés à des constructions pronominales réfléchies (*se développer*, *se structurer*, *se donner les moyens*) qui expliquent le suremploi du lemme *se* (+13). Significativement suremployées dans le paragraphe, les constructions pronominales, lorsque associées aux labels collectifs, peuvent légitimement être interprétées comme les traces d'une entreprise de réflexivité par laquelle le laboratoire met en relief sa capacité à se diriger *de lui-même*. Soulignons néanmoins que ce type de construction est fréquent dans l'écriture scientifique, qui privilégie des structures impersonnelles afin de mettre à distance les objets et les méthodes qu'elle décrit (par exemple : *il s'agit de*, *la recherche s'appuie sur*).

Ce sont enfin les marques d'une visée argumentative qui apparaissent suremployées dans le paragraphe, comme le suggèrent, d'une part, les suremplois des adverbes de négation (*ne* : +7,5 ; *pas* : +3,9) dont le fonctionnement est prioritairement restrictif comme dans l'extrait ci-dessous et d'autre part, les suremplois de différents connecteurs (*mais* : +6 ; *aussi* : +5 ; *ainsi* : +5 ; *donc* : +3) orientés vers la légitimation de l'activité et, souvent, vers la demande de moyens.

Ce programme de recherche qui se distribue en plusieurs étapes et s'échelonne sur plusieurs années ne peut se réaliser qu'avec le concours permanent des membres de l'équipe. Parmi ceux-ci, les étudiants de 3ème cycle constituent un groupe important. [...] Le centre d'étude de la néologie lexicale demande donc des crédits qui puissent leur être allouée pour les frais d'enquête propres à leur recherche particulière, les frais de dactylographie, etc ... (rapport ERA 1972)

Notons que les marqueurs d'opposition, dans les paragraphes du rapport, présentent des profils discrets (*cependant* : +2) et que le connecteur *mais* n'y opère pas comme un marqueur d'opposition mais comme un connecteur discursif additif corrélatif (*mais aussi*) structurant une énumération.

3.2.2 Dans l'espace de la liste : montrer les produits de l'activité et performer sa capacité à la mener et l'organiser

D'emblée, notons que très peu de catégories grammaticales présentent des indices excédentaires significatifs dans les listes, qu'elles soient verticales ou horizontales. Ce silence statistique est en soi signifiant et confirmerait que la liste relève d'un régime d'économie discursive où la valeur et les liens des éléments listés tiennent à leur simple présence sous la forme même de l'énumération.

Les listes présentent une consistance essentiellement nominale comme en attestent le suremploi des noms communs (liste verticale : +4 ; horizontale : +8) et des noms propres (liste verticale : +29 ; horizontale : +159). Par opposition à la dynamique verbale observable dans le paragraphe tourné vers l'action, les listes sélectionnent, énumèrent et cadrent des objets présentés comme le résultat stabilisé d'une activité.

Dans le cas des listes verticales, cette logique de monstration des « produits » de la recherche apparaît de manière particulièrement saillante. L'analyse des catégories grammaticales montre que cet espace textuel est avant tout dédié à la comptabilité des produits de la recherche, comme le suggèrent les indices de spécificité des données chiffrées (+40). La liste verticale est le lieu privilégié de l'énumération des publications, comme en témoignent les suremplois des noms propres et des guillemets, traces des normes bibliographiques ; elle opère comme un espace de preuve d'une activité accomplie : elle atteste l'existence d'un travail accompli, en donnant à voir (et à compter) des produits stabilisés (*thèse* : +6 ; *colloque* : +4 ; *communication* : +2).

Dans l'espace de la liste verticale peuvent néanmoins être confondus de façon stratégique des produits finis, des événements et des processus, comme dans l'extrait ci-

contre. Cadré par la mise en série qui l'intègre à une suite de « réalisations » hétérogènes (« plusieurs sortes »), le processus s'impose comme un résultat, effet renforcé par la nominalisation (« l'élaboration ») :

Les problèmes liés à la didactique des langues étrangères ont donné lieu à plusieurs sortes de réalisations :

- d'abord à la publication du N° 57 de la revue *Langages* (Larousse) [...]
- ensuite, la tenue d'un Colloque à l'Université de PARIS VIII VINCENNES [...].
- enfin, l'élaboration du programme triennal prioritaire a largement bénéficié des réflexions du groupe de travail attaché à la didactique des langues étrangères (R. PORQUIER, E. KOSKAS, notamment). (Rapport CRL 1980, p. 3)

Espace d'un inventaire duquel l'énonciateur efface les marques de sa présence, la liste verticale constitue paradoxalement un agencement où se (per)forme, dans la dimension verbo-visuelle, sa capacité organisationnelle. Savoir sélectionner, regrouper et ordonner des éléments pertinents, c'est aussi montrer que cette activité est maîtrisée, structurée et que le laboratoire l'organise. La liste fonctionne ainsi comme un opérateur de légitimation, au service d'une présentation efficace et normée du travail scientifique et de la structure qui la gère.

3.2.3 Dans les listes horizontales : suggérer, par une logique extensive, l'intensité de l'activité et la performance

Les lignes horizontales se distinguent des listes verticales principalement par un suremploi de la ponctuation suspensive, qui leur est statistiquement spécifique (+4,6). Opérateur d'extension, la ponctuation suspensive permet d'additionner des unités sans clore le sens. Le lemme *etc*, spécifique aux listes horizontales (+20), semble confirmer que la liste horizontale s'inscrit dans une dynamique d'accumulation continue. Ainsi, la liste horizontale suggère l'intensité de l'activité moins par ce qu'elle montre que par ce qu'elle laisse entendre. Par la multiplication des items, l'ouverture des séries et l'absence de hiérarchisation explicite, elle performe une image de productivité élevée et de compétence étendue.

La logique extensive à l'œuvre dans les listes horizontales apparaît confirmée par le suremploi de la classe des adjectifs (+19). L'analyse des types d'adjectifs à l'origine de cet excédent grammatical permet d'observer que les listes horizontales sont principalement introduites par des adjectifs subjectifs évaluatifs comme *différent* (+4), *varié* (+3), *divers* (+2) : ces derniers ont en commun d'évaluer moins des propriétés intrinsèques que le nombre et l'hétérogénéité des éléments convoqués. Placés généralement en position d'amorce de la liste horizontale, ces termes orientent la lecture vers une attente d'abondance et de pluralité avant même que les items ne soient explicitement donnés. L'adjectif fonctionne alors comme un opérateur de cadrage, annonçant une série dont la valeur tient précisément à sa capacité à agréger des éléments multiples et non strictement homogènes. La focale est alors mise sur la dynamique même de la série, qui performe une activité hautement productive, étendue et difficilement réductible à une énumération close.

L'extrait ci-dessous illustre la logique extensive soutenue par les listes horizontales, dans le cadre d'une justification des enjeux d'un projet scientifique, porteur d'une « avancée significative » qu'explicite le paragraphe final. Dès la première ligne de l'extrait, l'amorce *un certain nombre* installe un cadrage quantitatif flou mais orienté vers l'intensité de l'activité. Elle prépare l'apparition d'une liste exhaustive et étendue de langues. Une seconde liste horizontale énumère des unités linguistiques, en position finale desquelles apparaît un nouvel indicateur de non-clôture (*par exemples*), suggérant cette fois que les unités linguistiques observées dépassent celles qui sont effectivement listées – on remarquera d'ailleurs la marque fautive du pluriel. Une troisième liste, numérotée, met en exergue les enjeux de la méthodologie déployée : tandis que la numérotation surmarque la distinction des

items, la locution additive *mais aussi* insiste sur leur articulation, et le dépassement des pratiques scientifiques traditionnelles. Condensant dans un espace réduit une multiplicité de dimensions, d'objets ou de niveaux d'analyse, les listes horizontales participent à suggérer (plutôt qu'à démontrer) la richesse et la densité d'une activité à la mesure des ambitions posées dans le dernier paragraphe (*repousser les limites*).

Un certain nombre de langues ont été explorées du point de vue de l'hétérogénéité de leurs grammaires (français, anglais, russe, hébreu, indonésien, haoussa, japonais, inuktitut, hindi, khmer ancien, sanskrit).

De plus, selon l'hypothèse que les pratiques langagières agissent sur les grammaires, il devient fondamental de prendre en considération la diversité de ces pratiques langagières dans toute opération de modélisation de la langue. Sur ce point précis, le pôle [A] s'appuiera sur les savoir-faire de l'équipe [B] pour repérer dans des corpus des variations (entre autres, sociolinguistiques) inhérentes aux dynamiques des langues en utilisation, mais aussi pour annoter dans les corpus la distribution des unités de langue (morphèmes, intonèmes, patrons lexico-syntaxiques, structures prédicatives, formules discursives, par exemples [sic]). Cette approche linguistique qui prend en considération les variations langagières permet de réaliser, à un niveau de granularité fin, la modélisation 1) de faits de langues (pôle [B]), mais aussi 2) du fonctionnement des mécanismes langagiers (pôle [C]).

En nous encourageant à intégrer l'hétérogénéité des faits linguistiques dans notre démarche de modélisation, les chercheurs du pôle [A] contribueront à repousser les limites de validité écologique de nos modèles. [...]. Le dialogue constant entre sociolinguistique, linguistique de corpus et psycholinguistique incarné dans le projet de recherche porté par l'unité de recherche Modèles, Dynamiques, Corpus contribue dans ce sens une avancée significative dans le paysage actuel. (Rapport MODYCO 2018, p. 12, la numérotation alphabétique des pôles est nôtre)

La liste horizontale vient également à l'appui de séquences évaluatives [25] mettant en jeu des catégories du prescripteur (rayonnement, internationalisation, visibilité, par exemple), ce qu'illustre l'extrait suivant, dans lequel la liste, par l'agrégation d'éléments de spécification de différents ordres (formation à la recherche, montage de projet, relations avec des collègues à l'international) soutient la catégorisation « visibilité au niveau international » (reformulée en « particulièrement actif à l'international »). La double mise en exergue et en série d'éléments quantifiés perforce ainsi la qualité exprimée : « être particulièrement actif à l'international ».

Au niveau international, le laboratoire a su améliorer sa visibilité, notamment en encourageant les collaborations internationales dont certaines sont nourries par des projets internationaux financés ou des cotutelles de thèses de doctorat. Le laboratoire est particulièrement actif à l'international : 24 thèses en co-tutelles, 25 projets financés, et 3 accueil [sic] de chercheurs étrangers. (Rapport Modyco 2018, p. 7)

4 Conclusion

Nous avons montré que la liste, qu'elle soit verticale ou horizontale, constitue, dans le rapport d'activité de laboratoire, un agencement textuel participant à la visée et à la dimension argumentatives du genre. Les listes contribuent en effet à montrer les produits de la recherche, à structurer l'action passée et projetée, et à construire l'image d'un laboratoire organisé, productif et légitime.

Au terme de ce chantier, nous espérons avoir montré en quoi la liste constitue un objet pertinent pour une analyse du discours et une analyse textuelle aux prises avec l'élaboration d'une théorie des paliers de structuration (Adam, 2018). Nous illustrons ainsi l'intérêt méthodologique, en analyse outillée du discours, de prendre en compte la matérialité verbo-visuelle pour une meilleure connaissance des genres de discours. Une telle ouverture est destinée à être soutenue par le développement d'outils ergonomiques facilitant l'annotation et de visualisations textométriques permettant, comme le textogramme, d'observer finement

la texture de nos données et de les éclairer au prisme de la lecture différentielle qu'autorise le corpus en tant que mise en série.

Références

1. J.-M. Adam, *Le paragraphe : entre phrases et texte* (Armand Colin, Paris, 2018)
2. A. Rabatel, Listes et effets-listes. Énumération, répétition, accumulation. *Poétique* **167**(3), 259-272 (2011)
3. S. Milcent-Lawson, M. Lecolle, *Liste et effet liste en littérature* (Garnier, Paris, 2013)
4. R. Mahrer, La méthode liste. Textualité et créativité. *Genesis* **47**, 13-33, (2018)
5. M. Chandelier, P. Païssa, Enjeux argumentatifs des structures énumératives longues : analyse d'un corpus de tribunes consacrées à la biodiversité. *Argumentation et Analyse du Discours* **29** (2022).
6. M. Bakhtine, *Esthétique de la création verbale* (Gallimard, Paris, 1984)
7. M.-A. Paveau, L. Rosier, Grammaire de la liste. In *Le Sens en marge.*, Evrard I., Pierrard M., Rosier L., van Raemdonck D. (dir.), 113-133 (L'Harmattan, Paris, 2009)
8. A. Béguin-Verbrugge, *Images en texte, images du texte : dispositifs graphiques et communication écrite* (Presses universitaires du Septentrion, Lille, 2006)
9. J. Lefèvre, *La note de bas de page : dans les imprimés contemporains. Études linguistiques et textuelles* (Lambert-Lucas, Limoges, 2022)
10. R. Amossy, *L'argumentation dans le discours* (Nathan Université, Paris, 2000)
11. S. Lehmann, V. Lethier, É. Née, Vers l'intégration d'un nouveau régime en analyse du discours, le format. Le cas du rapport de laboratoire à l'université de Nanterre (1970-2018), *Communication et langages* **220**, 2024. <https://doi.org/10.3917/comla1.220.0061>
12. C. Facq-Mellet, S. Lehmann, É. Née, La construction discursive de l'ethos dans les rapports de laboratoire de l'université de Nanterre (1970-1989) : analyse comparative de deux profils énonciatifs. In *Travaux de linguistique du Cerlico* (Presses Universitaires de Rennes, Rennes, 2026)
13. É. Née, C. Oger, F. Sitri, Le rapport : opérativité d'un genre hétérogène. *Mots. Les langages du politique* **114** (2017). <https://doi.org/10.4000/mots.22752>
14. D. Maingueneau, Les unités topiques (chap. 6), in *Discours et analyse du discours : Une introduction*, 58-74 (Armand Colin, Paris, 2024)
15. C. Barats, É. Née, Structurer, encadrer et lisser l'évaluation : les écrits de l'AERES/HCERES. *Communication et langage* **203**, 5-25 (2020).
16. É. Née, S. Lehmann, Le rapport scientifique à l'université Paris Nanterre (1970-2018) : émergence, institutionnalisation et transformation d'un genre, *Cahiers de praxématique* **78** (2022). <https://doi.org/10.4000/praxematique.8256>
17. J. Rebeyrolle, M.-P. Péry-Woodley, Énumération et structuration discursive. In *Actes du CMLF 2014*, 3183-3196 (2014). <https://doi.org/10.1051/shsconf/20140801217>
18. U. Eco, *Vertige de la liste* (Flammarion, Paris, 2009).

19. C. Angotti, P. Chastang, V. Debiais, L. Kendrick (dir.), *Le pouvoir des listes au Moyen Âge I* (Éditions de la Sorbonne, Paris, 2019).
<https://doi.org/10.4000/books.psorbonne.54932>
20. M. Favriaud, Les problèmes de ponctuation générale soulevés par la poésie contemporaine. *Pratiques* **179-180** (2018). <https://doi.org/10.4000/pratiques.5101>
21. C. Luc, M. Mojahid, M.-P. Péry-Woodley, J. Virbel, Les énumérations : structures visuelles, syntaxiques et rhétoriques. In *Document électronique dynamique : actes du troisième colloque international sur le document électroniques (4-6 juillet 2000, Lyon)*, Gaio M. et Trupin É. (dir.), 21-40 (Europia, Paris, 2000)
22. A. Cormier, De l'objet textuel à l'énonciateur du texte. *Cahiers de praxématique* **78** (2022). <https://doi.org/10.4000/praxematique.8144>
23. P. Lafon, Sur la variabilité de la fréquence des formes dans un corpus. *Mots. Les langages du politique* **1**, 127-165 (1980). <https://doi.org/10.3406/mots.1980.1008>
24. P. Hamon, La mise en liste. In [3], 21-29.
25. É. Née, F. Sitri, M. Veinard, S. Fleury, Routines discursives et séquentialité dans des écrits professionnels : la mise au jour d'une séquence évaluative ? *Corpus* **17** (2017).
<https://doi.org/10.4000/corpus.2880>
26. U. Eco, Mes listes (chap. 7). In *Confessions d'un jeune romancier*, 139-231 (Grasset, Paris, 2011)
27. J. Lemarié, (Comprendre) les énumérations. In [19], 15-26
28. M. Colas-Blaise, Dynamiques de la mise en liste. Une approche sémio-linguistique. In [3], 33-44
29. J. Goody, *La raison graphique : la domestication de la pensée sauvage* (Éditions de Minuit, Paris, 1979)

ⁱ De l'anthropologie de l'écrit [29], à la stylistique [3] ou la linguistique de corpus [17], en passant par la sémiotique [18, 26, 28], la psycholinguistique [27], l'histoire et la littérature [19].

ⁱⁱ Initié en 2017 par Frédérique Sitri (UPEC/Céditec) et financé par le Labex *Les passés dans le présent*, ce projet a pour objectif de saisir les transformations de l'université depuis les années 1970 à partir de deux genres de discours produits par cette institution, les comptes rendus de conseil d'administration et les rapports de laboratoire.

ⁱⁱⁱ Le *Centre de néologie lexicale* (CNL) créé par Louis Guilbert en 1971, le *Centre de recherches linguistiques* (CRL), qui succède au CNL et qui est créé en 1979 par Michel Arrivé, puis le laboratoire *Modyco*, créé en 2001 par Bernard Laks et succédant au CRL.

^{iv} L'outil est disponible, sous licence libre, à l'adresse : <https://github.com/TimotheePremat/LisTEI/>

^v L'outil est disponible, sous licence libre, à l'adresse : <https://github.com/TimotheePremat/ShinyParts>

^{vi} Cette démarche s'inspire directement de la fonctionnalité « carte des sections » conçue et implémentée par André Salem dans l'un des premiers logiciels de lexicométrie, Lexico (<https://lexico.com/Produits.html>), et proposant des visualisations topologiques du texte.