

DIACHRONIE COMPARÉE DE FORMES MÉSO ET MACRO-SÉMANTIQUES DANS LE CORPUS GILLES DELEUZE

Sylvain LOISEAU
MoDyCo, Université Paris X

SOMMAIRE

1. Variables utilisées
 2. Analyse factorielle au palier des textes
 3. Du texte au paragraphe
 4. Comparaison des structures internes des textes
- Conclusion

Résumé : *Alors que les corpus richement annotés renouvellent la description sémantique, on souhaite dans cette contribution utiliser ce nouveau type d'objet empirique pour décrire la diachronie de formes sémantiques. Deux types de formes sémantiques en particulier sollicitent notre attention : le thème sémantique d'une part, le genre d'autre part. On essayera de décrire les traits caractéristiques de leur évolution, et l'articulation de leur évolution. Le corpus étudié comprend l'intégralité des commentaires philosophiques et des essais de Gilles Deleuze. L'empan diachronique s'étend de 1953 à 1994, c'est-à-dire de la philosophie d'après-guerre jusqu'à l'apparition de la génération philosophique d'après 1968. On s'attachera d'abord à justifier une périodisation qui distingue trois pôles : l'académisme des années 50, le militantisme des années 60 et 70, et le retour à un certain classicisme à partir des années 80. De nombreux traits contribuent à justifier cette périodisation, aussi bien dans les systèmes thématiques et dialogiques que, à un autre niveau, à travers des traits morphosyntaxiques et stylistiques.*

À partir d'une caractérisation des textes relevant de chacun des deux genres, le commentaire et l'essai, sur cet axe diachronique, la comparaison de la diachronie des genres permet d'observer des évolutions corrélées. La période centrale est particulière du fait de la disparition apparente du genre du commentaire de la production de Deleuze. On montrera comment les deux genres ont temporairement été confondus, dans un projet de renouvellement et de sommation des formes textuelles. Enfin, on s'attachera à montrer la corrélation d'une forme théorique caractéristique de la période du début des années soixante aux formes textuelles dans lesquelles elle s'inscrit. Le concept de différence est caractéristique d'une remise en cause de l'ontologie par la philosophie de l'après-guerre d'Algérie, commun notamment à Deleuze (Différence et répétition, 1969), Derrida (L'Écriture et la différence, 1979), et Lyotard (Le Différend, 1983). Sur le sous-corpus constitué de ces trois textes on exploitera les corrélations entre les différents systèmes linguistiques annotés pour montrer l'apparition et la construction d'une thématique partagée, centrale pour l'histoire des idées.

Introduction

Les corpus annotés peuvent-ils aider l'analyse diachronique ? Dans le cadre d'une recherche de méthodes, on souhaiterait faire porter cette question sur un point précis : celui de l'articulation de la description diachronique de formes sémantiques à des paliers d'analyse différents¹.

Le corpus étudié comprend l'intégralité des commentaires philosophiques et des essais de Gilles Deleuze. L'empan diachronique s'étend de 1953 à 1994, c'est-à-dire de la philosophie d'après-guerre à l'apparition de la génération philosophique d'après 1968 (les références des textes du corpus sont données en annexe).

¹ Nous remercions Céline Poudat pour son aide.

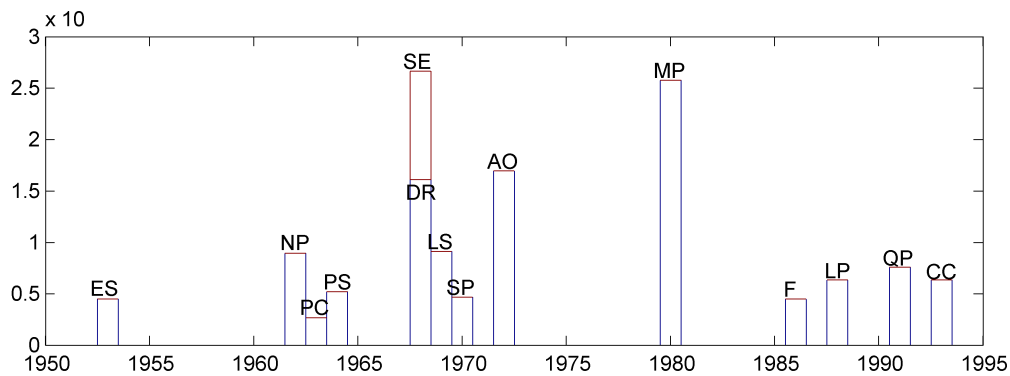


Figure 1 : distribution chronologique et taille (en nombre de mots) des textes du corpus

On se limitera dans les résultats présentés ici à l'utilisation de variables morphosyntaxiques pour se concentrer sur l'observation de l'évolution du système dialogique des textes du corpus. L'observation portera sur deux paliers différents : celui des textes et celui des paragraphes, afin d'essayer de décrire l'évolution de la structure globale du corpus et de la structure interne des textes.

1. Variables utilisées

Dans cette analyse on s'attache à l'observation des textes du corpus sur des critères essentiellement morphosyntaxiques. On a privilégié des variables reflétant le système verbal (temps et personne). Ont été ajoutés les ponctuations, les types de noms (singuliers, pluriels, ou propres), et l'opposition fini/indéfini (relevé sur les pronoms et les déterminants). Enfin on a également inclus les effectifs des notes, citations, mises en forme, titres et segmentations des textes en sections et sous-sections. De nombreux tests ont permis de montrer que l'hétérogénéité de ce jeu de variables ne fait pas sensiblement baisser la qualité de leur regroupement en facteurs : les deux premiers facteurs résument 51% de la variance quand sont utilisés seulement les temps et les personnes, tandis qu'ils résument 52% de la variance sur l'ensemble de ce jeu de descripteurs.

Les variables ont été constituées d'une part à partir d'un balisage des propriétés structurelles du corpus et d'autre part à partir d'une annotation morphosyntaxique du logiciel Cordial, importée dans l'annotation existante¹. Les étiquettes de Cordial ont été décomposées, dans le corpus, en traits minimaux (partie du discours, genre, nombre, caractère fini ou indéfini d'un pronom ou d'un déterminant, etc.). Les descripteurs utilisés ont été constitués « à façon », par un assemblage de ces traits. Par exemple, on a utilisé comme descripteur « noms communs singuliers », par un assemblage de plusieurs traits issus de l'analyse de Cordial (« Nom commun singulier » est la cooccurrence du trait « nom » et du trait « singulier » sur un mot).

Ce jeu de variables est sémiotiquement hétérogène : il relève à la fois de la syntaxe (à travers les temps et les personnes), de propriétés éditoriales ou macrostructurelles (segmentation en paragraphes, notes, etc.), et d'un niveau peut-être « rhétorique », à travers les ponctuations. Les paliers observés sont également très hétérogènes : les segmentations ou les notes relèvent du macrosémantique, tandis que l'opposition entre défini et indéfini par exemple est observée au palier des mots. Les ponctuations relèvent peut-être davantage du palier mésosémantique (palier de la période)².

Enfin, ce type de descripteurs pose des difficultés particulières. Les métriques sont très différentes : la moyenne la plus élevée est de 73,63 (nom singulier), tandis que la moyenne la plus faible est de 0,13 (point d'exclamation). Mais surtout, ces descripteurs forment des sous-ensembles (temps, personnes), et il est intéressant de pouvoir prendre en compte le caractère structurel de ces sous-ensembles, plutôt que de traiter uniformément les fréquences. À cette fin,

¹ Cette importation, ainsi que l'extraction des fréquences des traits, a été réalisée avec le logiciel CorpusReader (Loiseau 2006)

² Les descripteurs sont préfixés, dans cet article, par des accolades qui indiquent grossièrement leur origine sémiotique : {gram} indique un descripteur grammatical, {lm} un lemme, {ff} une forme fléchie, {c} un caractère de ponctuation, {tag} un élément de structuration. Ces préfixes ne prétendent cependant pas établir une typologie du système sémiotique ; ils reflètent originellement l'organisation matérielle de l'annotation du corpus.

on a utilisé la méthodologie établie par Céline Poudat (Poudat 2006), qui consiste à soumettre à l'analyse factorielle non pas des fréquences (absolues ou normées), mais des pourcentages dans différentes catégories comprenant les variables en distribution complémentaire. Ainsi, la valeur de la variable « imparfait » dans chacun des textes est le pourcentage des imparfaits dans l'ensemble des temps.

Pour les traits correspondant à des marques de segmentation (note, citation, titre, division, mise en forme), on n'a cependant pas pu opérer de regroupement en classes, ces variables n'étant pas en distribution complémentaire. On a alors utilisé un pourcentage calculé par rapport à la longueur des textes (mesurée en nombre de mots). L'interprétation des variables en dépend : la variable « nombre de notes » indique la « densité » en notes, tandis que la variable « nombre de paragraphes » indique davantage la longueur moyenne des paragraphes, puisque les textes sont entièrement segmentés en paragraphes. L'hétérogénéité du jeu d'étiquettes se répercute sur l'interprétation qu'elles permettent.

En résumé, le jeu d'étiquettes utilisé se répartit comme suit :

Catégorie	Etiquette	Nom
type de noms	{gram}ty.p	Nom propre
	{gram}nom-s	Nom commun singulier
	{gram}nom-p	Nom commun pluriel
ponctuations	{c}?	Point d'interrogation
	{c}!	Point d'exclamation
	{c}.	Point
	{c},	Virgule
	{c}...	Points de suspension
	{c}:	Deux-point
	{c};	Point-virgule
Verbes	{gram}t.p	Présent
	{gram}t.i	Imparfait
	{gram}t.pas	Passé simple
	{gram}t.fut	Futur
	{gram}t.subpre	Subjonctif présent
	{gram}t.subimp	Subjonctif imparfait
	{gram}t.con	Conditionnel
	{gram}t.imp	Imparfait
	{gram}t.partpas	Participe passé
	{gram}t.partpres	Participe présent
Personne en lemme	{lm}je	Lemme <i>je</i>
	{lm}tu	Lemme <i>tu</i>
	{ff}il	Forme fléchie <i>il</i>
	{lm}nous	Lemme <i>nous</i>
	{lm}vous	Lemme <i>vous</i>
	{ff}ils	Forme fléchie <i>ils</i>
	{lm}on	Lemme <i>on</i>
fini ou indéfini	{gram}fin.d	Traits défini (sur un pronom ou un déterminant)
	{gram}fin.i	Traits indéfini (sur un pronom ou un déterminant)

Tableau 1 : Descripteurs regroupés en classes

{tag}note	Nombres de note
{tag}hi	Nombre de mise en forme (gras, italique, ...)
{tag}head	Nombre de titre
{tag}div	Nombre de divisions et subdivisions
{tag}q	Nombre de citations
{tag}p	Nombre de paragraphes

Tableau 2 : Descripteurs hors classe

2. Analyse factorielle au palier des textes

Les variables ainsi calculées¹ sont soumises à une analyse en composante principale dans le logiciel DTM². Les valeurs propres des premiers axes représentent un pourcentage très

¹ La manipulation des variables s'est effectuée à l'aide du logiciel Matlab.

² (Lebart 2006)

satisfaisant de la variance totale. La stabilité et l'individualisation des axes issus des descripteurs sélectionnés sont confirmées par la longueur et le faible recouvrement des intervalles d'Anderson correspondants.

!	1	2	3	4	5	!
!	10.8796	6.7282	4.1454	2.9906	2.4524	!
!	32.00	19.79	12.19	8.80	7.21	!
!	32.00	51.79	63.98	72.78	79.99	!

Figure 2 : valeurs propres des cinq premiers axes

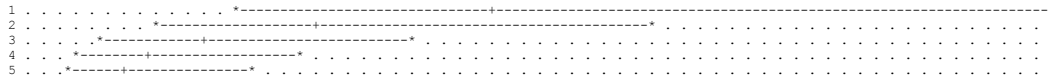


Figure 3 : représentation des intervalles d'Anderson des cinq premiers axes

La distribution des textes sur le premier facteur restitue notablement l'axe diachronique du corpus : l'origine de l'axe oppose les textes sur la coupure des années 1970. Les exceptions apparentes à cette distribution chronologique sont *Proust et les signes* (PS) d'une part, qui se positionne du côté des textes postérieurs, et *Spinoza et le problème de l'expression* (SE) d'autre part, qui se positionne avant *Différence et répétition* (DR). Dans le premier cas, on peut rapprocher le caractère irrégulier de la position de *Proust et les signes* (PS) de l'originalité de son objet dans le corpus. Dans le second, le recul sur le premier axe de *Spinoza et le problème de l'expression* (SE), est en partie un artefact de la numérotation des textes, puisque ce livre est paru la même année que *Différence et répétition* (DR) (et est issu, comme lui, d'une thèse de troisième cycle). La différence cependant sensible de leur position peut sans doute être rapportée à l'opposition entre genre du commentaire et genre de l'essai : *Différence et répétition* formule le programme qui sera réalisé dans les essais postérieurs. Les quatre derniers textes sont remarquablement regroupés. On peut noter, pour terminer, que l'opposition entre essais et commentaires est peut-être « en germe » dans le second axe : l'opposition entre les deux genres est contrariée seulement par la position de *Nietzsche et la philosophie d'une part* (NP), et celle de *Qu'est-ce que la philosophie ?* (QP) d'autre part.

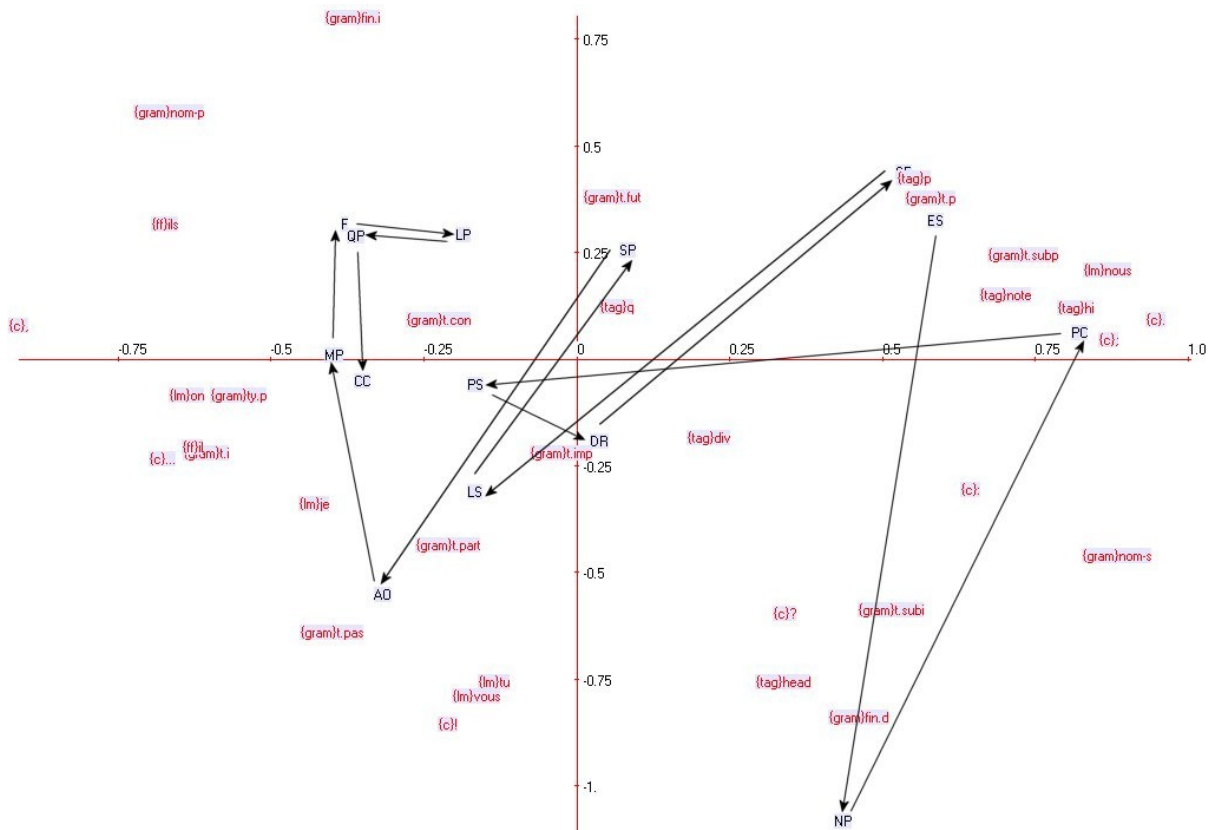


Figure 4 : ACP des textes du corpus. Les textes sont reliés dans l'ordre chronologique

Le premier axe semble organiser les variables autour d'une opposition entre un plus ou moins grand académisme : les marques de structuration des textes, les deux temps du subjonctif présent et passé, le *nous*, les ponctuations complexes comme les deux-points et le point-virgule, s'opposent notamment aux personnes *je*, *tu* et *vous*, aux points de suspension et d'exclamation. L'arbre de classification¹ donne une représentation plus fine de la proximité entre textes. On observe que la plus grande distance est entre les quatre premiers commentaires du corpus et les textes suivants. Si l'on coupe cet arbre pour obtenir quatre classes², les regroupements constitués respectent donc l'axe chronologique : seuls *Spinoza et le problème de l'expression* (SE) et *Proust et les signes* (PS) perturbent cette répartition, puisqu'ils sont inclus dans une classe alors qu'ils sont situés chronologiquement entre deux textes d'une autre classe. Les perturbations de l'axe chronologique semblent donc relever d'une opposition entre genres.

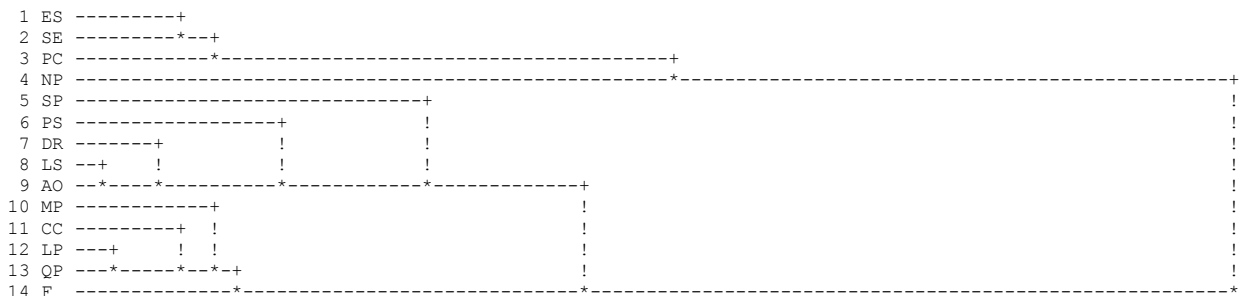


Figure 5 : Dendrogramme des textes

La classification des textes en quatre classes, issue de ce dendrogramme, permet d'observer, plus finement que sur les deux premiers axes factoriels, les variables spécifiques de chaque groupe.

¹ Classification opérée avec DTM, utilisant la méthode des voisins réciproques.

² Au-delà de quatre classes, chaque classe aurait moins de trois individus en moyenne dans l'hypothèse d'une répartition idéale des individus.

Classe	Individus-textes	Variables positives ¹	Variables négatives
Classe 1	ES, PC, SE	points, points virgules, notes, mises en formes, <i>nous</i> et subjonctif présent.	virgule, points de suspension, <i>il</i> , <i>on</i> , noms propres, et passé simple.
Classe 2	NP	Pas de variables caractéristiques.	Pas de variables caractéristiques.
Classe 3	PS, DR, LS, SP, AO	Participe passé, passé simple.	Futur.
Classe 4	MP, F, LP, QP, CC	<i>On</i> , indéfini, noms propres, virgule, conditionnel, points de suspension.	Subjonctif imparfait, noms communs singuliers, défini, nous, points-virgules.

Tableau 3 : les quatre classes issues d'une classification hiérarchique et leurs variables caractéristiques

Restituant en grande partie l'ordre chronologique, l'analyse factorielle permet d'une part de proposer une segmentation diachronique du corpus, et d'autre part de caractériser chaque sous-ensemble.

Les bornes temporelles des trois classes sans singleton sont 1953 et 1968 (15 ans) pour le premier, 1964 et 1972 (8) pour le second, et 1980 et 1994 pour le quatrième (14). Ces bornes sont en recouvrement partiel. Elles sont hétérogènes en genre, à l'exception de la première.

Les oppositions des descripteurs se concentrent principalement entre les classes 1 et 4 – les classes les plus éloignées chronologiquement – qui regroupent respectivement les premiers et les derniers textes du corpus. Du point de vue de la ponctuation, les points et points-virgules s'opposent aux virgules et aux points de suspension ; du point de vue des personnes, le *nous* s'oppose au *on* ; du point de vue des temps, le subjonctif présent s'oppose au conditionnel ; enfin, le défini s'oppose à l'indéfini. Ces oppositions semblent pouvoir être regroupées dans une opposition plus générale entre un régime plus assertif, et un régime moins académique, notamment par l'indétermination de la personne (*on*), la suspension des phrases (points de suspension), et le conditionnel. Si l'on revient à l'examen du plan factoriel, d'autres variables corroborent cette interprétation : la présence de toutes les marques de structuration (titres, note, division) du côté des textes de la première classe, davantage structurés et organisés.

Les textes de la classe 3 sont faiblement caractérisés. Le plan factoriel, plus riche mais moins sûr, permet de leur associer notamment le point d'exclamation, le *je*, le *tu* et le *vous*. Il oppose fortement le *je* et le *nous*, tandis que les marques de citation sont centrées : cette variation n'est pas due seulement à une variation dans le nombre des citations². L'hétérogénéité des variables utilisées permet peut-être de mieux contrôler l'interprétation d'une variable grâce à une autre.

Les variables utilisées permettent donc de dégager des ensembles de traits en opposition entre les textes corrélés à l'opposition sur l'axe chronologique. Ils ne permettent pas cependant de mener l'analyse à un niveau de précision supérieur à une opposition entre textes académiques (première période), et moins académiques (dernière période). Les textes de la période centrale semblent cependant se détacher à travers des variables, en particulier de personne, qui indiquent un régime dialogique très spécifique.

3. Du texte au paragraphe

La répartition obtenue est-elle similaire si l'on utilise comme individus non plus les textes, mais les paragraphes ? Afin de comparer ces deux paliers, et de prendre en compte plus finement l'hétérogénéité interne des textes, les paragraphes ont été utilisés comme individus, décrits par les mêmes variables que précédemment, et soumis à la même analyse factorielle. Les variables hors classe « nombre de paragraphes », « nombre de titres » et « nombre de divisions » ont été retirées, puisque les individus ne contiennent plus ces traits structurels. La variable « nombre de citation » a également été retirée pour des raisons pratiques : cette variable est manquante pour certains textes et la notation de son absence est plus difficile sur des individus nombreux.

Les valeurs propres des premiers facteurs sont, cette fois, beaucoup plus faibles (9%, 8%, 6%, 6%). Les intervalles d'Anderson suggèrent que les premiers facteurs sont peu stables et se

¹ On ne garde pour les variables positives et négatives que les variables d'une valeur test supérieure à 2.

² Elle peut naturellement être due à une variation dans leur contenu.

recouper partiellement. Les cinq premiers facteurs au moins semblent nécessaires à une restitution acceptable des variables initiales.



Figure 6 : Représentation des intervalles d'Anderson des cinq premiers axes

Une variable illustrative « titre du texte », placée sur chaque individu, permet de projeter sur le plan des deux premiers facteurs issus de l'analyse des paragraphes, les positions respectives des textes correspondants. On observe une structure très proche de la structure précédente : à nouveau, seuls *Spinoza et le problème de l'expression* (SE) et *Logique du sens* (LS) s'opposent à une partition chronologique des textes sur l'origine du premier facteur.

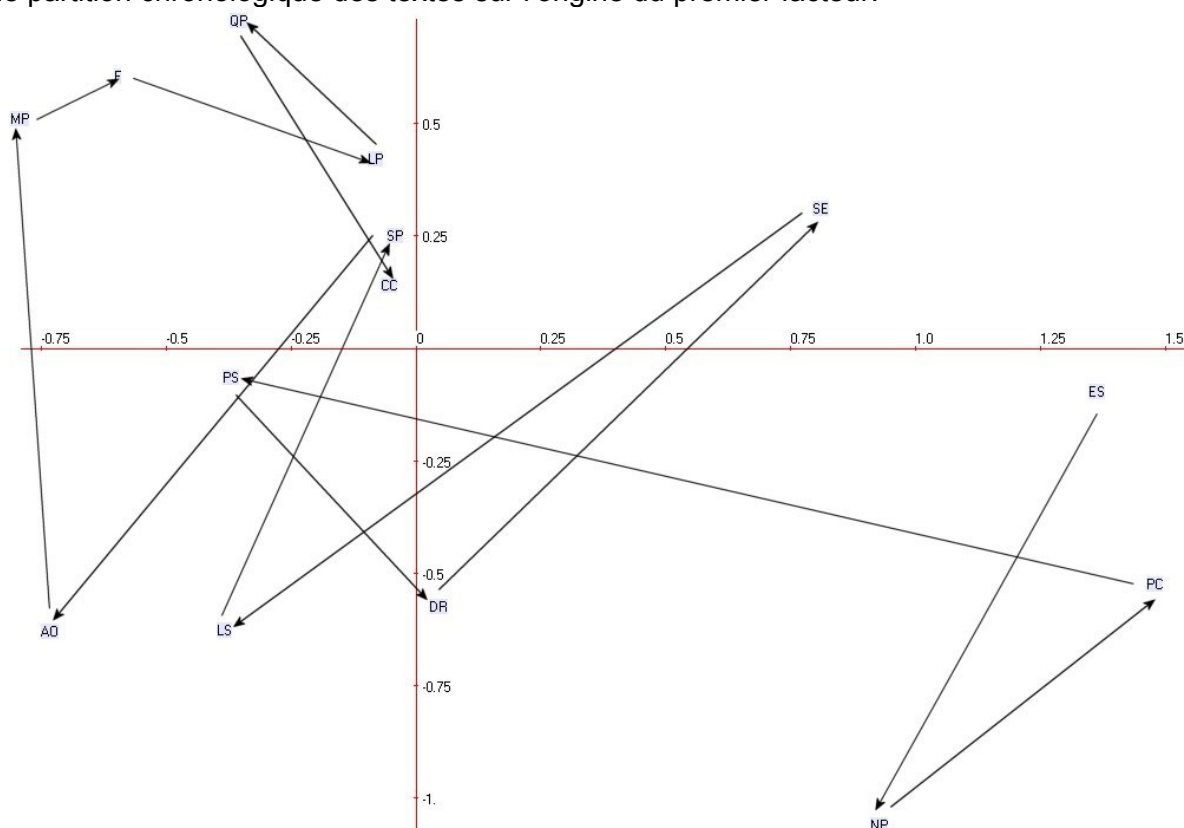


Figure 7 : positionnement de la variable illustrative « texte » sur les deux premiers facteurs de l'analyse en composante principale des paragraphes du corpus

La projection des variables sur les deux premiers facteurs est également proche des distributions observées précédemment.

Les deux premiers facteurs ont cependant une représentativité faible et l'examen de la classification ascendante est nécessaire. On observe une distribution identique des textes à travers les quatre classes, seuls sont écartés cette fois *Proust et les signes* (PS) et *Spinoza philosophie pratique* (SP), les deux commentaires de la classe 4, tandis que, au contraire, *Nietzsche et la philosophie* (NP) réintègre la classe attendue.

L'examen des variables caractéristiques de chaque classe fait apparaître une meilleure distribution des différentes personnes dans les trois classes et, ainsi, une meilleure répartition en ensembles interprétables d'un point de vue dialogique. La quatrième classe acquiert en particulier une caractérisation par la variable *vous*.

Classe		Modalités caractéristiques	Variables caractéristiques positives (personne)	Variables caractéristiques négative (personne)
1	1070	ES PC SE NP	<i>Nous</i>	<i>On, ils, vous</i>
2	1171	MP QP F LP	<i>Ils, on</i>	<i>Il, je, nous</i>
3	3			
4	1386	AO LS DR	<i>Il, vous</i>	<i>Nous, ils</i>

Figure 8 : répartition des catégories « noms du textes » dans les quatre classes issues d'une classification ascendante de tous les paragraphes du corpus

C'est cette structure dialogique, opposant les textes entre eux, que nous pouvons maintenant essayer de décrire plus précisément du point de vue de chacun des textes.

4. Comparaison des structures internes des textes

Afin d'affiner la comparaison, on a répété l'expérience au niveau des paragraphes pour trois textes du corpus issus des trois classes observées : *l'Anti-Cédipe*, *Empirisme et subjectivité*, et *Mille Plateaux*. Les paragraphes de *l'Anti-Cédipe*, répartis en quatre classes¹, font apparaître principalement une opposition structurée autour des personnes :

Classe	Nombre de paragraphes	Variables positives ²	Variables négatives
1	72	{gram}nom-p, {gram}fin.i, {gram}t.p, {c}, {lm}je	{gram}nom-s, [...], {c}!, [...],{c}:, {gram}t.subimp, {c}..., {gram}t.subpre
2	51	{gram}t.p, {lm}vous, {gram}fin.d, {lm}tu, {c}?, [{c}., {c}:]	{gram}m.partpres, {c}, {gram}t.partpas, {c}, {gram}fin.i, {ff}ils
3	94	{gram}t.partpas, {gram}nom-s, {gram}m.partpres, {gram}fin.d, {ff}il, {c}.	{gram}nom-p, {gram}fin.i, {gram}t.p, {gram}ty.p, {c}..., {lm}nous, {gram}t.subimp, {c}!, {lm}vous [...] [{ff}ils, {lm}je, {c}?, {lm}tu]
	72	{lm}nous, {c}!, {c}..., {gram}ty.p, [...], {gram}t.con, {gram}t.fut, [{tag}hi, {c}?, {c}:]	{ff}il, {gram}t.p, {c}., {gram}nom-p

Figure 9 : classification des paragraphes de *l'Anti-Cédipe*

La classe 3, la plus nombreuse par le nombre de paragraphes sélectionnés, est remarquable par son exclusion de toute personne autre que la troisième personne du singulier. Ce caractère impersonnel est renforcé par deux temps particulièrement nominaux (les participes présent et passé) et par les noms singuliers. Du point de vue des ponctuations, seul le point est positivement associé, tandis que tous les signes de ponctuation supposables fortement « dialogiques » sont exclus : points d'exclamations, points d'interrogation et points de suspension. Noms communs pluriels et noms propres sont également corrélés négativement.

Chacune des autres classes se spécialise sur une personne : la première du singulier pour la première classe, les secondes personnes du singulier et du pluriel pour la seconde classe, et la première du pluriel pour la dernière. Dans la seconde classe, les secondes personnes sont associées au défini, aux points d'interrogation, et s'opposent aux participes. Dans la quatrième classe, la première personne du pluriel est associée aux points d'exclamation, aux points de suspension, aux noms propres, au futur, au conditionnel, et, dans une moindre mesure, aux points d'interrogation et aux marques de mise en forme. Les variables opposant les paragraphes font donc fortement ressortir une opposition de régimes dialogiques : régime impersonnel (classe 1),

¹ Les valeurs propres des facteurs sont là encore relativement faibles : 11%, 9%, 7% et 6% pour les quatre premiers. Les intervalles d'Anderson sont par contre faiblement recouverts.

² Au-dessus d'une valeur test de 2. Des variables au-dessous de ce seuil sont notées entre crochets droits. Certaines variables sont exclues (noté [...]).

régime plus rare (51 occurrences) de l'adresse (classe 2 : présent, seconde personne, points d'exclamation), manifestation du philosophe – incluant éventuellement ses destinataires – (classe 4 : 1^{ère} personne du singulier, points d'exclamation et de suspension, temps modaux), et un régime dialogique moins caractérisable (première personne du singulier, noms propres, présent, virgule). *Empirisme et subjectivité*, opposé à l'*Anti-Cédipe* par le premier facteur de l'analyse précédente, montre une structure différente¹.

Classe	Nombre de paragraphes	Variables positives ²	Variables négatives
1	76	{gram}t.p, {gram}nom-s, {gram}t.fut, {gram}t.con, [...] {lm}on [...]	{gram}m.partpres, {gram}t.partpas, {gram}nom-p, {ff}ils [...], {lm}nous
2	2	{lm}vous, {c}?, {tag}note, {lm}je	{c}.
3	77	{gram}m.partpres, {gram}t.partpas, {gram}fin.d, {gram}nom-s	{gram}t.p, {gram}fin.i
	76	{gram}nom-p, {ff}ils, {c};, {c}: {gram}ty.p, [...]	{gram}t.con, {gram}t.fut, {tag}hi [...]

Figure 10 : classification des paragraphes de *Empirisme et subjectivité*

Ici, il est délicat d'interpréter les spécificités des classes comme des ensembles cohérents d'un point de vue dialogique. Les seules personnes présentes, si l'on met à part la troisième personne, sont dans une même classe ne comptant que deux individus (classe 2). Le présent, qui s'opposait précédemment au futur et au conditionnel, leur est maintenant corrélé. Seule la première classe peut être interprétable, le *on* remplaçant le *nous* dans le rôle de la manifestation du philosophe.

Classe	Nombre de paragraphes	Variables positives ³	Variables négatives
1	277	{gram}fin.i, {gram}nom-p, {gram}t.p, {lm}on, {lm}nous	{gram}fin.d, {gram}nom-s, {gram}ty.p, {gram}t.i, {ff}il
2	162	{gram}fin.d, {gram}nom-s, {gram}ty.p, {gram}m.partpres, {gram}t.partpas, {c}:	{gram}fin.i, {gram}nom-p, {lm}on
3	14	{gram}t.i, {gram}t.subimp, {gram}t.pas, {ff}il	{gram}t.p, {tag}note, {lm}nous, {gram}t.fut
	1		

Figure 11 : classification des paragraphes de *Mille Plateaux*

Les très faibles effectifs de la classe 3 rendent son interprétation délicate. L'opposition principale s'établit entre les classes 1 et 2 : on relève en particulier, comme caractéristiques de la première, le présent, le *on* et le *nous*, auxquels s'opposent dans la seconde les noms propres, et les participes. On retrouve dans l'association entre les participes, le défini, et les noms singuliers, une configuration de traits associée à l'absence de marques dialogiques fortes. La principale structure observable à partir de cette classification est l'opposition sur le critère de la présence de marques dialogiques.

Conclusion

Les trois expériences menées ont fait varier les paliers, du texte au paragraphe. Sur le même jeu de variables, on observe à la fois une classification du corpus restituant l'axe diachronique, et

¹ Les quatre premières valeurs propres sont plus faibles que celles de l'analyse de l'*Anti-Cédipe* : 9%, 8%, 7% et 5%.

² Au-dessus d'une valeur test de 2. Des variables au-dessous de ce seuil sont notées entre crochets droits. Certaines variables sont exclues (noté [...]).

³ Au-dessus d'une valeur test de 2. Des variables au-dessous de ce seuil sont notées entre crochets droits. Certaines variables sont exclues (noté [...]).

caractérisé par une opposition sur le registre académique, et une opposition des textes par la distribution de leurs paragraphes, où peut se lire la marque d'une organisation dialogique.

BIBLIOGRAPHIE

- DESCOMBES, V. 1979. *Le même et l'autre : Quarante-cinq ans de philosophie française (1933-1978)*, Minuit, Paris.
- LEBART, L., MORINEAU, A., PIRON, M. 2000. *Statistique exploratoire multidimensionnelle* (troisième édition), Paris, Dunod.
- LEBART, L. 2006. *Data and Text Mining* (DTM) <<http://www.lebart.org>>
- LOISEAU, S. 2005. Thématique et sémantique contextuelle d'un concept philosophique, in G. Williams (éd.), *La linguistique de corpus*, Rennes, Presses Universitaires de Rennes, pp. 129-140.
- LOISEAU, S., POUDAT, C., ABLALI, D. 2006. Exploration contrastive de trois corpus de sciences humaines, *JADT 2006*, Besançon, 19-21 avril 2006.
- LOISEAU, S. 2006. CorpusReader, un logiciel d'interrogation de corpus (<<http://panini.u-paris10.fr/CR>>)
- POUDAT, C. 2006. *Étude contrastive de l'article scientifique de revue linguistique dans une perspective d'analyse des genres*, Thèse de doctorat, Université d'Orléans.
- RASTIER, F. 2001. L'Être naquit dans le langage : un aspect de la mimésis philosophique, *Methodos*, vol. I, n. 1, Lille, Presse du septentrion, pp. 103-132.
- RASTIER, F. 2005. Enjeux épistémologiques de la linguistique de corpus, in G. Williams (éd.), *La linguistique de corpus : Actes des deuxièmes Journées de la linguistique de corpus*, Rennes, Presses Universitaires de Rennes, pp. 31-47.
- VALETTE, F. 2003. Conceptualisation and Evolution of Concepts. The example of French Linguist Gustave Guillaume, in K. Fløttum & F. Rastier (éds.), *Academic discourse -- multidisciplinary approaches*, Oslo, Novus, pp. 55-74.

RÉFÉRENCES DES TEXTES DU CORPUS

- [ES] Deleuze G. (1953 [1998]). *Empirisme et subjectivité*, Paris, PUF.
- [NP] Deleuze G. (1962). *Nietzsche et la philosophie*, Paris, PUF.
- [PC] Deleuze G. (1963 [1994]). *La philosophie critique de Kant*, Paris, PUF.
- [PS] Deleuze G. (1964 [1996]), *Proust et les signes*, Paris, PUF, coll : Quadrige.
- [SE] Deleuze G. (1968). *Spinoza et le problème de l'expression*, Paris, Minuit.
- [DR] Deleuze G. (1968 [2000]). *Différence et répétition*, Paris, PUF.
- [LS] Deleuze G. (1969). *Logique du sens*, Paris, Les éditions de Minuit, coll : Critique.
- [SP] Deleuze G. (1970 [1981]). *Spinoza Philosophie pratique*, Paris, Minuit.
- [AO] Deleuze G., Guattari F. (1972). *L'Anti-Œdipe*, Paris, Minuit.
- [MP] Deleuze G., Guattari F. (1980). *Mille Plateaux*, Paris, Minuit.
- [F] Deleuze G. (1986). *Foucault*, Paris, Minuit.
- [LP] Deleuze G. (1988). *Le pli – Leibniz et le baroque*, Paris, Minuit.
- [QP] Deleuze G., Guattari F. (1991). *Qu'est-ce que la philosophie ?*, Paris, Minuit.
- [CC] Deleuze G. (1993). *Critique et clinique*, Paris, Minuit.