

## Annexes

### *Annexe 1 : Composition du corpus ASLF*

Pour faciliter la représentation, nous n'indiquons ici que les identifiants des revues, ces dernières étant présentées section 2.1.1.2.

<b>ID</b>	<b>Nom<sup>1</sup></b>	<b>Prénom</b>	<b>Titre article</b>	<b>ID Revue</b>
1	Bouquet	Simon	Y a-t-il une théorie saussurienne de l'interprétation?	1
2	Kanellos	Ioannis	De la vie sociale du texte. L'intertexte comme facteur de la coopération interprétative	1
3	Rastier & Pincemin		Des genres à l'intertexte	1
4	Siblot	Paul	Les palimpsestes du texte ou les fantômes de l'interdiscours	1
5	Moirand	Sophie	Les indices dialogiques de contextualisation dans la presse ordinaire	1
6	Tanguy & Thlivitis		Parcours interprétatifs (inter) textuels : vers une assistance informatique	1
7	Kleiber	Georges	Remarques sur la dénomination	2
8	Anscombe	Jean-Claude	Dénomination, sens et référence dans une théorie des stéréotypes nominaux	2
9	Collinot	André	Dénomination d'un objet social dans un discours encyclopédique	2
10	Petit	Gérard	Pour une conception lexicologique de la dénomination	2
11	Humbley	John	Quelques enjeux de la dénomination en terminologie	2
12	Boisson	Claude	Dénomination et « vision »	2
13	Constantin de Chanay	Hugues	La dénomination : perspective discursive et interactive	2
14	Siblot	Paul	De la dénomination à la nomination. Les dynamiques de la signifiante nominale et le propre du nom	2
15	Pérennec	Marcel	La construction du signifié	3
16	François	Jacques	LE PROFIL ASPECTUEL ET PARTICIPATIF DES PREDICATIONS : PERSPECTIVES ALGORITHMIQUES INTERACTIVES	3
17	Courtois	Blandine	Buts et méthodes de l'élaboration des dictionnaires électroniques du LADL	3
18	Cortès	Colette	Effets sur le lexique des mécanismes de la métonymie et de la métaphore	3
19	Salazar	Béatrice	La formation diminutive en espagnol	3
20	Neyrod	Dominique	LES DÉRIVÉS LEXICALISÉS espagnols EN -ILLO/-ILLA,-ÓN : UNE CONTRIBUTION A L'ÉTUDE DE L'ORGANISATION DU LEXIQUE ET DES PROCESSUS DE création LEXICALE	3
21	Cortès & Szabo		Un exemple de catégorisation des noms, des verbes et des adjectifs par le langage naturel : les adverbes allemands en -weise	3
22	Durieux	Christine	TEXTE, CONTEXTE, HYPERTEXTE	3
23	Rochard	Michel	UNE APPROCHE TRADUCTOLOGIQUE DE LA	3

<sup>1</sup> En cas de co-écriture, les noms des auteurs sont séparés par une esperluette et les prénoms ne sont pas mentionnés afin de limiter toute confusion.

			TERMINOLOGIE ET DE LA RÉVISION	
24	Sablayrolles	Jean-François	Néologismes : une typologie des typologies	4
25	Beciri	Hélène	Dérivés, composés... A propos des affixes	4
26	Delplanque	Carine	Néologismes : Le statut phraséologique des composés économiques allemands	4
27	Durieux	Christine	PSEUDO-SYNONYMES EN LANGUE DE SPECIALITE	4
28	Mathieu	Yvette Yannick	Un classement sémantique des verbes psychologiques	4
29	Kriegel	Sibylle	Les constructions pronominales dans deux textes du latin tardif	4
30	Cortès	Colette	Étude traductologique (français-allemand) des constructions réfléchies dans des textes spécialisés	4
31	Cadiot	Pierre	Espaces et prépositions	5
32	Leeman	Danielle	Dans un juron, il sauta sur ses pistolets. Aspects de la polysémie de la préposition dans	5
33	Ibrahim	Amr Helmy	Les prépositions comme traces ou équivalents d'un support	5
34	Marque-Pucheu	Christiane	Source, inférence et position du locuteur dans les énoncés comportant selon	5
35	Martinot	Claire	Premières prépositions : dans ou à	5
36	Pierrard	Michel	Grammaticalisation et contexte : l'extension des emplois de Comme	5
37	Vandeloise	Claude	Quand dans quitte l'espace pour le temps	5
38	Luzzati	Daniel	Impératif, indicatif et subjonctif : de la morphologie à l'interaction	6
39	Grobet	Anne	Pour une approche linguistique ET discursive de l'organisation informationnelle	6
40	Manno	Giuseppe	A propos de quelques actes 'pseudo-directifs' de la communication écrite	6
41	Kwon-Pak	Song-Nim	Par terre vs sur le sol	6
42	Daussaint- Doneux	Isabelle	Le fonctionnement de la conjonction mais dans le dialogue romanesque	6
43	Schneider	Jean	Les concepts fondamentaux et la terminologie des traités orthographiques	7
44	Lo Piparo	Franco	Le théorème de Pythagore dans la linguistique grecque	7
45	Ildefonse	Frédérique	L'absence des complétives de l'analyse grammaticale antique (II)	7
46	Romashko	Sergej A.	Vers l'analyse du dialogue en Russie	8
47	Archaimbault	Sylvie	Un texte fondateur pour l'étude du dialogue : de la parole dialogale (L. Jakubinskij)	8
48	Ivanova	Irina	Spécificités de l'étude du dialogue dans la linguistique russe	8
49	de Fornel & Léon		L'analyse de conversation, de l'ethnométhodologie à la linguistique interactionnelle	8
50	Škiljan	A.	Traduction « rétrospective » et traduction « prospective » des textes grammaticaux anciens	9
51	Wouters	A.	Les conjonctions dubitatives : la définition (originale) de la Technè retrouvée ?	9
52	Laurendeau	P.	Condillac contre Spinoza : une critique nominaliste des glottognoses	9
53	Combettes	Bernard	L'évolution de la forme en <i>-ant</i> : Aspects syntaxiques et textuels	10
54	Halmøy	Odile	Les formes gérondives dans les .XV. Joies de mariage et autres textes du XV <sup>e</sup> siècle	10
55	Arnavielle	Teddy	Le participe, les formes en <i>-ant</i> : Positions et propositions	10
56	Kindt	Saskia	Le participe présent en emploi adnominal comme prétendu équivalent de la relative en Qui	10
57	Lipsky	Angela	Pour une description sémantique et morpho-syntaxique du participe français et allemand	10
58	Duffley	Patrick J.	Les conditions de production de l'effet de sens 'imperfectif' avec la forme en <i>-ing</i> de l'anglais	10

59	De Carvalho	Paulo	Gérondif, participe présent et adjectif déverbal en morphosyntaxe comparative	10
60	Bergounioux	Gabriel	Esquisse d'une histoire négative de l'endophasie avec une attention presque exclusive pour les productions en langue française consacrées à cette question.	11
61	Puech	Christian	Langage intérieur et ontologie linguistique à la fin du XIXe siècle	11
62	Carroy	Jacqueline	Une enquête, ses enjeux et ses limites	11
63	Friedrich	Janette	La discussion du langage intérieur par L.S. Vygotskij	11
64	Rabatel	Alain	LES REPRESENTATIONS DE LA PAROLE INTERIEURE Monologue intérieur, discours direct et indirect libres, point de vue	11
65	Philippe	Gilles	LE PARADOXE ENONCIATIF ENDOPHASIQUE ET SES PREMIERES SOLUTIONS FICTIONNELLES	11
66	Bergounioux	Gabriel	Endophasie et linguistique Décomptes, quotes et squelette.	11
67	Culioli	Antoine	A propos de même	12
68	De Vogüé & Franckel		Identité et variation de l'adjectif grand.	12
69	Péroz	Pierre	Le mot clé Variations sémantiques et régularités des fonctionnements	12
70	Doumbe Akwa	Betote	Les emplois du mot raison, entre singularité et régularité	12
71	Romero-Lopes	Márcia Cristina	Identité et variation du verbe jouer	12
72	Jalenques	Pierre	Etude sémantique du préfixe RE en français contemporain : A propos de plusieurs débats actuels en morphologie dérivationnelle	12
73	Paillard	Denis	Contribution a l'analyse du préfixe sous- combine avec des bases verbales.	12
74	Thuillier	François	À l'emporte-pièce : de la métaphore à l'outil.	12
75	Dalle	Patrice	La place de la langue des signes dans le milieu institutionnel de l'éducation : enjeux, blocages et évolution	13
76	El Khomsi	Brigitte	Présentation d'une expérience pédagogique d'enseignement du français dans une filière bilingue en lycée	13
77	Brugeille	Jean-Louis	L'enseignement de la LSF dans le cadre d'un enseignement en langue des signes au collège	13
78	Fusellier-Souza	Ivani	Apprentissage institutionnel d'une troisième langue par les apprenants sourds : Discussion autour d'une approche bilingue dans l'enseignement d'une langue vivante	13
79	Quipourt	Christine	Interpréter en Langue des Signes : un acte militant ?	13
80	Feuillet	Jack	L'hypothétique dans les langues slaves	14
81	Corblin	Francis	Une approche cognitive des conditionnelles: hypothèse et monotonie	14
82	Trévisé	Anne	A propos de repérages fictifs : variété des formes et construction du sens	14
83	Merle	Jean-Marie	Genèse et interprétation des repères hypothétiques en milieu toncal – entre arbitraire et pertinence	14
84	Filippi-Deswelle	Catherine	Etude énonciative de <i>if</i> et <i>though</i> antéposés	14
85	De Vogüé	Sarah	Le champ des subordonnées dites conditionnelles du français : conditions, éventualités, suppositions et hypothèses	14
86	Paris	Marie-Claude	Condition et conditionnelles en chinois contemporain	14
87	Gadet	Françoise	Le terme « relâchement » en sociolinguistique	15
88	Laks	Bernard	De la variation et des variantes : à propos du relâchement	15
89	Lyche	Chantal	Affaiblissement consonantique en cadien	15
90	Berit Hansen	Anita	Le E caduc interconsonantique en tant que variable sociolinguistique. Une étude en région parisienne	15
91	Armstrong & Boughton		Absence de repères régionaux et relâchement de la prononciation	15

92	Léglise	Isabelle	Variations prosodiques et syntaxiques en français parlé: le cas des anaphoriques nominaux	15
93	Chevrot & Beaud & Varga		L'apprentissage des unités phonologiques variables : l'exemple du /R/ post-consonantique final en français	15
94	Malderez	Isabelle	Variation phonétique et compétences alphabétiques : le traitement des voyelles moyennes antérieures dans la production écrite d'enfants de CE1 et CE2	15
95	Sabio	Frédéric	Les difficultés de la notion de mot: l'exemple des liaisons graphiques dans les textes d'enfants	15
96	Mondada	Lorenza	Les effets théoriques des pratiques de transcription	15
97	Bilger	Mireille	Au sujet de la représentation de la langue parlée	15
98	Lazard	Gilbert	La nature des invariants interlangues	16
99	Pottier	Bernard	La typologie et les universaux	16
100	Hénault-Sakhno & Sakhno		Typologie des langues et sémantique lexicale diachronique : le problème des universaux	16
101	Vallée & Rousset & Boë		Des lexiques aux syllabes des langues du monde. Typologies, tendances et organisations structurelles	16
102	Lemaréchal	Alain	Typologie pré-greenbergienne, morphologie et cognition : « flexionalismes » et « flexionalité »	16
103	Rousseau	André	Réflexions sur les catégories des unités linguistiques : comparaison de l'allemand et du chinois	16
104	Creissels	Denis	Remarques sur la notion de passif et l'origine des constructions passives	16
105	Drettas	Georges	Morphosyntaxe de la diathèse moyenne (médiopassive) en grec contemporain	16
106	Brahim	Ahmed	Passif et moyen dans les langues du pourtour méditerranéen	16
107	Combettes	Bernard	L'émergence d'une catégorie morphosyntaxique : les déterminants du nom en français	16
108	Abeillé & Godard		Deux types de prédicats complexes dans les langues romanes	16
109	Noyau	Colette	Typologie et dynamiques des langues : les études acquisitionnelles sur les processus de morphologisation temporelle	16
110	Anscombe	Jean-Claude	La nuit, certains chats sont gris, ou la généricité sans syntagme générique	17
111	Anis	Jacques	Les titres des Fables de la Fontaine : « définitude », généricité, narrativité	17
112	Bosredon	Bernard	Les titres de Magritte : surprise et convenance discursive	17
113	Martin	Robert	Sur le statut sémantique de il y a	17
114	Kleiber	Georges	Marqueurs référentiels et théorie du centrage	17
115	Lavieu & Leeman		« Tout le monde me montre au doigt... ». Remarques sur le complément de moyen introduit par la préposition à	17
116	Francis	Renaud	Les bases de la collaboration réciproque	17
117	Dostie & de Sève		Du savoir à la collaboration. Étude pragma-sémantique et traitement lexicographique de t'sais	18
118	Kisseleva & Paillard		Les Mots du discours : garant et point de perspective.	18
119	Rossari	Corinne	Pour une approche lexicale des relations de discours : l'exemple de donc	18
120	Richard	Élisabeth	"Félix est beau, mais beau !": du dit au dire	18
121	Bacha	Jacqueline	Marqueurs exclamatifs et aspect verbal	19
122	Gamallo Otero	Pablo	L'interprétation d'expressions nominales analysée comme un processus métonymique	19
123	Daladier	Anne	Une nouvelle syntaxe pour une sémantique constructive du français	19

124	Godart-Wendling	Béatrice	Comment ça réfère ?	19
125	Miche	Elisabeth	L'articulation entre les structures linguistique, textuelle et situationnelle du discours	19
126	Burger	Marcel	Scènes d'actions radiophoniques et prises de rôles : informer, débattre, divertir	19
127	Erlich	Frances D.	Les fonctions pragmatiques globales d'une publication antiraciste française	19
128	Benetti	Laurence	La structure des syntagmes nominaux complexes: problèmes de reconnaissance	20
129	Gapany & Zay		Les relatives parenthétiques : problèmes de reconnaissance et de délimitation	20
130	Fryba-Reber	Anne-Marguerite	Le français avancé d'Henri Frei: une notion contestable illustrée par le traitement de la relative	20
131	Benninger	Céline	Mètre, mètre carre, mètre cube et litre : que mesurent ces unités de grandeurs uni-, bi- et tridimensionnelles ?	20
132	Choi-Jonin	Injoo	La préposition "avec" : opérateur de (dé)composition	20
133	Hug	Marc	Voici et voilà : enseignements d'un corpus d'emplois réels	20
134	Riegel	Martin	L'interprétation de tel adjectival : une variable de caractérisation liée contextuellement	20
135	Reichler-Béguelin	Marie-José	Contrôle du sujet zéro de l'infinitif et programmation de la période	20
136	Berrendonner	Alain	Redoublement actantiel et nominalisations	20
137	Mondada	Lorenza	Planification syntaxique des énoncés et séquentialité de la conversation	20
138	Crévenat-Werner	Danièle	Les centres d'intérêts dans un corpus linguistique : une analyse quantitative	20
139	Jeanneret & Nobelis		Relatives co-énoncées : conversation et syntaxe	20
140	Bonhomme	Marc	La syntaxe de l'image publicitaire et la genèse du sens: une extrapolation des modèles linguistiques ?	20
141	Chanet	Catherine	Dialogue homme-machine et observables : les actes de langage en question	20
142	Roos	Eva	Structures fonctionnelles et hiérarchiques de messages laissés sur répondeur automatique	20
143	Crévenat-Werner	Danièle	Les contextes de contexte La notion de contexte dans les dictionnaires	21
144	Kerbrat-Orecchioni	Catherine	Texte et contexte	21
145	Delhay	Corinne	Texte, contexte, contextualisation : A-t-on progressé ?	21
146	Vaxelaire	Béatrice	La notion de contexte en phonétique : un aperçu	21
147	Didelot-Zerr	Corinne	L'influence du contexte géographique sur la prosodie : analyse phonétique	21
148	Bonnot & Bursztejn		A propos du rôle du contexte dans l'acquisition du langage chez l'enfant	21
149	Pallaud	Berthille	Les contextes de l'acquisition de l'écriture	21
150	Charolles	Michel	Quand intervient le contexte dans la résolution des ambiguïtés ?	21
151	Maingueneau	Dominique	Contexte et scénographie	21
152	Rossignol	Christian	Contexte et circonstances dans la sémiotique juridique	21
153	Fenoglio	Irène	Question du contexte et événement d'énonciation	21
154	Schmoll	Patrick	Production et interprétation du sens : la notion de contexte est-elle opératoire ?	21
155	Piotrowski	David	Le T.A.L.N. Faillite des ambitions théoriques, succès de la « raison pratique »	22
156	Berthelin	Jean-Baptiste	Hypertexte universitaire : des principes aux applications quotidiennes	22

157	Slodzian	Monique	WordNet et EuroWordNet : questions impertinentes sur leur pertinence linguistique	22
158	Pincemin	Bénédicte	Sémantique interprétative et analyses automatiques de textes : que deviennent les sèmes ?	22
159	Habert & Folch & Illouz		Sortir des sens uniques : Repérer les mots « mouvants » dans le domaine social	22
160	Rousseau	André	L'irréel ou le défi de l'impossible dans la langue	23
161	Samain	Didier	Le langage et l'idiome Les partitions de l'espace grammatical au vu de quelques pathologies	23
162	Gelbert	Gisèle	L'impossible du langage pathologique	23
163	Fortis	Jean-Michel	La modularité du langage : une démonstration impossible ?	23
164	Sabouraud	Olivier	Les impossibles du langage : la place des déficits neurologiques, aphasiques et non-aphasiques	23
165	Bronckart	Jean-Paul	Possibles de l'activité langagière et impossibles de « la langue »	23
166	Christol	Alain	Verbal d'obligation et thèmes de futur en grec	24
167	Gebert	Lucyna	Aspect, impératif et futur en polonais et en russe	24
168	Lessan Pezechki	Homa	L'expression de l'avenir dans la langue persane : une étude contrastive avec le français et le portugais	24
169	Laurendeau	Paul	L'alternance futur simple / futur périphrastique : une hypothèse modale	24
170	Bourdin	Philippe	À propos des "futurs" ventifs et itifs : remarques sur un paradoxe	24
171	Cornish	Francis	L'accessibilité cognitive des référents, le Centrage d'attention et la structuration du discours : une vue d'ensemble	25
172	Walker	Marilyn A.	Vers un modèle de l'interaction du Centrage avec la structure globale du discours	25
173	Péry-Woodley	Marie-Paule	Cadrer ou Centrer son discours ? Introduceurs de cadres et Centrage	25
174	Gundel & Hedberg & Zacharsky		Statut cognitif et forme des anaphoriques indirects	25
175	de Mulder	Walter	Démonstratifs et accessibilité	25
176	Benninger & Theissen		L'anaphore pronominale in absentia : le cas de certains	26
177	Bosveld-de Smet	Léonie	Du pluriel au massif	26
178	Van de Velde	Danièle	Quelques variétés de pluriels de noms abstraits	26
179	Bianco & Schnedecker		Reprise pronominale plurielle et construction cognitive des "complexes référentiels" à partir des structures SN1 et SN2 verbe et SN1 verbe avec SN2 : rôle du coordonnant et du sémantisme du verbe	26
180	Brissaud & Chevrot		Acquisition de la morphographie entre 10 et 15 ans : le cas du pluriel des formes verbales en /E/	26
181	Mann & Thompson		Deux perspectives sur la théorie de la structure rhétorique (RST)	27
182	Rossari	Corinne	Les relations de discours : approches rhétoriques, approches pragmatiques et approches sémantiques	27
183	Busquets & Vieu & Asher		La SDRT : une approche de la cohérence du discours dans la tradition de la sémantique dynamique	27
184	Luc & Virbel		Le modèle d'architecture textuelle : fondements et expérimentation	27
185	Hickmann	Maya	Espace et sémantique des verbes : perspective cognitive et inter-langues	28
186	Le Ny & Franquart-Declercq		Flexibilité des significations, traits sémantiques et compréhension des métaphores verbales	28
187	Brion	Cécile	Phrasal Verbs en anglais : réflexions syntactico-sémantiques	28

188	Castagne	Eric	Verbes et prépositions : réflexions sur leur contribution syntactico-sémantique dans l'organisation du français et de l'italien	28
189	Tyvaert	Jean-Emmanuel	Le Verbe comme entité proto-linguistique : une archéologie des mots lexicaux et une histoire du déploiement systématique des parties du discours dans la production verbale	28
190	Blanche-Benveniste	Claire	Phrase et construction verbale	29
191	Berrendonner	Alain	Les deux syntaxes	29
192	Di Cristo	Albert	Prosodie et discours : modélisation et codage de la constituance prosodique	29
193	Lacheret & Victorri		La période intonative comme unité d'analyse pour l'étude du français parlé : modélisation prosodique et enjeux linguistiques	29
194	Seguin	Jean-Pierre	Le métalangage de l'oralité dans les théories de la ponctuation au XVIIIème siècle	29
195	Béguelin	Marie-José	Clause, période ou autre ? La phrase graphique et la question des niveaux d'analyse	29
196	Combettes	Bernard	Aspects de la grammaticalisation de la phrase complexe en ancien et en moyen français	29
197	Neveu	Franck	Du prédicat second à la clause. Sur le rang syntaxique de quelques types de détachements	29
198	François & Manguin		L'unité-phrase dans la traduction littéraire : les corrections d'alignement et leur évaluation	29
199	Roulet	Eddy	Le problème de la définition des unités à la frontière entre le syntaxique et le textuel	29
200	Nølke	Henning	Pour un traitement modulaire de la syntaxe transphrastique	29
201	Bonnotte & Fayol		Approche développementale de la gestion des relations sémantiques entre locutions temporelles, types de procès et temps grammaticaux du passé. Une étude en production et en compréhension	30
202	Cordier	Françoise	Compréhension et production de verbes. Quelques données en psychologie du développement	30
203	François	Jacques	Désémantisation verbale et grammaticalisation. (se) voir employé comme outil de redistribution des actants	30
204	Gayral	Françoise	Traitement automatique et polysémie des verbes	30
205	Gosselin	Laurent	Le statut du temps et de l'aspect dans la structure modale de l'énoncé. Esquisse d'un modèle global	30
206	Ibrahim	Amr Helmy	Une classification des verbes en six classes asymétriques hiérarchisées	30
207	Le Ny	Jean-François	La sémantique des verbes et la représentation des situations	30
208	Marquant-Thiébaud	Monique	Le savoir sur l'action. Un exemple : déplacer quelque chose. Perspective développementale	30
209	Paillard	Denis	A propos des verbes "polysémiques" : identité sémantique et principes de variation	30
210	Manguin	Jean-Luc	Construction d'espaces sémantiques associés aux verbes de déplacement d'objets à partir des données des dictionnaires informatisés des synonymes.	30
211	Sarda	Laure	L'expression du déplacement dans la construction transitive directe	30
212	Rolshoven & Remberger & Lalande		La genèse du passé composé: une interprétation générative d'un changement de valence	31
213	Blumenthal	Peter	Le centrage du verbe transitif	31
214	Jacob	Daniel	Pour une nouvelle définition des fonctions des cas	31
215	Koch	Peter	« Il ne me faut plus nule rien. » – changement sémantique , metataxe et réanalyse	31
216	Stein	Achim	Valence sémantique et définition lexicographique	31
217	Waltereit	Richard	Le rapport dépendancier entre adjectif et nom : données	31

			syntaxiques et structures conceptuelles	
218	Ville-Ometz	Fabienne	La préposition à et les verbes de transfert de disposition	31
219	Theissen	Anne	Petite incursion dans la jungle topicale	32
220	Cappeau &		Partition et topicalisation ; il y en a « stabilisateur » de sujets et de topiques indéfinis	32
221	Crévenat-Werner	Danielle	Cinq Belges dont un Wallon : un dont déviant ou non ?	32
222	Combettes & Prévost		Évolution des marqueurs de topicalisation	32
223	Debaisieux	Jeanne-Marie	Contraintes syntaxiques et discursives des emplois de quant à et en ce qui concerne en français parlé	32
224	Guimier	Claude	Les adverbes de domaine en -wise de l'anglais moderne	32

## *Annexe 2 : Composition du corpus « Auteurs »*

ID	Auteur	Titre	Référence	Date
225	Authier-Revuz J.	Modalisation autonymique et discours autre, quelques remarques	Modèles Linguistiques, Vol. 35, XVIII, fasc.1	1997
226	Authier-Revuz J.	Deux mots pour une chose ; trajets de non-coïncidence	Annales Littéraires de l'Université de Besançon n 701, Presses Universitaires Franc-Comtoises	2000
227	Authier-Revuz J.	La strate méta-énonciative, lieu d'inscription du sujet dans son dire : enjeux théoriques et descriptifs d'une approche littérale. L'exemple des modalités irréalisantes du dire.		
228	Authier-Revuz J.	Du Dire "en plus" : dédoublement réflexif et ajout sur la chaîne	Figures d'ajout - phrase, texte, écriture (J. Authier-Revuz, M.C. Lala, eds.), Presses de la Sorbonne Nouvelle p. 147-167	2002
229	Authier-Revuz J.	Aux risques de l'allusion	L'Allusion dans la Littérature, M. Murat éd., coll. Colloques de la Sorbonne, Presses Universitaires de Paris-Sorbonne, p.209-235.	2000
230	Authier-Revuz J.	Énonciation, méta-énonciation. Hétérogénéités énonciatives et problématiques du sujet	Les sujets et leurs discours - Énonciation et interaction, R. Vion (éd.), Publications de l'Université de Provence, 1998, p. 63-79.	
231	Authier-Revuz J.	"Comme disent les humains" : un exil langagier	in "Le signe et la lettre. Hommage à Michel Arrivé" L'harmattan, Paris	2002
232	Authier-Revuz J.	La représentation du discours autre : un champ multiplement hétérogène	in : Le discours rapporté dans tous ses états, J.M. Lopez Munoz, S. Marnette, L. Rosier eds, L'Harmattan, Paris, p. 35-53.	2004
233	Authier-Revuz J.	Le guillemet, un signe de "langue écrite" à part entière	in "A qui appartient la ponctuation ?", Actes du colloque international et interdisciplinaire de Liège, 13-15 mars 1997, J.M. Defays, L. Rosier, F. Tilkin, eds, Duculot 1998, p. 373-388.	1998
234	Authier-Revuz J.	Le Fait autonome : Langage, Langue, Discours – Quelques repères	Actes du colloque : Le fait autonymique dans les langues et les discours, SYLED, Université de la Sorbonne Nouvelle, Paris 3, 5-7 octobre 2000, sur le site internet < <a href="http://www.cavi.univ-paris3.fr/ilpga/autonymie/actes.htm">http://www.cavi.univ-paris3.fr/ilpga/autonymie/actes.htm</a> >,	2000



			à paraître in Parler des mots	
235	Barbérís JM	"Rue X". La grémémisation à l'œuvre dans la parole	Faits de langue 9	1997
236	Barbérís JM	Identité, ipséité dans la déixis spatiale : "ici" et "là", deux appréhensions concurrentes de l'espace	Info grammaticale	
237	Barbérís JM	"Faire parler les siècles" : familiarité et français parlé dans La légende des siècles		
238	Barbérís JM	Représenter l'espace de la ville en contexte interculturel : l'"impasse" identitaire	CDP	
239	Barbérís JM	Etre là sans y être. Mise en scène de la présence et de l'absence dans Così fan tutte		
240	Barbérís JM	Interjection : de l'affect à la parade, et retour		
241	Barbérís JM	Le sujet et sa praxis dans l'expression de l'espace : les énoncés de "mouvement fictif"	Langages 127	
242	Barbérís JM	La dynamique de la référence spatiale. Elaborer des cartes cognitives en langue parlée	Modèles linguistiques 8	
243	Barbérís JM	"Faire voir" : la communication des connaissances, entre dire et montrer	Cahiers de praxématique 28	1998
60	Bergounioux	Esquisse d'une histoire négative de l'endophasie avec une attention presque exclusive pour les productions en langue française consacrées à cette question	LF 132	2001
66	Bergounioux	Endophasie et linguistique Décomptes, quotes et squelette	LF 132	2001
285	Bergounioux	Dans l'enseignement supérieur et hors de l'Université : Albert Dauzat		
286	Bergounioux	LA LANGUE ET LE CERVEAU Esquisse d'une histoire de l'aphasiologie d'un point de vue linguistique (XIXe-XXe siècle)		
287	Bergounioux	L'UNIVERSITE, DE LA LIBERATION A LA CRISE (1945-1968)	Modèles linguistiques	
288	Bergounioux	LA SEMANTIQUE DANS LE CHAMP DE LA LINGUISTIQUE FRANCOPHONE JUSQU'À 1916	Sémiotiques	
244	Calvet LJ	L'argot et la « langue des linguistes ». Des origines de l'argotologie aux silences de la linguistique		
245	Calvet LJ	Le plurilinguisme alexandrin		
246	Calvet LJ	Les langues dans les relations internationales. L'in vivo et l'in vitro		
247	Calvet LJ	Aux origines de la sociolinguistique. la conférence de sociolinguistique de l'UCLA (1964)		
248	Calvet LJ	Les politiques linguistiques en Afrique francophone. État des lieux du point de vue de la politologie linguistique		
249	Calvet LJ	Le signe linguistique. De Saussure à Lacan	conf	
250	Calvet LJ	La France a-t-elle une politique linguistique?	conf	
251	Calvet LJ	DIVERSITE DES USAGES ET USAGES DE LA DIVERSITE : L'aioli	conf	

		et l'Académie		
53	Combettes B	L'évolution de la forme en -ant : aspects syntaxiques et textuels	LANGAGE 149	2003
107	Combettes B	L'émergence d'une catégorie morphosyntaxique : les déterminants du nom en français	LINX 45	2001
196	Combettes B	Aspects de la grammaticalisation de la phrase complexe en ancien et en moyen français	Verbum 24	2002
222	Combettes B	Évolution des marqueurs de topicalisation	PRAX 37	2001
252	Combettes B	Texte argumentatif et structures syntaxiques en Moyen Français : la topicalisation d'une proposition	Linx	
253	Combettes B	Aspects diachroniques des constructions à détachement		
254	Combettes B	La grammaticalisation d'un déterminant indéfini : quelque en moyen français	Scolia	
255	Combettes B	La thématization dans une perspective diachronique : Structures textuelles et système linguistique		
256	Combettes B	Thématisation, topicalisation, et éléments non-référentiels : le cas de l'adjectif détaché	PRAX 30	1998
257	Combettes B	Grammaticalisation de la phrase complexe et évolution de la cataphore		
258	Combettes B	Topicalisation d'une proposition et réfutation : approche diachronique		
259	Combettes B	AU CONTRAIRE, EN REVANCHE, PAR CONTRE : ASPECTS DIACHRONIQUES		
16	François	Le profil aspectuel et participatif des prédications : perspectives algorithmiques interactives	CIEL 94	1994
198	François	L'unité-phrase dans la traduction littéraire: les corrections d'alignement et leur évaluation	Verbum 24	2002
203	François	Désémantisation verbale et grammaticalisation (se) voir employé comme outil de redistribution des actants	SynSém 2	2001
260	François	La compréhension de l'articulation chronologique des séquences narratives: un exercice de linguistique cognitive	Verbum 4	1993
261	François	La compositionnalité de l'aspect et les tests de classement aspectuel	Etudes lexicales 1	1993
262	François	La gauche et la droite dans un corpus de déclarations d'accidents de la circulation	CIRLEP 7	1996
263	François	L'évaluation de la proximité sémantique entre prédications dans le champ onomasiologique du transfert d'avoir	H. Dupuy-Engelhardt (dir. 1996), Questions de méthode et de délimitation en sémantique lexicale. Actes d'EUROSEM 1994, Presses Universitaires de Reims, pp.65-90	1996
264	François	La sémantique	Cours de psychologie édité par R. Ghiglione & J.F. Richard (Volume 4, 2° édition 1999, p. 54-107)	1999
144	Kerbrat-Orecchioni C.	Texte et contexte	Scolia	1996
265	Kerbrat-Orecchioni C.	Suivez le guide ! Les modalités de l'invitation au voyage dans les guides touristiques : l'exemple de l'île		

		d'Aphrodite"		
266	Kerbrat-Orecchioni C.	Converser au féminin		
267	Kerbrat-Orecchioni C.	L'humour au quotidien		
268	Kerbrat-Orecchioni C.	Politesse en deçà des Pyrénées, impolitesse au delà : retour sur la question de l'universalité de la (théorie de la) politesse		
269	Kerbrat-Orecchioni C.	La rhétorique interpersonnelle : Coopération et politesse dans la communication d'itinéraire	CDP	
270	Kerbrat-Orecchioni C.	Les lectures : vingt ans après		
271	Kerbrat-Orecchioni C.	Le malentendu dans la communication ordinaire : Sources, traitement, effets		
7	Kleiber	Remarques sur la dénomination	PRAX 36	2001
114	Kleiber	Marqueurs référentiels et théorie du centrage	LINX 47	2002
272	Kleiber	Contexte, où es-tu?	RSP 1	1997
273	Kleiber	Anaphore associative, thèse lexicostéréotypique : oui, mais...	PRAX 24	1995
274	Kleiber	Au générique : tout ça pour ça	CIRLEP 9	1997
275	Kleiber	Sur le sens du sens : objectivisme et constructivisme	KELLER, D., DURAFOUR, J.-P., BONNOT, J.-F. et SOCK, R. (éds), Percevoir : monde et langage, Bruxelles, Mardaga, 335-370.	2001
276	Kleiber	Noms propres et noms communs : un problème de dénomination	Meta, 41 - 4	1996
277	Kleiber	On a tous un petit vélo dans la tête ou L'anaphore en roue libre ou Les différents cycles de l'anaphore ou encore Est-ce que vraiment l'anaphore est un vélo ?	La grammaire en roue libre. Variations cyclo-linguistiques offertes à Marc Wilmet, Bruxelles, Maison de la Francité (20 juin 2003), 7-16	2003
300	Lamizet B.	Sémiotique de la main. Codes de la main et médiations du corps	Degrés	
301	Lamizet B.	Y A-T-IL UN "PARLER JEUNE" ?	Cahiers de sociolinguistique	
302	Lamizet B.	LE MIROIR MÉDIATÉ. PSYCHANALYSE ET MÉDIAS	Etudes de com?	
303	Lamizet B.	LA SÉMIOTIQUE INSTANTE. INTRODUCTION À LA SÉMIOTIQUE POLITIQUE	Semiotica?	
304	Lamizet B.	Sémiotique et sciences sociales	congrès de l'Association Internationale de Sémiotique	
305	Lamizet B.	TERMINOLOGIE DE LA SÉMIOTIQUE	congrès de l'Association Internationale de Sémiotique	
306	Lamizet B.	ALTERCULTURALITÉ ET INTERCULTURALITÉ. L'INTERCULTURALITÉ URBAINE	congrès de l'Association Internationale de Sémiotique	
307	Lamizet B.	SÉMIOTIQUE RELIGIEUSE DE L'IDENTITÉ	congrès de l'Association Internationale de Sémiotique	
5	Moirand S	Les indices dialogiques de contextualisation dans la presse ordinaire	CDP	
308	Moirand S	Le dialogisme, entre problématiques énonciatives et théories discursives	Cahiers de praxématique 43	2005
309	Moirand S	L'impossible clôture des corpus médiatiques. La mise à jour des observables entre catégorisation et contextualisation	Tranel 40	2004

310	Moirand S	Du traitement différent de l'intertexte selon les genres convoqués dans les événements scientifiques à caractère politique	Semen	2001
311	Moirand S	Formes discursives de la diffusion des savoirs dans les médias	Hermès	1997
312	Moirand S	Variations discursives dans deux situations contrastées de la presse ordinaire	Carnets du Cediscor 6	1998
313	Moirand S	Éléments de théorisation d'une linguistique du discours. L'exemple du discours sur les sciences dans les médias	Modèles linguistiques	1999
314	Neveu F	Des marges de la phrase aux marges de la grammaire; sur la grammatisation de "l'annexe syntaxique" en français	Colloque	
315	Neveu F	Sur la relation partie/tout et la désignation indistincte dans Fin de partie - référence et contexte au théâtre		
316	Neveu F	Grammaire de l'adresse. Aspects de la discontinuité syntaxique		
317	Neveu F	De la syntaxe à l'image textuelle. Ponctuation et niveaux d'analyse linguistique		
318	Neveu F	Quelle syntaxe pour l'apposition? Les types d'appariement des appositions frontales et la continuité référentielle		
319	Neveu F	La glose et le système appositif		
320	Neveu F	L'apposition : concepts, niveaux, domaines		
	Neveu F	Du prédicat second à la glose - sur le rang syntaxique de quelques types de détachements		
64	Rabatel A	LES REPRESENTATIONS DE LA PAROLE INTERIEURE Monologue intérieur, discours direct et indirect libres, point de vue	LF	2001
279	Rabatel A	Cas de belligérance entre perspectives du narrateur et du personnage :neutralisation ou mise en résonance des points de vue ?	LINX	2000
280	Rabatel A	De l'influence de la fréquence itérative sur l'accroissement de la profondeur de la perspective.Un retour critique sur l'omniscience narrative et sur la restriction de champ du personnage.	Protée	2000
281	Rabatel A	Valeurs représentative et énonciative du « présentatif » C'est et marquage du point de vue	LF	2000
282	Rabatel A	LA VALEUR DE « ON » PRONOM INDEFINI/PRONOM PERSONNEL DANS LES PERCEPTIONS REPRESENTEES		2001
283	Rabatel A	FONDUS ENCHAINES ENONCIATIFS complémentarités et distorsions des points de vue représenté, raconté et asserté au fil du texte	Poétique	2001
284	Rabatel A	LA VALEUR DELIBERATIVE DES CONNECTEURS ET MARQUEURS TEMPORELS mais,cependant, maintenant, alors, et DANS	RF?	2001

		L'EMBRAYAGE DU POINT DE VUE. Propositions en faveur d'un continuum argumentativo-temporel		
278	Rabatel A.	Mais dans les énoncés narratifs un embrayeur du point de vue et un organisateur textuel		1999
3	Rastier	Des genres à l'intertexte	PRAX 33	1999
321	Rastier	Les fondations de la sémiotique et le problème du texte	RSP 5	1999
322	Rastier	De l'antinomie spatio-temporelle aux zones anthropiques	CIRLEP 7	1996
323	Rastier	Philologie numérique	CIRLEP 15	1999
324	Rastier	Sémiotique et sciences de la culture	LINX 44	2001
325	Rastier	Ontologie(s)	Revue des sciences et technologies de l'information, série : Revue d'Intelligence artificielle, Vol. 18, n°1, 2004, p. 15-40	2004
326	Rastier	Doxa et lexique en corpus - pour une sémantique des "idéologies"	Actes des Journées Scientifiques en linguistique 2002-03, J. Pauchard et F.Canon-Roger (éds.), CIRLLLEP, Presses Universitaires de Reims, N° 22, 2004	2004
327	Rastier	L'action et le sens pour une sémiotique des cultures	Journal des anthropologues, n°.85-86, mai 2001, pp. 183-219	2001
328	Rastier	L'être naquit dans le langage Un aspect de la mimesis philosophique	Methodos , v. I, n°1, 2001, Lille, Presses du Septentrion, pp. 103-132	2001
292	Rosier L	Méandres de la circulation du terme polyphonie		
293	Rosier L	La presse et les modalités du discours rapporté : l'effet d'hyperréalisme du discours direct surmarqué		
294	Rosier L	Les locutions « Nous autres » et « Nous autres + (X) » : Mais de quels "autres" s'agit-il ?		
295	Rosier L	La circulation des discours à la lumière de « l'effacement énonciatif » : l'exemple du discours puriste sur la langue	Langages 156	2004
296	Rosier L	Sortes d'invectives !ouApproche syntactico-sémantique des termes espèce et sorte		
297	Rosier L	De l'insulte aux mots doux en français : parcours de discours		
298	Rosier L	L'imparfait ventriloque ?	Cahiers Chronos	2005
299	Rosier L	Interjection, subjectivité, expressivité et discours rapporté à l'écrit : petits effets d'un petit discours	CDP34	2000
4	Siblot	Les palimpsestes du texte ou les fantômes de l'interdiscours	PRAX 33	1999
14	Siblot	de la dénomination à la nomination Les dynamiques de la signifiance nominale et le propre du nom	PRAX 36	2001
289	Siblot	La linguistique peut-elle traiter de la « représentation des connaissances dans le lexique » ?	PRAX 21	1993
290	Siblot	L'ÉLOQUENCE DES SILENCES D'une absence de nomination comme déni de réalité	PRAX 23	1994
291	Siblot	Variations sur un thème qui « toujours déjà » prédique	PRAX 30	1998

### *Annexe 3 : Composition du corpus « Mécanique »*

[XVème Congrès Français de Mécanique]

ID	année	Thème	Titre	Nom <sup>12</sup>
0002	2001	Acoustique	Moteur thermoacoustique annulaire à ondes progressives	JOB
0003	2001	Acoustique	Détermination des distributions de contrainte dans une membrane en silicium couplée à une cavité – application au calcul de la sensibilité de microphones piézorésistifs	DURAND
0004	2001	Acoustique	MESURE DE LA DISPERSION DES ONDES DANS UN MILIEU	DERIBLE
0005	2001	Acoustique	POREUX	ECOTIERE
0006	2001	Aérodynamique hydrodynamique	ÉCOULEMENT EN CANAL AVEC OBSTACLE	GUIOT
0007	2001	Aérodynamique hydrodynamique	ÉTUDE DES EFFETS D'UN FORT NIVEAU DE TURBULENCE SUR L'AÉRODYNAMIQUE DES ÉOLIENNES À AXE HORIZONTAL	LAVERNE
0008	2001	Aérodynamique hydrodynamique	L'ONDE DE SOUFFLE : DE LA COLUMBIAD À ARIANE 5	VARNIER
0009	2001	Aérodynamique hydrodynamique	ÉCOULEMENT TRIDIMENSIONNEL DANS UNE CAVITÉ ROTOR-STATOR cylindrique avec éclatement tourbillonnaire	Serre
0010	2001	Aérodynamique hydrodynamique	INFLUENCE DES CONDITIONS D'ENTRÉE SUR LES TAUX DE DÉFORMATIONS ET LE MÉLANGE INDUITS DANS UNE GÉOMÉTRIE ANNULAIRE PAR UN ÉCOULEMENT TOURBILLONNAIRE NON ENTRETENU	PRUVOST
0011	2001	Aérodynamique hydrodynamique	UNE MÉTHODE DE CALCUL POUR LA LIGNE PORTANTE COURBE ET EN FLÈCHE ISSUE DE LA THÉORIE GÉNÉRALISÉE	GALLOIS
0012	2001	Biomécanique	ACCÉLÉRATIONS DANS LE RÉSEAU VEINEUX DU MEMBRE INFÉRIEUR AU COURS DE LA MARCHÉ STATIONNAIRE	AUBERT
0013	2001	Biomécanique	ACOUSTIQUE ULTRASONORE DES MILIEUX POREUX : APPLICATION À LA CARACTÉRISATION DE L'OS TRABÉCULAIRE	CARDOSO
0014	2001	Biomécanique	COMPRESSIBILITÉ ET ORTHOTROPIE DES VAISSEAUX ARTERIELS	CANTERI
0015	2001	Biomécanique	ÉTUDE COMPARATIVE DES TESTS D'ÉVALUATION MÉCANIQUE DU SAUT VERTICAL	BARBEDETTE
0016	2001	Biomécanique	ÉTUDE DYNAMIQUE DE L'INTERACTION ENTRE MOLECULES D'ADHESION à la surface cellulaire	ZHAO
0017	2001	Biomécanique	ÉTUDE EXPERIMENTALE DU MOUVEMENT ET DE LA déformation d'une capsule dans un écoulement de poiseuille	Collé-Paillet
0018	2001	Biomécanique	MODELISATION DU COMPORTEMENT MÉCANIQUE DE LA CAROTIDE HUMAINE: APPLICATION A L'HYPERTENSION	ZIDI

<sup>2</sup> On ne présente ici que le premier auteur de l'article.

0019	2001	Biomécanique	MODELISATIONS EXPERIMENTALE ET NUMERIQUE D'UN écoulement physiologique dans un tube coudé	DEPLANO
0020	2001	Conception - Production	AIDE AU CHOIX DE PROCÉDÉS DE TRAITEMENT DE SURFACE : APPLICATION À L'OPTIMISATION DE LA DURÉE DE VIE DE COMPOSANTS MÉCANIQUES	EYNARD
0021	2001	Conception - Production	AMÉLIORATION DE L'ÉTAT DE SURFACE MICROGÉOMÉTRIQUE EN RAINURAGE CINQ AXES	BOUJELBENE
0022	2001	Conception - Production	APPORT DE L'OUTIL INFORMATIQUE POUR LA CONCEPTION DES ASSEMBLAGES DANS LES CHARPENTES BOIS	IMBERDIS
0023	2001	Conception - Production	CONCEPTION OPTIMALE DES RESSORTS DE TORSION	PAREDES
0024	2001	Conception - Production	CONTRIBUTION DE LA LOGIQUE FLOUE À LA COTATION FONCTIONNELLE	LELU
0025	2001	Conception - Production	DÉTERMINATION DES CONTRAINTES DE DIMENSIONNEMENT : APPLICATION AUX BUTÉES DES MONTAGES DE ROULEMENTS	BELLOY
0026	2001	Conception - Production	INFLUENCE DE LA FLEXIBILITÉ DU SERRE-FLAN LORS D'UNE OPÉRATION D'EMBOUITISSAGE	MARTINEZ
0027	2001	Conception - Production	LA PROSPECTIVE TECHNOLOGIQUE ET TRIZ, DE NOUVEAUX OUTILS POUR LE CONCEPTEUR	CRUBLEAU
0028	2001	Dynamique des structures et des machines	APPROXIMANTS DE PADÉ ET ANALYSE MODALE	LARDIES
0029	2001	Dynamique des structures et des machines	COMPORTEMENT DYNAMIQUE DES JOINTS DANS LES ASSEMBLAGES COMPLEXES : APPROCHE UNIFIÉE CONDENSATION / SIMPLIFICATION	HUANG
0030	2001	Dynamique des structures et des machines	COMPORTEMENT DYNAMIQUE D'UN ROTOR SOUMIS A UN CHOC : simulation - expérimentation	DUCHEMIN
0031	2001	Dynamique des structures et des machines	DEUX METHODES D'IDENTIFICATION MODALE avec réduction d'espace	RAFFY
0032	2001	Dynamique des structures et des machines	IDENTIFICATION MODALE DANS LE DOMAINE TEMPOREL : application à des structures industrielles	LARDIES
0033	2001	Dynamique des structures et des machines	METHODE EXPERIMENTALE D'IDENTIFICATION ET DE LOCALISATION des diamètres nodaux d'un disque en rotation	BERGER
0034	2001	Dynamique des structures et des machines	RECHERCHE D'UNE METHODE ORIGINALE POUR L'IDENTIFICATION DES MODES PROPRES D'UN MONTAGE DYNAMIQUE.	HAUGOU
0035	2001	Ecoulements polyphasiques	COMPARAISON DE DIFFERENTS MODELES DE DISPERSION EN simulation lagrangienne d'écoulements gaz-solide en conduite	PETRISSANS
0036	2001	Ecoulements polyphasiques	CROISSANCE ET DETACHEMENT DE BULLES dans un écoulement cisailé	DUHAR
0037	2001	Ecoulements polyphasiques	EFFET DE STRUCTURES COHERENTES SUR LA DISPERSION DE goutellettes au voisinage d'une couche de cisaillement	FOHANNO

0038	2001	Ecoulements polyphasiques	ETUDE EXPERIMENTALE D'ECOULEMENTS GAZ / LIQUIDE DANS UN MODULE plan de nanofiltration	DUCOM
0039	2001	Ecoulements polyphasiques	Etude expérimentale des propriétés locales d'un écoulement à bulles piloté par la gravité à grands taux de vide	GARNIER
0040	2001	Ecoulements polyphasiques	EVALUATION DE LA REPARTITION DE FRACTION VOLUMIQUE DE PHASE SOLIDE AU SEIN D'UN ECOULEMENT SEDIMENTAIRE	BOUNHOURE
0041	2001	Ecoulements polyphasiques	INFLUENCE D'UNE PHASE DISPERSEE SUR L'ECOULEMENT DANS LES MELANGEURS STATIQUES	HIRECH
0042	2001	Ecoulements polyphasiques	LOI DE TATE : ANALYSE CRITIQUE ET RESTRICTIONS	GEORGESCU
0043	2001	Ecoulements polyphasiques	SIMULATION FOUR-WAY D'UN ECOULEMENT GAZ-SOLIDE NON ISOTHERME PAR APPROCHE EULERIENNE-LAGRANGIENNE	MOISSETTE
0044	2001	Endommagement - Rupture - Fatigue	CALCUL À LA RUPTURE APPLIQUÉE AU MULTICOUCHE : EXEMPLE DE DALLE EN BÉTON ARMÉ RÉPARÉE PAR DU MATÉRIAU COMPOSITE	LIMAM
0045	2001	Endommagement - Rupture - Fatigue	CHRONOLOGIE DE LA RUPTURE STATIQUE ET DYNAMIQUE D'UN ANNEAU EN VERRE FLOAT	N'YOUNGUE
0046	2001	Endommagement - Rupture - Fatigue	ENDOMMAGEMENT ANISOTROPIQUE DU BÉTON SOUMIS À DES CHARGEMENTS DE FATIGUE	ALLICHE
0047	2001	Endommagement - Rupture - Fatigue	MÉTHODE THERMODYNAMIQUE DE DE DONDER ET ENDOMMAGEMENT	MARTIN
0048	2001	Environnement - Milieux poreux	AMELIORATION DES CARACTERISTIQUES MECANQUES DES ARGILES par électro-osmose : essais en centrifugeuse	BEDDIAR
0049	2001	Environnement - Milieux poreux	CONVECTION NATURELLE DANS UN CANAL PLAN MUNI D'UNE SUCCESSION DE MATRICES POREUSES	BOUHADEF
0050	2001	Environnement - Milieux poreux	ETUDE PAR GAMMAMETRIE DOUBLE-SOURCE de sols gonflants non-saturés	ROLLAND

### *Annexe 4 : Composition du corpus ASLA*

ID	Titre article	Nom	Prénom	Revue	Vol.	Issue
1	Pragmatics markers revisited with a focus on you know in adult and adolescent talk	Erman	Britt	Journal of Pragmatics	33	9
2	Japanese reportive evidentiality and the pragmatics of retelling	Mushin	Ilana	Journal of Pragmatics	33	9
3	Modal particles and their functional equivalents: A speech-act-theoretic approach	Waltereit	Richard	Journal of Pragmatics	33	9
4	'She knows more about Hong Kong than you do isn't it': Tags in Hong Kong conversational English	Cheng & Warren		Journal of Pragmatics	33	9
5	Dui bu dui as a pragmatic marker: Evidence from Chinese classroom discourse	Chen & He		Journal of Pragmatics	33	9



7	Jordanian college students' responses to compliments : A pilot study	Farghal		Journal of Pragmatics	33	9
8	Subjective and objective modality: Interpersonal and ideational functions in the English modal auxiliary system	Verstraete	Jean-Christophe	Journal of Pragmatics	33	10
9	Presupposition can be a bluff: How abstract nouns can be used as presupposition triggers	Schmid	Hans-Jörg	Journal of Pragmatics	33	10
10	In defence of a code: Linguistic meaning and propositionality in verbal communication	Chapman	Siobhan	Journal of Pragmatics	33	10
11	Dialectics of dialect and dialectology: Culture, structure and ideology of 'the Kyôto dialect' of Japanese	Koyama	Wataru	Journal of Pragmatics	33	10
12	Cooperating with Grice: A cross-disciplinary metaperspective on uses of Grice's cooperative principle	Lindblom	Kenneth	Journal of Pragmatics	33	10
14	Making sense of nonsense: Fabrication, ambiguity, error and clarification in the organization of experience in ordinary conversation	Dornelles & Garcez		Journal of Pragmatics	33	11
17	Illocutionary force and degrees of strength in language use	Sbisà	Marina	Journal of Pragmatics	33	12
18	Questions and strategic orientation in verbal conflict sequences	Gruber	Helmut	Journal of Pragmatics	33	12
19	Describing conversational dominance	Itakura	Hiroko	Journal of Pragmatics	33	12
20	Arguing about the future: on indirect disagreements in conversations	Georgakopoulou	Alexandra	Journal of Pragmatics	33	12
21	Linguistic identities in talk-in-interaction: Order in bilingual conversation	Gafaranga	Joseph	Journal of Pragmatics	33	12
22	The interactional organization of pharmacist consultations in a hospital setting: A putative structure	Pilnick	Alison	Journal of Pragmatics	33	12
23	Is semantics still possible?	Berg	Jonathan	Journal of Pragmatics	34	4
24	The demise of a unique concept of literal meaning	Ariel	Mira	Journal of Pragmatics	34	4
25	What pragmatics can tell us about (literal) meaning: A critical note on Bach's theory of implicature	Vicente	Begona	Journal of Pragmatics	34	4
26	Literally speaking	Israel	Michael	Journal of Pragmatics	34	4
27	Literal meaning, minimal propositions, and pragmatic processing	Bezuidenhout & Cooper		Journal of Pragmatics	34	4
28	A new look at literal meaning in understanding what is said and implicated	Gibbs	Raymond W.	Journal of Pragmatics	34	4
29	Literal vs. figurative language: Different or equal?	Giora	Rachel	Journal of Pragmatics	34	4
30	The pragmatics of conducive questions in academic discourse	Piazza	Roberta	Journal of Pragmatics	34	5
31	Managing rapport in talk: Using rapport sensitive incidents to explore the motivational concerns underlying the management of relations	Spencer-Oatey	Helen	Journal of Pragmatics	34	5
32	German compliment responses	Golato	Andrea	Journal of Pragmatics	34	5
33	Conflict management in Japanese public affairs talk shows	Honda	Atsuko	Journal of Pragmatics	34	5

34	The phonology of epenthetic stops: implications for the phonetics-phonology interface in optimality theory	Warner	Natasha	Linguistics	40	1
35	Computing perspective: the pluperfect in Dutch	Oversteegen & Bekker		Linguistics	40	1
36	Self-mockery in Japanese	Suzuki	Satoko	Linguistics	40	1
37	Dynamic versus static phonotactic conditions in prosodic morphology	Silverman	Daniel	Linguistics	40	1
38	Does MANNER count or not? Manner-of-motion verbs revisited	Iwata	Seizi	Linguistics	40	1
39	Why quantitative?	Eddington	David	Linguistics	40	2
40	On the diphthong/hiatus contrast in Spanish: some experimental results	Hualde & Prieto		Linguistics	40	2
41	The time-course of inflexional morphological priming	Dominguez et al.		Linguistics	40	2
42	Leísmo in contemporary Spanish American educated speech	Demello	George	Linguistics	40	2
43	Le: from pronoun to intensifier	Cacoullos	Rena Torres	Linguistics	40	2
44	Spanish evidence for pitch-accent structure	Face	Timothy L.	Linguistics	40	2
45	Is there a Spanish imperative intonation revisited: local considerations	Willis	E. W.	Linguistics	40	2
46	A fresh look at quantity sensitivity in Spanish	Bàrkányi	Z.	Linguistics	40	2
47	Spanish diminutive formation without rules or constraints	Eddington	David	Linguistics	40	2
48	Word reconstruction and consonant features in English and Spanish	Marks et al.		Linguistics	40	2
49	Semantic transparency as a predictor of copula choice in second-language acquisition	Geeslin	K.L.	Linguistics	40	2
50	Anaphora: lexico-textual structure, or means for utterance integration within a discourse? A critique of the functional-grammar account	Cornish	Francis	Linguistics	40	3
51	The gray area of supine clauses	Hill	Virginia	Linguistics	40	3
52	The role of morphology in gender determination: evidence from Modern Greek	Ralli	A.	Linguistics	40	3
53	Cognition, epistemic scale, and functions of the Old Japanese question particle ka	Shinzato	Rumiko	Linguistics	40	3
54	Semantic construction for reinterpretation phenomena	Egg	Markus	Linguistics	40	3
55	Conjunct/disjunct marking in Awa Pit	Curnow	T.J.	Linguistics	40	3
56	The Text of Performance and the Performance of Text in the Electronic Edition	Best	Michael	Computers and the humanities	36	3
57	Mutability, Medium and Character	Grigar	Dene	Computers and the humanities	36	3
58	Core Galore: Literary Theory and Computer Games	Rockwell	Geoffrey	Computers and the humanities	36	3
59	Computer-mediated Texts and Textuality: Theory and Practice	Schreibman	Susan	Computers and the humanities	36	3
60	A New Computer-assisted Literary Criticism?	Siemens	Raymond G.	Computers and the humanities	36	3
61	Animating the Language Machine: Computers and Performance	Soules	Marshall	Computers and the humanities	36	3
62	The Question Concerning Theory: Humanism, Subjectivity and Computing	Van Pelt	Tamise	Computers and the humanities	36	3

63	Industrial Text and French Neo-Structuralism	Winder	William	Computers and the humanities	36	3
64	Statistical Morphological Disambiguation for Agglutinative Languages	Tur et al.		Computers and the humanities	36	4
65	Stylistic Constancy and Change Across Literary Corpora: Using Measures of Lexical Richness to Date Works	Smith & Kelly		Computers and the humanities	36	4
66	Korean Combinatory Categorical Grammar and Statistical Parsing	Cha et al.		Computers and the humanities	36	4
67	Meta-Interpretation and Hypertext Fiction: A Critical Response	Gardner	Colin	Computers and the humanities	37	
68	Beyond the Web: TEI, the Digital Library, and the Ebook Revolution	Gibson & Ruotolo		Computers and the humanities	37	
69	Putting the Dialogue Back Together Re-Creating Structure in Letter Publishing	Eide	Øyvind	Computers and the humanities	37	
70	The Time Course of Language Change	Juola	Patrick	Computers and the humanities	37	
71	Analyzing the Order of Items in Manuscripts of The Canterbury Tales	Spencer et al.		Computers and the humanities	37	
72	Categorisation Techniques in Computer-Assisted Reading and Analysis of Texts (CARAT) in the Humanities	De Pasquale & Menunier		Computers and the humanities	37	
73	Research and Representation: The M.Phil in History and Computing	Spaeth	Donald	Computers and the humanities	37	
74	The Pompey Project: Digital Research and Virtual Reconstruction of Rome's First Theatre	Beacham & Denard		Computers and the humanities	37	
75	Another Perspective on Vocabulary Richness	Hoover	David L.	Computers and the humanities	37	
76	Vocabulary in Interviews as Related to Respondent Characteristics	Härnqvist et al.		Computers and the humanities	37	
77	The Were-Subjunctive in British Rural Dialects: Marrying Corpus and Questionnaire Data	Hardie & Mc Enery		Computers and the humanities	37	
78	Humble servants of the discipline? Self-mention in research articles	Hyland	Ken	ESP	20	3
79	Impersonality in the research article as revealed by analysis of the transivity structure	Martinez	Iliana A.	ESP	20	3
80	A comparative study of metaphor in Spanish and English financial reporting	Charteris-Black & Ennis		ESP	20	3
81	When professors don't cooperate: a critical perspective on EAP research	Braine	George	ESP	20	3
82	Using student performance data to develop an English course for clinical training	Shi et al.		ESP	20	3
83	Using computerized corpus analysis to investigate the textlinguistic discourse moves of a genre	Upton & Connor		ESP	20	4
84	Doctor and patient questions as a measure of doctor-centredness in UAE hospitals	Ibrahim	Y.	ESP	20	4
85	Aligning genre and practice with learning in Higher Education: an interdisciplinary perspective for course design and teaching	Mavor & Trayner		ESP	20	4
86	Designing EAP reading courses at the university level	Spector-Cohen et al.		ESP	20	4
87	Introduction to research articles: variations across disciplines	Samraj	B.	ESP	21	1

88	Visual discourse in scientific conference papers. A genre-based study	Rowley-Jolivet	Elizabeth	ESP	21	1
89	Oral Communication: the workplace needs and uses of business graduate employees	Crosling & Ward		ESP	21	1
90	From needs analysis to curriculum development: designing a course in health-care communication for immigrant students in the USA	Bosher & Smalkoski		ESP	21	1
91	Empirical examination of EFL reader's use of rhetorical information	Lahuerta Martinez	Ana Cristina	ESP	21	1
92	The WWW for academic purposes: old study skills for new?	Slaouti	Diane	ESP	21	2
93	Thesis and dissertation writing: an examination of published advice and actual practice	Paltridge	Brian	ESP	21	2
94	Analysing adjectives in scientific discourse: an exploratory study with educational applications for Spanish speakers at advanced university level	Soler	Viviana	ESP	21	2
95	Genre analysis of business letters of negotiation	Pinto dos Santos	V.B.M.	ESP	21	2
96	The fly's perspective discourse in the daily routine of a business manager	Louhila-Salminen	L.	ESP	21	3
97	Negotiating meaning in seminar-type discussion and EAP	Basturkmen	Helen	ESP	21	3
98	The China strategy: a tale of two case leaders	Jackson	J.	ESP	21	3
99	Strategy and self-regulation as contributors to improving students' cognitive model in an ESL program	Chien Ching	Lee	ESP	21	3
100	Developing legal writing materials for English second language learners: problems and perspectives	Candlin et al.		ESP	21	4
101	Dovetailing language and content: teaching balanced argument in legal problem answer writing	Bruce	Nigel	ESP	21	4
102	Knowledge and agency: a study of 'metaphenomenal discourse' in textbooks from three disciplines	Moore	Tim	ESP	21	4
103	"It is interesting to note that...": a comparative study of anticipatory 'it' in student and published writing	Hewings & Hewings		ESP	21	4
107	The baby, the bathwater and the ``language instinct" debate	Cowley	Stephen J.	Language Sciences	23	1
108	What's wrong with grammaticalization?	Campbell	Lyle	Language Sciences	23	
109	Is there such a thing as "grammaticalization?"	Joseph	Brian D.	Language Sciences	23	
120	Argument linking and the commercial transaction frame in English, Russian and Japanese	Croft et al.		Language Sciences	23	

## Annexe 5 : Balisage TEI

### Description générale du TEI header

→ en-tête TEI. Ne doit pas être confondu avec le *prologue* (déclaration XML ou SGML) et la DTD.

Le header fait l'objet d'un chapitre entier : <http://www.tei-c.org/P4X/HD.html>, et d'extension dans le tag set "Corpora".

Tout document conforme à la TEI doit être préfixé d'un *header* (teiHeader), identifiant et décrivant le contenu du document, ainsi que les stratégies d'encodage appliquées au texte.

Le teiHeader se découpe en quatre composantes principales :

- **<fileDesc>** description bibliographique complète du document (titre, auteur, éditeur, conditions juridiques de distribution, etc.) **SEULE COMPOSANTE OBLIGATOIRE**

- **<encodingDesc>** décrit la relation entre le texte balisé et sa(ses) source(s), c-à-d les stratégies d'encodage adoptées, les niveaux d'encodage, les balises utilisées, etc. Egalement les systèmes typologiques employés pour caractériser les différents textes.

- **<profileDesc>** informations classificatoires et contextuelles du texte (sujet, langues, situation de production, individus participants, etc.).

- **<revisionDesc>** mémoire du document, où sont consignées les différentes révisions.

Dans un corpus, le document peut avoir un "corpus header" en plus du header de chaque texte. Dans ce cas les deux headers n'ont pas la même fonction : le corpus header permet de noter ce qui s'applique à tous les textes, le header de chaque texte ne notant que ses spécificités ou des exceptions par rapport à ce qui est défini pour tous.

#### 1. <fileDesc>

Contient une description bibliographique complète du document. Trois éléments sont obligatoires, quatre sont facultatifs.

#### Composantes obligatoires

**<titleStmt>** (position 1 dans le header)

(title statement). Titre du document + les responsables de son contenu intellectuel.

**Balises :**

**<title>** pour le titre, les sous-titres d'un article, d'une revue, d'un livre, de séries... avec éventuellement les attributs **level** et **type** ; **<author>** pour le ou les auteur(s) ; **<respStmt>** Céline Poudat [mais que dans le header du corpus ?] avec deux éléments **<resp>** description de la nature de la responsabilité ; **<name>** nom propre, donc Céline Poudat ;

**<sponsor>**; **<funder>**; **<principal>** (chercheur principal responsable de la création d'un document électronique.

**<publicationStmt>** (position 4)

Informations concernant la publication ou la distribution du document électronique.

<publisher> nom de l'organisation responsable de la publication/distribution du document ; <distributor> nom de la personne ou autre responsable de la distribution ; <authority> autre autorité que le publisher ou distributor.

Au moins un des éléments ci-dessus doit être inclus :

<pubPlace> lieu de publication; <address> adresse postale ou autre; <idno> numéro standard ou non utilisé pour identifier un élément bibliographique – par exemple avec <type> ISBN; <availability> restrictions d'usage, distribution, copyright, etc. ; <date> date dans tout format.

### <sourceDesc> (position 7)

Informations sur le ou les documents source. Peut contenir une simple description ou une référence bibliographique du texte source.

<bibl> contient une référence bibliographique des éléments non balisés ; <biblStruct> liste des sources ; <biblFull> citation bibliographique entièrement structurée dans laquelle toutes les composantes du fichier TEI sont mentionnées ; <listBibl> listes bibliographiques en tout genre. *Ex.* :

```
<sourceDesc> <biblStruct lang='FR'><monogr><author>Eug&egrave;ne Sue</author> <title>Martin, l'enfant  
trouv&eacute;</title> <title type='sub'>M&eacute;moires d'un valet de chambre</title> <imprint> <pubPlace>Bruxelles  
et Leipzig</pubPlace> <publisher>C. Muquardt </publisher> <date value="1846">1846</date>  
</imprint></monogr></biblStruct></sourceDesc>
```

## Composantes optionnelles

### <editionStmt> (position 2)

(edition statement). Optionnel mais recommandé. Informations relatives à l'édition du texte.

<edition> particularités de l'édition du texte ; <respStmt> responsabilité particulière, hors auteur, éditeur, etc. avec <resp> et <name>.

### <extent> (position 3)

optionnel. Taille approximative du document électronique (taille du fichier, nombre de mots, phrases, etc.)

### <seriesStmt> (position 5)

Information concernant la ou les séries à laquelle le document appartient. Par exemple, les volumes des issues d'une revue, les revues des articles.

Doit contenir au moins l'un des éléments ci-dessous :

<title> pour le titre, les sous-titres d'un article, d'une revue, d'un livre, de séries... avec éventuellement <level> et <type> ; <idno> numéro standard ou non utilisé pour identifier un élément bibliographique – par exemple avec <type> ISBN ; <respStmt> responsabilité particulière, hors auteur, éditeur, etc. ; <resp> et <name> éventuellement avec <type> ;

### <notesStmt> (position 6)

Pour rajouter des notes pour un texte particulier. Pas employé pour le moment, mais infos sur <http://www.tei-c.org/P4X/HD.html>)

## 2. *<encodingDesc>*

Méthodes et stratégies d'encodage. Plusieurs des éléments suivants, dans l'ordre :

### **<projectDesc>** (position 1)

But et objectifs de l'encodage, petite description.

### **<samplingDecl>** (position 2)

Raisons et méthodes utilisées pour regrouper les textes, informations comme taille des échantillons (si normalisation par exemple), méthodes de sélection, etc.

### **<editorialDecl>** (position 3)

Principes éditoriaux utilisés pour l'encodage. Description ou jeu de balises : **<correction>** comment et dans quelles circonstances des corrections ont été effectuées dans le document → **<status>** et **<method>** mais surtout pour les textes scannés ; **<normalization>**; **<quotation>** ; **<hyphenation>** ; **<segmentation>** (...).

### **<tagsDecl>** (position 4)

Choix des balises.

**<rendition>** information sur l'interprétation des balises ; **<tagUsage>** information sur l'usage d'un élément spécifique dans un texte avec **<occurs>** nombre d'occurrences de l'élément dans le texte, **<ident>** nombre d'occurrences de l'élément dans le texte avec une valeur distincte de l'attribut global *id* et **<render>** interprétation de l'élément.

[utile pour marquer par exemple les italiques à valeur particulière, cf. marquent les exemples, les citations, les éléments d'origine étrangère, etc.]

**<refsDecl>** références canoniques ; **<classDecl>** taxonomies, codes classificatoires, cf. références Dewey (...).

## 3. *<profileDesc>*

Description détaillée des aspects non bibliographiques du texte, comme les langues et les sous-langages, la situation de production du texte, les participants et le contexte.

Trois composantes optionnelles :

### **<creation>**

Information concernant la création du texte. Par exemple :

**<creation>** **<date value="1992-08">**August 1992**</date>** **<rs type="city">**Taos, New Mexico**</rs>** **</creation>**

### **<langUsage>**

**<langUsage>** description des langues, sous-langages et registres des textes.

**<language>** caractérise une seule langue ou un seul sous-langage dans un texte avec éventuellement **<wsd>** *writing system declaration*, utile pour les exemples en langue rare et **<usage>** pourcentage approximatif du texte dans cette langue.

**<textClass>**

cf. taxonomies Dewey, mots-clés...

#### **4. <revisionDesc>**

Pas utilisé.



## Balises adoptées

### *Formatage/typographie*

De manière générale, `<hi rend= 'IT'>` pour les italiques, `<hi rend= 'BO'>` pour les gras et `<hi rend= 'UL'>` pour les soulignés. Le balisage est effectué dans Word, avec la fonction rechercher-remplacer. Une fois les éléments balisés, il est conseillé de supprimer leur formatage. Les éléments à la fois gras et soulignés par exemple doivent être balisés `<head rend="BO and UL">`.

Si le formatage porte sur un élément qui est déjà balisé (comme un titre, un exemple ou un élément de langue étrangère, etc.) : `<head rend="BO">`, `<foreign lang="DE" rend="IT">...`

### *Eléments de structuration du document*

#### Texte

De manière générale chaque texte est encadré des balises suivantes:

```
<TEI.2><teiHeader type="texte"></teiHeader><text>texte</text></TEI.2>
```

#### `<front>`, `<body>` et `<back>`

Structure par défaut de la TEI pour chaque texte. 3 niveaux généraux :

- `<front>` pour les éléments contenant des éléments précédant le texte, comme le titre, le nom de l'auteur, la préface, etc. ;
- `<body>` pour le texte, sans front ni back ;
- `<back>` pour les appendices, les annexes, etc. suivant la partie principale du texte, ainsi que pour les notes de bas de page.

#### `<head>`

Tous les titres et sous-titres sont encadrés de balises `<head>`, quels que soient leurs niveaux.

#### `<div>`

#### Balisage général des niveaux avec `<div>`

La structure de l'article est de manière générale marquée de balises `<div>`, le niveau est en attribut (plus simple pour automatiser les fermetures de niveaux – même balise `</div>` pour chaque fin de niveau, et non `<div1>`, `<div2>`, etc., plus complexe à automatiser).

*Ex.* : `<div type='1'>` pour les divisions de premier niveau, `<div type='2'>` pour les divisions de second niveau, etc.

**Balisage** : exploitation des niveaux dans Word pour le balisage. Certaines revues, comme RSP ou les cahiers de Praxématique ont une feuille de style. Dans ce cas, il convient

simplement de vérifier la pertinence des niveaux, puis de baliser les parties avec la fonction rechercher-remplacer (baliser chaque titre avec **<head>** en le faisant précéder de deux balises, l'une de fin de niveau, l'autre de début).

*Exemple* : [rechercher] titre de niveau 2 → [remplacer par] `</div><div type='2'><head>titre de niveau 2</head>`

### Attributs particulier des **<div>**

Balilage des introductions et des conclusions: **<div type='introduction'>**, **<div type='conclusion'>**. Adapter la valeur de l'attribut en fonction des spécificités des textes. Par exemple, **<div type='annonce de plan'>** (Cortès, CIEL 94).

### **<list>**

Sont considérées comme des *listes* les éléments précédés de puces, de tirets ou de numéros, *i.e.* les éléments formellement identifiables.

Utilisation du tag **<list>**. Les éléments de la liste ont été encodés avec **<item>**. Eventuellement, titre de la liste balisé (**<head>**). Les listes peuvent être imbriquées, si une même liste en contient d'autres.

### **<figure>**

Les tableaux, figures, schémas ou photos sont tous balisés **<figure>** avec un attribut précisant leur type.

Un tableau sera ainsi balisé de la manière suivante :

```
<figure type="tableau">  
  <figDesc />  
</figure>
```

**<figDesc>** étant ici vide. Si possible, inclure une courte description du tableau, ou de la figure. Si l'élément a un titre, l'inclure entre deux balises **<head>** avant **<figDesc>**.

Les éléments doivent être supprimés du texte et stockés dans un fichier séparé (un fichier par revue). Les éléments seront conservés dans l'ordre, et par article.

### Références bibliographiques

Placées dans la partie **<back>** du texte, les références bibliographiques à la fin des articles sont balisées avec le tag **<listBibl>**, suivi du titre de la bibliographie (références, bibliographie, etc.) avec **<head>**, contenant une liste de **<bibl>** avec un identifiant NOM\_DE\_L\_AUTEUR CODE ANNEE (deux chiffres). Par exemple : **<bibl id='DIDAY82'>**. Les éléments auteurs (tag **<author>**) et titres (tag **<title>**) sont balisés.

Balilage d'une référence à un ouvrage :

```
<bibl id='DIDAY82'><author>E. Diday et al.</author><title>éléments d'analyse de  
données</title>Paris : Dunod, Bordas, 1982</bibl>
```

Balilage d'une référence à un article avec **<analytic>** et **<monogr>** (si possible, sinon, balilage identique à celui de l'ouvrage) :

```
<biblStruct id="NEL80">
```

<analytic>

<author>Nelson, T. H.</author>

<title>Replacing the printed word: a complete literary system.</title>

</analytic>

<monogr>

<title>Information Processing '80: Proceedings of the IFIPS Congress, October 1980</title>

<editor>Simon H. Lavington</editor>

<imprint>

<publisher>North-Holland</publisher>

<pubPlace>Amsterdam</pubPlace>

<date>1980</date>

</imprint>

<biblScope>pp 1013&ndash;23 </biblScope>

</monogr>

Si possible, balisage des références dans le texte.

### Exemple :

He was a member of Parliament for Warwickshire in 1445, and died March 14, 1470 (according to <bibl><author>Kittredge</author>, <title>Harvard Studies</title> <biblScope>5. 88ff</biblScope></bibl>).

### <note>

Les notes sont placées dans la partie <back> du texte. <note type= 'bas de page'> ou <note type= 'fin d'article'>.

Une balise <ptr /> est positionnée autour de chaque renvoi, et au début de chaque note (un compteur permettra ensuite de les retrouver dans les textes). Balisage également facilement automatisable avec Word.

### <p>

<p> pour les paragraphes. Il est préférable de les baliser à la fin, afin d'éviter les éventuels chevauchements de balises.

## Citations

Utilisation générale de la balise <q> pour les citations, ce qui permet de baliser également les éléments marqués de tirets, en italique dans le texte.

Idéalement, il faudrait distinguer les citations attribuées par l'auteur à une source externe au texte (le plus souvent en italique, ou marqués par un retrait, une police particulière – généralement plus petite) avec la balise <quote>, et les éléments non pris en charge par l'auteur (et souvent marqués en italique, mais également entre guillemets) avec la balise <q>.

## Exemples

Les exemples, non documentés par la TEI, ont un statut d'attribut. S'ils ont un statut de paragraphe : `<p ana="exemplum">`, sinon utilisation du tag `<seg>` : `<seg ana="exemplum">`.

Distinction des exemples repris d'autres textes avec : **"exemplum quoted"** (exemples généralement suivis d'une référence).

```
<p ana="exemplum quoted">
```

**(9c) « Les divergences de conception de la tâche au sein des groupes de l'article 2, qui interviennent dans la répartition leader/collaborateur(s), et d'un groupe à l'autre, aboutissent à des réalisations et à des types d'échanges différents (...) » (J. Arditti et M. Vasseur,**

```
<hi rend="IT">Langages</hi>
```

```
134, 1999 :17).
```

```
</p>
```

Les autres types d'exemples seront simplement balisés avec **"exemplum"**.

## Autres éléments

- `<foreign>` pour les éléments du texte de langue différente, avec la langue en attribut, cf. `<foreign lang='FR'>` (voir la documentation des langues dans le header du corpus) ;
- `<c type="symbole">?</c>` pour les symboles linguistiques, comme \*, ?, etc. précédant des énoncés le plus souvent à valeur d'exemple.

## Annexe6 : Extrait de texte balisé (début du texte 19)

```
- <TEI.2>
```

```
- <teiheader type="texte">
```

```
- <fileDesc>
```

```
- <titleStmt>
```

```
<title>Version numérisée de Un exemple de catégorisation des noms, des verbes et des adjectifs par le langage naturel : les adverbes allemands en -weise de C. Cortès et Helge Szabó</title>
```

```
- <respStmt>
```

```
<resp>Collection et balisage par</resp>
```

```
<name>Céline Poudat</name>
```

```
</respStmt>
```

```
</titleStmt>
```

```
- <publicationStmt>
```

```
<distributor>Céline Poudat</distributor>
```

```
</publicationStmt>
```

```
- <sourceDesc>
```

```
- <biblStruct>
```

```
- <analytic>
```

```
<author>Colette Cortès et Helge Szabó</author>
```

```
<title>Un exemple de catégorisation des noms, des verbes et des adjectifs par le langage naturel : les adverbes allemands en -weise</title>
```

```

</analytic>
- <monogr>
  <title>Cahiers du CIEL</title>
- <imprint>
  <biblScope type="num">94-95</biblScope>
</imprint>
  <biblScope type="pages">31</biblScope>
</monogr>
</biblStruct>
</sourceDesc>
</fileDesc>
</teiheader>
- <text>
- <front>
- <head rend="BO">
  Un exemple de catégorisation des noms, des verbes et des adjectifs par le
  langage naturel : les adverbes allemands en
  <hi rend="IT">-weise</hi>
  <ptr />
</head>
- <p>
  <hi rend="BO">Colette CORTÈS</hi>
</p>
- <p>
  <hi rend="BO">C.I.E.L.</hi>
</p>
- <p>
  <hi rend="BO">Université de Paris 7</hi>
</p>
- <p>
  <hi rend="BO">Helge SZABÓ</hi>
</p>
- <p>
  <hi rend="BO">Université de Paris IV</hi>
</p>
</front>
- <body>
- <div type="introduction">
  <p>Les grandes classifications, notamment nominales, sont construites
  habituellement à partir de "classes d'objets" (classifications dans les domaines
  de la biologie et de la chimie, champs lexicaux par exemple des liens de
  parenté, etc... ). Nous étudierons ici une classification inscrite dans le langage
  naturel car elle sert de fondement à la dérivation d'adverbes allemands
  productifs actuellement. Nous verrons que, dans tous les mécanismes de
  catégorisation par le langage naturel que nous décrirons, paradigmatique et
  syntagmatique sont indissociables.</p>
- <p>
  Nous considérons que les adverbes allemands qui se terminent par
  <hi rend="IT">-weise</hi>
  sont formés d'un lexème-racine, d'un élément de jonction que nous
  appellerons désormais : "joncteur" (de l'allemand "Fugen-Element" = fr.
  joncteur, marqueur de jonction ) et du suffixe -
  <hi rend="IT">weise</hi>
  (ex.

```

**<hi rend="IT">kilo-weise (au kilo ), stell-en-weise( par endroits, par ci, par là ), beispiel-s-weise (par exemple ), klug-er-weise (intelligemment avec jugement du locuteur sur le dire</hi>**

**). Une observation minutieuse de ces adverbes fait apparaître, sur le plan fonctionnel, plusieurs classes de comportement des adverbes eux-mêmes, auxquelles correspondent, sur le plan lexical, plusieurs classes de lexèmes-racines sur lesquels les adverbes sont formés: nous assistons ainsi à une sorte de catégorisation des lexèmes-racines par le langage naturel, dont nous voudrions dégager dans cet article tous les paramètres. Le classement auquel nous parvenons s'appuie sur deux ensembles de tests syntaxiques, sémantiques et pragmatiques, l'un portant sur les adverbes, l'autre sur les racines ; il montre de façon étonnante que les joncteurs sont des marquages relativement réguliers des classes de comportement à la fois des adverbes, et des lexèmes-racines.**

**</p>**

**<p>Au cours du travail, nous (re )définirons les notions de : adverbe, racine, suffixe et joncteur.</p>**

**<p>Plan</p>**

**- <p>**

**1. Classification des adverbes. Tests syntaxiques, sémantiques et pragmatiques portant sur le statut des adverbes en**

**<hi rend="IT">-weise</hi>**

**.**

**</p>**

**<p>1.1. Modèle d'emploi des adverbes et exemples attestés.</p>**

**<p>1.2. Commentaire sur les tests.</p>**

**<p>2. Classification des racines.</p>**

**<p>2.1. Classes de noms.</p>**

**<p>2.2. Classes de verbes</p>**

**<p>2.3. Classes d'adjectifs</p>**

**<p>2. 4. Catégorisation "naturelle".</p>**

**- <p>**

**3. Valeur sémantico-syntaxique du suffixe**

**<hi rend="IT">-weise</hi>**

**. et rôle du joncteur.**

**</p>**

**- <p>**

**3.1. Valeur du suffixe -**

**<hi rend="IT">weise</hi>**

**</p>**

**- <p>**

**3.2. Joncteur**

**<ptr />**

**</p>**

**</div>**

**- <div type="1">**

**- <head rend="BO and small caps">**

**1 Tests syntaxiques, sémantiques et pragmatiques portant sur les adverbes en**

**<hi rend="IT">-weise</hi>**

**</head>**

**<p>Des tests nous ont permis de distinguer plusieurs classes d'adverbes, en fonction de critères syntaxiques (emploi comme épithète (test 3 ) incidence de l'adverbe (tests 9, 10, 11 ) ), de critères sémantiques ( gradation (test 1 ) paraphrase (test 2 ) ou de critères pragmatiques (tests 4, 5, 6, 7, 8 ) ).</p>**

**<p>Ils nous ont permis de distinguer 8 classes de comportement, comme il apparaît dans le tableau 1. Pour faciliter la lecture, nous ferons référence à chaque classe de comportement par le numéro qui lui correspond dans le classement définitif, mais nous voudrions insister sur le fait que ce tableau est, dans la chronologie de notre travail, un résultat et non un point de départ.</p>**

**<p>Nous présenterons tout d'abord un exemple de chaque type d'adverbe employé dans une phrase (les phrases numérotées sont des phrases attestées, les autres constituent une sorte de cadre prototypique d'emploi de l'adverbe ), puis nous donnerons un tableau des résultats des tests et enfin un commentaire sur chaque type de test.</p>**

**- <div type="2">**

**<head rend="BO and IT">1.1. Modèles d'emploi des adverbes suivis d'exemples attestés</head>**

**- <p>**

**Classe 1**

**<ptr />**

**</p>**

**<p>Das Material</p>**

**- <p>**

**Die Ware wird**

**<hi rend="UL">Xweise</hi>**

**verkauft/ verwendet.**

**</p>**

**- <p>**

**- <seg ana="exemplum">**

**(1 ) Die spanische Seite war verärgert, das Presseecho verheerend. Radio COPE, vielgehört, sprach von bösem Imperialismus: "Wir haben Spanien**

**<hi rend="UL">scheibchenweise</hi>**

**verkauft, jetzt erhalten wir die Quittung."**

**<hi rend="IT">Die Zeit. Nr 25. 17. Juni 1994</hi>**

**.**

**</seg>**

**</p>**

**[...]**

## ***Annexe 7 : Exemples de sorties des étiqueteurs Brill, MBT et TnT après la première phase d'entraînement***

### **Brill Tagger**

Antoine/NC:sg Culioli /NC:sg A/ADV:1st propos/ADV:2nd de/PREP même /NC:sg Order/FGW is/FGW not/FGW sufficient/VER:pres . /ABR What/FGW is/FGW required/FGW is/FGW something/FGW much/FGW more/FGW complex/FGW . /ABR It/FGW is/FGW order/FGW entering/FGW upon/FGW novelty/FGW ; /NN so/FGW that/FGW the/FGW massiveness/FGW of/FGW order/FGW does/FGW not/FGW degenerate/FGW into/FGW mere/FGW repetition/FGW ; /NN and/FGW so/FGW that/FGW the/FGW novelty/FGW is/FGW always/FGW reflected/FGW upon/FGW a/FGW background/FGW system/FGW . /ABR A./ABR N./ABR Whitehead/FGW ,/PON:comma Process/FGW and/FGW reality/FGW . /ABR /NN Dans/PREP un/DET:indef article/NC:sg intitulé/VER:pper Le/DET:def système/NC:sg sublogique/ADJ:sg des/DTC:pl prépositions/NC:pl en/PREP latin/ADJ:sg ,/PON:comma publié/VER:pper en/PREP 1949/NUM:dat (/PON repris/VER:pper dans/PREP le/PRO:clit tome/VER:pres 1/NUM:car

de/PREP Problèmes/NC:pl de/PREP linguistique/ADJ:sg générale/ADJ:sg ,/PON:comma Paris/NP ,/PON:comma 1996/NUM:dat )/PON ,/PON:comma E./ABR Benveniste/NP écrit/VER:pres en/PREP conclusion/NC:sg : /SYM "/PON Tous/DET:indef les/DET:def emplois/NC:pl de/PREP prae/NC:sg se/PRO:refl tiennent/VER:pres ainsi/ADV dans/PREP une/DET:indef définition/NC:sg constante/ADJ:sg . /ABR Nous/PRO:pp1pl avons/VER:aux:pres voulu/VER:pper montrer/VER:inf sur/PREP un/DET:indef exemple/NC:sg que/PRO:rel ,/PON:comma dans/PREP l'étude/NC:sg des/DTC:pl prépositions/NC:sg quels/ADV:int que/SUB:2nd soient/VER:subp l'idiome/NC:sg et/COO l'époque/NC:sg considérés/VER:pper ,/PON:comma une/DET:indef nouvelle/ADJ:sg technique/ADJ:sg de/PREP la/DET:def description/NC:sg est/VER:pres nécessaire/ADJ:sg et/COO devient/VER:pres possible/ADJ:sg pour/SUB restituer/VER:inf la/DET:def structure/NC:sg de/PREP chacune/PRO:indef des/DTC:pl prépositions/NC:pl et/COO intégrer/VER:inf ces/DET:dem structures/NC:pl dans/PREP un/DET:indef système/NC:sg général/ADJ:sg . /ABR La/DET:def tâche/NC:sg entraîne/VER:pres l'/DET:def obligation/NC:sg de/PREP réinterpréter/VER:inf toutes/DET:indef les/DET:def données/NC:pl acquises/ADJ:pl et/COO de/PREP refondre/VER:inf les/DET:def catégories/NC:pl établies/ADJ:pl ./PON:sep " /NN Dans/PREP l'analyse/NC:sg qui/PRO:rel suit/VER:pres ,/PON:comma j'essaierai/VER:fut de/PREP poursuivre/VER:inf à/ADV:1st propos/ADV:2nd de/PREP même/ADV l'/DET:def entreprise/NC:sg (/PON la/DET:def "/PON tâche/NC:sg "/PON )/PON que/SUB Benveniste/NP posait/VER:impf comme/ADV nécessaire/ADJ:sg . /ABR Ce/PRO:dem faisant/PARPRES ,/PON:comma je/PRO:pp1sn m'abstiendrai/VER:fut ,/PON:comma autant/ADV que/SUB possible/ADJ:sg ,/PON:comma de/PREP remarques/NC:pl générales/ADJ:pl d'ordre/NC:sg théorique/ADJ:sg et/COO j'adopterai/VER:fut un/DET:indef mode/NC:sg de/PREP présentation/NC:sg métalinguistique/ADJ:sg qui/PRO:rel ne/ADV:neg:1st devrait/VER:cond pas/ADV:neg:2nd dérouter/VER:inf le/DET:def lecteur/NC:sg pour/PREP peu/ADV qu'il/ADJ:sg accepte/VER:pres de/PREP lire/VER:inf le/DET:def texte/NC:sg . /ABR /NN I./ABR Je/PRO:pp1sn commencerai/VER:fut par/PREP étudier/VER:inf les/DET:def emplois/NC:pl de/PREP même/ADJ:sg postposé/VER:pper . /ABR Voici/ADV quelques/DET:indef exemples/NC:pl . /ABR Nous/PRO:pp1pl nous/PRO:clit rencontrerons/VER:fut aujourd'hui/ADV même/ADJ:sg et/COO ici/ADV même/ADJ:sg . /ABR C'/PRO:dem est/VER:pres le/DET:def genre/NC:sg même/ADJ:sg de/PREP bibelot/NC:sg que/SUB je/PRO:pp1sn déteste/VER:pres . /ABR L'/DET:def idée/NC:sg même/ADJ:sg de/PREP les/PRO:clit rencontrer/VER:inf me/PRO:clit fait/VER:pres horreur/NC:sg . /ABR L'/DET:def opération/NC:sg marquée/VER:pper par/PREP même/ADJ:sg peut/VER:mod:pres être/VER:aux:inf décrite/VER:pper comme/ADV suit/VER:pres : /NN (a)/LS II/PRO:pp3msn s'/PRO:refl agit/VER:pres d'/PREP opposer/VER:inf une/DET:indef (/PON ou/COO plusieurs/DET:indef )/PON occurrence/NC:sg (/PON s/SYM )/PON de/PREP temps/NC:sg ,/PON:comma lieu/NC:sg ,/PON:comma objet/NC:sg ,/PON:comma représentation/NC:sg subjective/ADJ:sg ,/PON:comma etc./ABR à/PREP d'autres/DET:indef occurrences/NC:pl dans/PREP un/DET:indef certain/DET:indef domaine/NC:sg . /ABR

## MBT Tagger

Antoine//UNKNOWN Culioli//NP A//UNKNOWN propos//UNKNOWN  
de//UNKNOWN même//PRO:refl Order//UNKNOWN is//UNKNOWN not//UNKNOWN  
suffisant//ADJ:sg ./PON:sep What//UNKNOWN is//UNKNOWN required//UNKNOWN  
is//UNKNOWN something//UNKNOWN much//UNKNOWN more//ADJ:sg  
complex//UNKNOWN It//UNKNOWN is//UNKNOWN order//UNKNOWN  
entering//UNKNOWN upon//UNKNOWN novelty//UNKNOWN ;//UNKNOWN  
so//UNKNOWN that//UNKNOWN the//UNKNOWN massiveness//UNKNOWN



of//UNKNOWN order//UNKNOWN does//UNKNOWN not//UNKNOWN  
degenerate//UNKNOWN into//UNKNOWN mere//UNKNOWN repetition//UNKNOWN  
://UNKNOWN and//UNKNOWN so//UNKNOWN that//UNKNOWN the//UNKNOWN  
novelty//UNKNOWN is//UNKNOWN always//UNKNOWN reflected//UNKNOWN  
upon//UNKNOWN a//UNKNOWN background//UNKNOWN system//UNKNOWN  
./PON:sep A//UNKNOWN ./PON:sep N//UNKNOWN ./PON:sep Whitehead//UNKNOWN  
./PON:comma Process//UNKNOWN and//UNKNOWN reality//UNKNOWN  
Dans//UNKNOWN un//UNKNOWN article//UNKNOWN intitulé//UNKNOWN  
Le//UNKNOWN système//UNKNOWN sublogique//ADJ:sg des//UNKNOWN  
prépositions//UNKNOWN en//UNKNOWN latin//UNKNOWN ./PON:comma  
publié//ADJ:sg en//UNKNOWN 1949//UNKNOWN (/UNKNOWN repris//UNKNOWN  
dans//UNKNOWN le//ADJ:sg tome//UNKNOWN 1//UNKNOWN de//UNKNOWN  
Problèmes//UNKNOWN de//UNKNOWN linguistique//UNKNOWN générale//ADJ:sg  
./PON:comma Paris//UNKNOWN ./PON:comma 1996//UNKNOWN )//UNKNOWN  
./PON:comma E//UNKNOWN ./PON:sep Benveniste//NP écrit//UNKNOWN  
en//UNKNOWN conclusion//UNKNOWN ://UNKNOWN "//UNKNOWN Tous//DET:indef  
les//DET:def emplois//UNKNOWN de//UNKNOWN prae//UNKNOWN se//ADJ:sg  
tiennent//UNKNOWN ainsi//UNKNOWN dans//UNKNOWN une//DET:indef  
définition//UNKNOWN constante//ADJ:sg ./PON:sep Nous//UNKNOWN avons//UNKNOWN  
voulu//UNKNOWN montrer//UNKNOWN sur//UNKNOWN un//UNKNOWN  
exemple//UNKNOWN que//ADJ:sg ./PON:comma dans//UNKNOWN l'étude//UNKNOWN  
des//UNKNOWN prépositions//UNKNOWN quels//UNKNOWN que//ADJ:sg  
soient//UNKNOWN l'idiome//UNKNOWN et//ADJ:sg l'époque//ADJ:sg  
considérés//UNKNOWN ./PON:comma une//DET:indef nouvelle//UNKNOWN  
technique//ADJ:sg de//ADJ:sg la//DET:def description//UNKNOWN est//UNKNOWN  
nécessaire//ADJ:sg et//UNKNOWN devient//UNKNOWN possible//ADJ:sg  
pour//UNKNOWN restituer//UNKNOWN la//DET:def structure//UNKNOWN  
de//UNKNOWN chacune//UNKNOWN des//UNKNOWN prépositions//UNKNOWN  
et//UNKNOWN intégrer//UNKNOWN ces//UNKNOWN structures//UNKNOWN  
dans//UNKNOWN un//UNKNOWN système//UNKNOWN général//ADJ:sg ./PON:sep La//LS  
tâche//ADJ:sg entraîne//UNKNOWN l'//DET:def obligation//UNKNOWN de//UNKNOWN  
réinterpréter//UNKNOWN toutes//UNKNOWN les//DET:def données//UNKNOWN  
acquises//UNKNOWN et//UNKNOWN de//UNKNOWN refondre//UNKNOWN les//DET:def  
catégories//UNKNOWN établies//UNKNOWN ./PON:sep "//UNKNOWN Dans//UNKNOWN  
l'analyse//UNKNOWN qui//UNKNOWN suit//UNKNOWN ./PON:comma  
j'essaierai//UNKNOWN de//UNKNOWN poursuivre//UNKNOWN à//UNKNOWN  
propos//UNKNOWN de//UNKNOWN même//PRO:refl l'entreprise//UNKNOWN  
(/UNKNOWN la//DET:def "//UNKNOWN tâche//UNKNOWN "//UNKNOWN  
)//UNKNOWN que//UNKNOWN Benveniste//NP posait//UNKNOWN comme//UNKNOWN  
nécessaire//ADJ:sg ./PON:sep Ce//UNKNOWN faisant//UNKNOWN ./PON:comma  
je//UNKNOWN m'abstiendrai//UNKNOWN ./PON:comma autant//UNKNOWN que//ADJ:sg  
possible//ADJ:sg ./PON:comma de//UNKNOWN remarques//UNKNOWN  
générales//UNKNOWN d'ordre//UNKNOWN théorique//ADJ:sg et//UNKNOWN  
j'adopterai//UNKNOWN un//UNKNOWN mode//ADJ:sg de//ADJ:sg  
présentation//UNKNOWN métalinguistique//ADJ:sg qui//UNKNOWN ne//ADJ:sg  
devrait//UNKNOWN pas//UNKNOWN dérouter//UNKNOWN le//ADJ:sg  
lecteur//UNKNOWN pour//UNKNOWN peu//UNKNOWN qu'il//ADJ:sg  
accepte//UNKNOWN de//UNKNOWN lire//UNKNOWN le//ADJ:sg texte//UNKNOWN  
./PON:sep I//UNKNOWN Je//NP commencerai//UNKNOWN par//UNKNOWN  
étudier//UNKNOWN les//DET:def emplois//UNKNOWN de//UNKNOWN même//PRO:refl

postposé//UNKNOWN Voici//UNKNOWN quelques/DET:indef exemples/UNKNOWN  
./PON:sep Nous/UNKNOWN nous/UNKNOWN rencontrerons//UNKNOWN  
aujourd'hui//UNKNOWN même//PRO:refl et//UNKNOWN ici//UNKNOWN même//ADJ:sg  
./PON:sep C'est//UNKNOWN le//ADJ:sg genre/UNKNOWN même//PRO:refl  
de//UNKNOWN bibelot//UNKNOWN que//ADJ:sg je//UNKNOWN déteste//ADJ:sg  
./PON:sep L'idée//UNKNOWN même//PRO:refl de//UNKNOWN les/DET:def  
rencontrer//UNKNOWN me//ADJ:sg fait//UNKNOWN horreur/UNKNOWN ./PON:sep  
L'opération//UNKNOWN marquée/UNKNOWN par//UNKNOWN même//PRO:refl  
peut//UNKNOWN être/UNKNOWN décrite/UNKNOWN comme/UNKNOWN  
suit/UNKNOWN ://UNKNOWN (a)//NUM:par II/UNKNOWN s'agit//UNKNOWN  
d'opposer//UNKNOWN une/DET:indef (/UNKNOWN ou//UNKNOWN plusieurs/DET:indef  
)//LS occurrence(/UNKNOWN s//UNKNOWN )//NUM:par de//UNKNOWN  
temps//UNKNOWN ./PON:comma lieu/UNKNOWN ./PON:comma objet//UNKNOWN  
./PON:comma représentation/UNKNOWN subjective/ADJ:sg ./PON:comma  
etc//UNKNOWN à//UNKNOWN d'autres//UNKNOWN occurrences//UNKNOWN  
dans//UNKNOWN un//ADJ:sg certain/DET:indef domaine//UNKNOWN ./PON:sep

<utt>

## TnT Tagger

Antoine NP Culioli NP A PREP propos NC:pl de PREP même ADJ:sg Order FGW is FGW  
not FGW sufficient NC:sg . PON:sep What FGW is FGW required FGW is FGW something  
FGW much FGW more FGW complex. FGW It FGW is FGW order FGW entering FGW  
upon FGW novelty FGW ; FGW so FGW that FGW the FGW massiveness FGW of FGW  
order FGW does FGW not FGW degenerate FGW into FGW mere FGW repetition FGW ;  
FGW and FGW so FGW that FGW the FGW novelty FGW is FGW always FGW reflected  
FGW upon FGW a FGW background FGW system FGW . PON:sep A SYM . PON:sep N  
SYM . PON:sep Whitehead FGW , PON:comma Process FGW and FGW reality. FGW Dans  
PREP un DET:indef article NC:sg intitulé VER:pper Le DET:def système NC:sg sublogique  
ADJ:sg des DTC:pl prépositions NC:pl en PREP latin NC:sg , PON:comma publié VER:pper  
en PREP 1949 NUM:dat ( PON repris VER:pper dans PREP le PRO:clit tome VER:pres 1  
NUM:car de PREP Problèmes NC:pl de PREP linguistique NC:sg générale ADJ:sg ,  
PON:comma Paris NP , PON:comma 1996 NUM:dat ) PON , PON:comma E. ABR  
Benveniste NP écrit VER:pres en PREP conclusion NC:sg : PON " PON Tous DET:indef les  
DET:def emplois NC:pl de PREP prae NC:sg se PRO:refl tiennent VER:pres ainsi ADV dans  
PREP une DET:indef définition NC:sg constante ADJ:sg . PON:sep Nous PRO:pp1pl avons  
VER:aux:pres voulu VER:pper montrer VER:inf sur PREP un DET:indef exemple NC:sg que  
PRO:rel , PON:comma dans PREP l'étude NC:sg des DTC:pl prépositions NC:pl ,  
PON:comma quels ADV:int que SUB:2nd soient VER:aux:subp l'idiome VER:pper et COO  
l'époque NC:sg considérés VER:pper , PON:comma une DET:indef nouvelle ADJ:sg  
technique ADJ:sg de PREP la DET:def description NC:sg est VER:pres nécessaire ADJ:sg et  
COO devient VER:pres possible ADJ:sg pour SUB restituer VER:inf la DET:def structure  
NC:sg de PREP chacune PRO:indef des DTC:pl prépositions NC:pl et COO intégrer VER:inf  
ces DET:dem structures NC:pl dans PREP un DET:indef système NC:sg général ADJ:sg .  
PON:sep La DET:def tâche NC:sg entraîne VER:pres l' DET:def obligation NC:sg de PREP  
réinterpréter VER:inf toutes DET:indef les DET:def données NC:pl acquises ADJ et COO de  
PREP refondre VER:inf les DET:def catégories NC:pl établies ADJ . PON:sep " PON Dans  
PREP l'analyse NC:sg qui PRO:rel suit VER:pres , PON:comma j'essaierai VER:fut de PREP  
poursuivre VER:inf à PREP propos NC:pl de PREP même ADJ:sg l'entreprise NC:sg ( PON  
la DET:def " PON tâche NC:sg " PON ) PON que SUB Benveniste NP posait VER:impf

comme ADV nécessaire ADJ:sg . PON:sep Ce PRO:dem faisant PARPRES , PON:comma je PRO:pp1sn m'abstiendrai VER:fut , PON:comma autant ADV que SUB possible ADJ:sg , PON:comma de PREP remarques NC:pl générales ADJ d'ordre VER:pres théorique ADJ:sg et COO j' PRO:pp1sn adopterai VER:fut un DET:indef mode NC:sg de PREP présentation NC:sg métalinguistique ADJ:sg qui PRO:rel ne ADV:neg:1st devrait VER:mod:cond pas ADV:neg:2nd dérouter VER:inf le DET:def lecteur, ??? pour PREP peu ADV qu'il ADJ:sg accepte VER:pres de PREP lire VER:inf le DET:def texte NC:sg . PON:sep I. ABR Je NP commencerai VER:fut par PREP étudier VER:inf les DET:def emplois NC:pl de PREP même ADJ:sg postposé ADJ:sg . PON:sep Voici ADV quelques DET:indef exemples NC:pl . PON:sep Nous PRO:pp1pl nous PRO:clit rencontrerons VER:fut aujourd'hui ADV même ADV et COO ici ADV même ADJ:sg . PON:sep C'est ADV le DET:def genre NC:sg même ADJ:sg de PREP bibelot FGW que SUB je PRO:pp1sn déteste VER:pres . PON:sep L'idée NC:sg même ADJ:sg de PREP les PRO:clit rencontrer VER:inf me PRO:clit fait VER:pres horreur NC:sg . PON:sep L'opération NC:sg marquée VER:pper par PREP même ADJ:sg peut VER:mod:pres être VER:aux:inf décrite VER:pper comme ADV suit VER:pres : PON (a) LS Il PRO:pp3msn s'agit VER:pres d'opposer VER:inf une DET:indef ( PON ou COO plusieurs DET:indef ) PON occurrence NC:sg ( PON s SYM ) PON de PREP temps NC:pl , PON:comma lieu NC:sg , PON:comma objet NC:sg , PON:comma représentation NC:sg subjective ADJ:sg , PON:comma etc. ABR à PREP d'autres ADJ occurrences NC:pl dans PREP un DET:indef certain DET:indef domaine NC:sg . PON:sep

## *Annexe 8 : Etiqueteurs recensés*

### **BRILL**

*Sources:*

- Page de Eric Brill : <http://www.cs.jhu.edu/~brill>

WinBrill – InaLF:

[http://www.inalf.cnrs.fr/scripts/mep.exe?HTML=mep\\_winbrill.txt?OUVRIR\\_MENU=1](http://www.inalf.cnrs.fr/scripts/mep.exe?HTML=mep_winbrill.txt?OUVRIR_MENU=1)

- English POS Tagger en ligne: [http://www.cst.dk/online/pos\\_tagger/uk/index.html](http://www.cst.dk/online/pos_tagger/uk/index.html)
- Lecomte, Josette. *Le Catégoriseur Brill 14-JL5/WinBrill-0.3*. InaLF/CNRS, décembre 1998.

### *Présentation générale de l'étiqueteur*

Créé par Eric Brill dans le cadre de sa thèse (1993, université de Pennsylvanie), le tagger Brill est fondé sur les travaux de Bloomfield (1933) et Harris (1946, 1951). Reposant sur l'idée que l'étude d'une langue peut se fonder sur l'observation de faits linguistiques et indépendamment d'une théorie linguistique particulière, le tagger doit, pour fonctionner, être entraîné sur un corpus de taille restreinte étiqueté manuellement et à partir duquel il infère des règles d'étiquetage (distribution "extensionnelle"). Les mots inconnus sont traités à partir d'une hypothèse naïve sur la structure du langage Enfin, une analyse de la distribution est effectuée afin de réduire les erreurs d'étiquetage.

Il est possible d'entraîner Brill sur tout type de corpus étiqueté<sup>3</sup> ; le tagger présente un intérêt incontestable pour étudier les genres dans la mesure où il est possible de l'adapter au traitement d'un genre particulier, celui de l'article scientifique par exemple et de choisir le jeu d'étiquettes souhaité. Brill est effectivement davantage adapté au codage des textes en "discours" qu'en "langue", comme le souligne Lecomte dans sa description de Brill (Inalf). En outre, l'étiqueteur est multilingue et fonctionne avec d'autres langues que l'anglais.

En ce qui concerne le français, c'est l'Inalf qui a entraîné Brill sur la base Frantext, et il est possible de signer une convention pour pouvoir profiter des résultats de cet apprentissage.

### ***Fonctionnement de l'étiqueteur (Lecomte, chapitre 4, p. 22)***

À partir d'un sous-corpus échantillon étiqueté manuellement, le système crée sa base de connaissances. Quatre fichiers serviront de paramètres dans la commande d'étiquetage:

1. **fichier LEXIQUE**: contient une liste de mots (une entrée par ligne) associés à une liste de codes pouvant contenir tout caractère hormis l'espace (séparateur). Le premier code associé est "*le plus probable*". Ceux qui suivent sont ceux qui ont été également rencontrés, pas de contrainte d'ordre. La taille maximale du lexique semble limitée à 100 000 mots, la taille du mot à 100 caractères et celle d'une étiquette à 20 caractères.

#### ***Exemples:***

réduit	VCJ:sg SBC:sg ADJ1PAR:sg VPAR:sg ADJ2PAR:sg
bon_gré_,_mal_gré	ADV
répondant	VNCNT SBC:sg
actives	ADJ:pl
indiquent	VCJ:pl
cette_fois_-_ci	ADV
énonça	VCJ:sg
*Catherine	SBP:sg

2. **fichier des REGLES LEXICALES**: traitement des mots inconnus.
3. **fichier des REGLES CONTEXTUELLES**: règles contextuelles qui serviront à affiner l'étiquetage en contexte (local assez réduit: 6 max. – étiquettes PD ou mots).

#### ***Exemples:***

DTN:sg PRV:sg NEXTTAG	VCJ:sg
PRV:sg PRV:pl NEXT1OR2TAG	VCJ:pl
ADJ2PAR:sg ADJ1PAR:sg PREV1OR2OR3TAG	ECJ:sg
SBC:sg ADJ:sg PREVTAG	SBC:sg
PRV:sg PRO:sg WDPREVTAG	PREP elle

<sup>3</sup> Pour des langues positionnelles plutôt que casuelles.

VPAR:sg ADJ1PAR:sg PREV1OR2OR3TAG ECJ:sg DTN:sg PRO:sg NEXTTAG REL SBC:sg VCJ:sg PREV1OR2TAG PRV:sg ADJ2PAR:sg VPAR:sg PREV1OR2OR3TAG ACJ:sg SUB\$ SUB WDPREVTAG PREP que
---

4. **fichier de BIGRAMMES** (= **paire de mots adjacents**): traitement des mots inconnus, contient la liste de toutes les paires de mots repérées dans le corpus à étiqueter (ex.: *presque bleu, résultait qu', constaté que, etc.*).

**Corpus échantillon employé à l'INALF**: environ 400 000 occurrences, textes essentiellement littéraires mais mélange de genres (ouvrages sur la biologie et l'écologie).

Normes BRILL minimales: décolage des ponctuations, une phrase par ligne.

La version proposée par l'INALF inclut le logiciel FLEMM, qui permet d'obtenir le lemme des mots analysés.

### *Du jeu d'étiquettes disponible*

#### ***Pour le français, choix de l'INALF:***

Lexiques de l'INALF: 100 codes différents pour 5 catégories du discours: verbes conjugués (45), verbes non conjugués (3), adjectifs (4), substantifs (4), mots grammaticaux (45);

Étiquettes adoptées pour Brill: 49 (tableau 1, annexe).

#### ***En résumé, distinction:***

- du **nombre** des PD: sg. et pl.
- des **contractions** des déterminants: N pour non contracté, C pour contracté;
  - des **pronoms de genre indéterminé** – pour le *il* impersonnel, code ++;
  - des verbes **être** et **avoir** des "**autres verbes**" pour distinguer ces derniers des auxiliaires, dans la mesure où l'étiquetage s'effectue au niveau du mot.
- des **verbes conjugués** des **verbes non conjugués** (gérondif, participe présent, passé);
- des **participes passés après "être"** des **participes passés adjectivaux** (non après auxiliaire);
  - pour **QUE**, codage en tant que "subordonnant possible", code SUB\$ - trop ambigu;

étiquetage des **cardinaux, coordonnants, mots étrangers, interjections (onomatopées, etc.), préfixes détachés, prépositions, particules non indépendantes, relatifs (pronoms, adjectifs ou adverbes), subordonnants, symboles ou signes mathématiques.**

- étiquetage des signes de ponctuation (une quinzaine de codes).

**ADJECTIFS ADJ:xx** : trait **nombre**, sg. ou pl. pour adjectifs qualificatifs, ordinaux, possessifs et indéfinis (par opp. aux déterminants). Les *adjectifs adverbialisés* considérés comme des adjectifs et non comme des adverbes (il crachait noir);

**ADVERBES ADV** : concerne tous les adverbes, ainsi que les locutions adverbiales, reconnues comme telles dans une phase de pré-traitement et marquées avec un underscore (d'\_ailleurs). Les *adverbes prédéterminants* sont codés comme des adverbes (*presque/ADV* tous les mois), de même que les superlatifs et comparatifs. *Beaucoup* et *peu* sont codés en tant qu'adverbes possibles, mais pas *rien*.

**CARDINAUX CAR** : tous les numéraux cardinaux écrits en chiffres ou en lettres, ainsi que les dates. Pas de subdivision des cardinaux selon leur fonction. En outre, il n'était pas possible de distinguer l'indéfini *un* du cardinal.

**COORDONNANTS COO**: conjonctions de coordination usuelles mais aussi expressions comme *c'est-à-dire*, *non\_seulement*, *mais\_encore*, *soit...soit*, etc. Chaque élément de coordonnant discontinu reçoit un code de coordonnant normal.

**DETERMINANTS DTN:xx ou DTC:xx** : introducteurs de syntagmes nominaux. Distinction entre les déterminants normaux et les déterminants contractés (avec préposition ou partitif). Valeur de trait **nombre** sg. ou pl. *De* considéré comme ambigu (marqueur partitif ou préposition); *du* étiqueté **DTC:sg**, *des* **DTC:pl**. Problème pour *tout*, *toute*, *toutes*, *tous*, cf. p.11.

**INTERJECTIONS INJ**: surtout utilisées en discours direct, et à moins de correspondre au thème de recherche de l'article, il est peu probable d'en avoir dans les textes.

**PARTICIPES PASSES VPAR:xx ADJ1PAR:xx ou ADJ2PAR:xx:**

- Sont notés **VPAR:sg ou pl** (participes passés de tout verbe hormis *être* et *avoir*), **EPAR:sg** (participe passé du verbe *être* comme *j'ai été*) et **APAR:sg ou pl** (participe passé du verbe *avoir*) les participes passés placés en syntagme verbal derrière l'auxiliaire "avoir".
- Sont notés **ADJ1PAR:sg ou pl** les participes passés placés derrière un verbe "être" uniquement, de façon à pouvoir éventuellement récupérer la voix passive.
- Sont notés **ADJ2PAR:sg ou pl** les participes passés placés dans TOUS les autres contextes (nominaux ou ambigus ou derrière un verbe autre que "être" ou "avoir"). Ex.: il dormait *assis*, une fosse tout *équipée*, etc.

**PARTICULES PUL**: éléments n'ayant pas d'existence autonome et qui n'ont pas été regroupés dans une locution composée figée. Ex: *quant* (à, aux), éléments comme *-t-* dans une succession Verbe-pronom.

**PONCTUATIONS**: l'étiquette reprend le signe de ponctuation. Ex: **?/!** ou **///**. Le trait d'union reste à l'intérieur du mot, n'est étiqueté comme signe de ponctuation que si le mot composé n'est pas reconnu.

**PREFIXES PFX:** code les préfixes détachés, isolés entre deux blancs comme *entre, ex, micro...* Etiquette non productive au niveau des règles (le système n'affecte jamais cette étiquette de lui-même à un mot du texte).

**PREPOSITIONS PREP:** prépositions simples (*à, de, par, jusqu', etc.*). Pour les prépositions complexes, une particule PUL peut être associée à une préposition PREP. Les locutions prépositionnelles les plus fréquentes sont également codées. Les prépositions déictiques (*voici, voilà, revoici, revoilà*) ne reçoivent pas d'étiquette particulière.

**PRONOMS PRV:xx et PRO:xx:** concerne les éléments pronominaux à l'exclusion des relatifs (REL). Sous-catégorisation en Nombre (sg, pl ou ++ pour *en* et *y, s' et se*). Distinction des pronoms supportés ou non par le verbe: **PRV:xx** (*je, tu, il elle, nous, etc. + me, te se, leur, les, la, le, lui + en et y + -t-on, -t-il, etc.*) et **PRO:xx** pour tous les autres pronoms (démonstratifs, possessifs, anaphoriques, etc. comme *moi, ceci, sein, plusieurs, maints, etc.*).

**RELATIFS REL:** pronoms introduisant une subordonnée relative, complète ou non, avec antécédent exprimé.

**SUBORDONNANTS SUB et SUBS:** conjonctions de subordination, simples (*que, lorsque, quand, comme, etc.*) et locutions conjonctives (*parce que, afin que, etc.*). **Comme** et **que** sont considérés comme des cas particuliers car ambigus: *comme* peut être adverbe exclamatif (*comme il est beau!*) et *que* est très ambigu (relatif, subordonnant et adverbe exclamatif ou restrictif). **SUBS** concerne donc **que** et le code ainsi dans les cas d'ambiguïté maximale.

**SUBSTANTIFS SBC:xx et SBP:xx:** distinction des noms communs (SBC) et des noms propres (SBP). Sont codés noms communs les noms sans majuscules, avec majuscules pour les noms de peuples et d'habitants, et les adjectifs substantivés. Les noms propres ont été codés ainsi à partir du code \* de Frantext, de leur présence dans le lexique avec un code "nom propre", ainsi que de leur passage par le traitement des mots inconnus, qui affecte à tout mot inconnu commençant par une majuscule un code nom propre. Des ambiguïtés demeurent, liées aux problèmes d'homonymie et des mots composés.

**VERBES:** distinction entre les auxiliaires de temps et de voix et les verbes pleins. Les modaux ne sont pas distingués et sont considérés comme des verbes pleins.

pos 1	pos2 et 3	pos 4et5	Nbre gramm.
<b>A</b> (verbe <i>avoir</i> ) <b>E</b> (verbe <i>être</i> ) <b>V</b> (tout autre verbe)	<b>CJ</b> (forme conjuguée) <b>PAR</b> (forme participe passé)		<b>:sg</b> <b>:pl</b>
<b>A</b> (verbe <i>avoir</i> ) <b>E</b> (verbe <i>être</i> ) <b>V</b> (tout autre verbe)	<b>NC</b> (forme non conjuguée)	<b>FF</b> (infinitif) <b>NT</b> (gérondif ou participe présent)	

**Remarques:** CJ neutralise les indications de temps, de mode et de personne.

**Exemples:** ils *retombaient*/VCJ:pl ; elle l'*aura*/ACJ:sg vouluVPAR:sg; ils *sont*/ECJ:pl partis/ADJ1PAR:pl.; sans *vouloir*/VNCFF aller/VNCFF; n'*ayant*/ANCNT aucune idée de..., etc.

RESIDUS **ABR**, **FGW** ou **SYM**: concerne les mots difficiles à classer: abréviations (**ABR**), mots étrangers (**FGW**) et symboles (**SYM**).

**Pour l'anglais, 1/exploitation de 8 systèmes d'annotation différents unifiés par le projet AMALGAM (Automatic Mapping Among Lexico-Grammatical Annotation):**

(<http://www.scs.leeds.ac.uk/amalgam/amalgam/amalghome.htm>)

Les linguistes/groupes de recherche emploient des systèmes d'annotation grammaticale différents pour traiter leurs corpus, ce qui limite les possibilités d'échange et de partage des corpus.

Si la TEI est un moyen de pallier ce dernier problème, le projet AMALGAM a développé un ensemble d'algorithmes de mapping fusionnant les huit systèmes d'annotation suivants :

- Brown Corpus
- International Corpus of English
- London-Lund Corpus
- Lancaster-Oslo/Bergen Corpus
- UNIX parts
- Polytechnic of Wales Corpus
  - Spoken English Corpus
- University of Pennsylvania Corpus

A partir de ces systèmes d'annotation a été développé un système hybride entraîné avec Brill Tagger. Il est ainsi possible de tagger un texte, et de choisir les systèmes d'annotation souhaités. L'utilisateur envoie son texte par mail en spécifiant le(s) type(s) d'étiquetage qu'il souhaite.

Cette possibilité est très intéressante, dans la mesure où certains de ces systèmes, destinés à l'annotation de transcriptions orales, sont peu pertinents pour notre étude. De fait, le système d'annotation du London-Lund Corpus a été développé pour le codage de transcriptions de dialogues, le Polytechnic of Wales Corpus est fondé sur des transcriptions de dialogues interactifs avec des enfants et le Spoken English Corpus porte, comme son nom l'indique, sur de l'anglais parlé.

Le projet AMALGAM a d'ailleurs dû résoudre les problèmes générés par les spécificités des systèmes d'annotation, et le codage du LLC a par exemple dû être adapté au traitement de textes écrits.

**Pour l'anglais, 2/Utilisation du site <http://rayuela.lsi.uned.es/cgi-bin/ircourse/brill.perl>, qui utilise les étiquettes du Penn Treebank**



Source: Santorini, Beatrice. *Part-of-Speech Tagging Guidelines for the Penn Treebank Project*. March 15, 1991. [téléchargeable sur le site de TreeTagger]

Voici la liste des tags (sans les étiquettes concernant les marques de ponctuation):

**CC Conjonction de coordination:** inclut *and, but, nor, or, for* et *yet*, de même que les opérateurs mathématiques *plus, minus, less, times* (dans le sens de *multiplied by*) et *over* (dans le sens de *divided by*).

**CD Nombres cardinaux**

**DT Déterminants:** inclut les articles *a(n), every, no* et *the*, les déterminants indéfinis *another, any, some, each, either, neither, that, these, this* et *those*, de même que des occurrences de *all* et *both* lorsqu'ils ne précèdent pas un déterminant ou un pronom possessif.

**EX "there" existentiel:** le *there* qui entraîne une inversion du verbe et du sujet logique d'une phrase.

**FW Mots d'origine étrangère** (pas de liste particulière)

**IN Préposition ou conjonction de subordination:** Pas de distinction entre les deux – toujours possible de la retrouver néanmoins, car une préposition est un IN qui précède un SN ou un SP, et une conjonction de subordination un IN qui précède une proposition. A noter que la préposition *to* a son propre tag.

**JJ Adjectifs**

**JJR Adjectifs comparatifs:** adjectifs suffixés en *-er, more* et *less* dans certains cas.

**JJS Adjectifs superlatifs:** adjectifs suffixés en *-est, most* et *least* dans certains cas.

**LS marqueurs d'éléments dans une liste:** inclut les lettres et les numéraux lorsqu'ils sont employés dans une liste.

**MD Verbes modaux:** inclut tous les verbes qui ne prennent pas un *-s* à la troisième personne du singulier: *can, could, (dare), may, might, must, ought, shall, should, will* et *would*.

**NN Noms au singulier** (ou *mass nouns*).

**NNS Noms au pluriel.**

**NP Noms propres au singulier**

**NPS Noms propres au pluriel**

**PDT Prédéterminant:** inclut les éléments suivants lorsqu'ils précèdent un article ou un pronom possessif: *all, both, half, many, quite, rather, such, etc.*

**POS Marques de possession en position finale:** *'s* ou *'*: ces éléments sont écartés des mots auquel ils se rattachent comme s'ils étaient eux-mêmes des mots. Ex: John/NP *'s*/POS *idea*.

**PP Pronoms personnels**

**PP\$ Pronoms possessifs:** Inclut les pronoms possessifs *my, your, his, her, its, one's, our* et *their*. Les pronoms possessifs *mine, yours, his, hers, ours* et *theirs* sont étiquetés PP (pronoms personnels).

**RB Adverbes:** la plupart des mots inclus dans cette catégorie se terminent par *-ly*. Les adverbes de degré comme *quite, too* et *very* y sont également inclus, de même que les modificateurs comme *enough* ou *indeed* et les adverbes de négation comme *not, n't* ou *never*.

**RBR Adverbes comparatifs:** adverbes suffixés en *-er* qui n'ont pas un sens comparatif strict. Exemples : *later* dans *we can always come back later, later* sera étiqueté RB et non RBR.

**RBS Adverbes superlatifs**

**RP Particules:** inclut un ensemble de mots monosyllabiques pour la plupart susceptibles d'être redoublés en tant qu'adverbes directionnels ou prépositions. Un mot est une particule (RP) plutôt qu'une préposition :

- s'il peut aussi bien précéder ou suivre un complément d'objet. Exemple : *She told off/RP her friends/She told her friends off/RP*
- s'il fait partie d'un nom dérivé d'une collocation verbe-particule. Exemple : *to break down* → *breakdown* ; *to be left over* → *leftovers*, etc.

- enfin, s'il accepte un accent en fin de proposition. Exemple : *Why don't you come **by**/RP (vs. Real bargains are hard to come **by**/IN)?*

**SYM Symboles:** symboles ou expressions mathématiques, scientifiques ou techniques qui ne sont pas des mots anglais. SYM ne doit toutefois pas être utilisé pour tagger toutes les expressions techniques. Par exemple, les noms de symboles et d'unités de mesures (abréviations incluses) devront être étiquetés en tant que noms.

**TO "To"**

**UH Interjection** inclut *my, oh, please, see, uh, well, yes, etc.*

**VB Verbes, forme "de base":** cette catégorie inclut les impératifs, les infinitifs et les subjonctifs.

**VBD Verbes au passé:** inclut la forme conditionnelle du verbe *to be* (If I were/VBD rich...)

**VBG Verbes au gérondif ou au participe présent**

**VBN Verbes au participe passé**

**VBP Verbes au présent, formes autres que la troisième personne du singulier.**

**VBZ Verbes au présent, troisième personne du singulier**

**WDT Wh-determiner:** inclut *which* et *that* (employé en tant que pronom relatif).

**WP Wh-pronoms:** inclut *what, who* et *whom*.

**WP\$ Pronom possessif en wh-:** inclut *whose*.

**WRB Wh-adverbes:** inclut *how, where, why, etc.* *when* est étiqueté WRB dans son acception temporelle. Dans le sens de IF, il sera étiqueté IN (ex.: *when*/WRB *he finally arrived*/I like it *when*/IN *you make dinner for me*).

### ***Du codage des tags et de leur conversion en balises XML***

La conversion des tags Brill en balises XML est simple, tous les tags étant suffixés au mot et commençant par un slash /.

*Exemple:*

It/PRP is/VBZ generally/RB the/DT case/NN that/IN special/JJ volumes/NNS include/VBP studies/NNS on/IN a/DT narrowly/RB defined/VBN topic/NN of/IN linguistic/JJ investigation/NN from/IN a/DT single/JJ theoretical/JJ standpoint.

## **CORDIAL ANALYSEUR**

### *Fonctionnement général de l'étiqueteur*

Cordial 8 (Société Synapse-Développement) est un correcteur orthographique et grammatical du français. La version "Cordial Analyseur", qu'il est possible d'acquérir à des fins de recherche, intègre entre autres fonctionnalités un étiqueteur morphosyntaxique. Les résultats de l'étiquetage (texte étiqueté, chiffres absolus, pourcentages, etc.) sont distribués dans différents fichiers en format texte, facilement réexploitables dans d'autres cadres d'analyse (méthodes de statistique multidimensionnelle exploratoire -analyse factorielle, classification automatique, etc.).

Il est en outre possible de configurer l'étiquetage et de sélectionner les informations souhaitées parmi les 27 types de données proposées (lemme, type syntaxique, numéro de paragraphe, équivalents sémantiques, etc.).

Cordial a été largement exploité dans des cadres d'analyse et d'exploration de corpus (Malrieu & Rastier 2000, Beauvisage 2000, etc.) et est également utilisé en amont de traitements automatiques par différents outils (TypTex, Lexter, etc.).

Etant donné son statut de logiciel grand public, les variables qu'il propose sont plus grammaticales (grammaire scolaire traditionnelle) que linguistiques – a fortiori, aucune documentation renseignant les catégories utilisées n'est disponible.

## MBT TAGGER

Sources:

<http://ilk.uvt.nl/cgi-bin/tstchunk/demo.pl>

- Daelemans, Walter et al. *MBT: Memory-Based Tagger, version 1.0, Reference Guide*. ILK Technical Report –ILK 02-09. October, 29, 2002.
- Zavrel, Jakub et Daelemans, Walter. *Recent Advances in Memory-Based Part-of-Speech Tagging*. [retrouver référence précise] 1999
- Daelemans, Walter and Zavrel, Jakub. *Part-of-Speech Tagging of Dutch with MBT, a Memory-Based Tagger Generator*. 1996 [retrouver référence précise].

### Fonctionnement général de l'étiqueteur

L'étiqueteur MBT est fondé sur un système d'**apprentissage** combinant deux méthodes d'étiquetages largement utilisées: l'étiquetage stochastique, et l'étiquetage par règles (cf. Brill tagger).

À l'instar de Brill, MBT peut s'appliquer à tout type de corpus annoté. MBT fonctionne à partir de **techniques d'apprentissage et de classification**:

Un ensemble d'**exemples** est stocké dans une mémoire; chaque exemple contient un mot (ou sa représentation lexicale), son contexte (antérieur et postérieur) et la catégorie grammaticale à laquelle il est associé dans chaque contexte.

Une nouvelle phrase sera donc analysée de la manière suivante: pour chaque mot de la phrase, le tagger cherchera un exemple d'emploi analogue dans la mémoire et en déduira sa catégorie grammaticale – à partir de ses plus proches voisins.

Les tags sont donc considérés comme des variables, qui seront assignées aux mots à partir de **méthodes de classification**. MBT emploie une mesure de similarité qui considère le nombre de tags potentiels qu'il est possible d'associer à chaque mot et qui pondère l'importance de chaque catégorie.

MBT est un **générateur de tagger**: à partir d'un corpus étiqueté selon le système d'annotation désiré, un POS tagger est généré. **Trois structures de données** sont extraites automatiquement du corpus fourni en entrée: un **lexique** (permettant d'associer des tags aux mots) et **deux bases de cas** (*casebases*) pour les mots connus et inconnus.

- Construction du **lexique**: la définition lexicale d'un mot correspond à la fréquence des tags auxquels il est associé dans le corpus de départ. Par exemple, la définition du mot *once* pourra être: *RB:330;IN:77*, i.e. *once* a été étiqueté 330 fois en tant qu'adverbe et 77 fois comme une préposition/conjonction de subordination.
- Construction des **bases de cas**: le système mémorise les informations qui sont associées aux *mots connus* (présents dans le corpus de départ), c'est-à-dire leurs contextes gauche et droits et les étiquettes correspondantes. Les étiquettes correspondant aux *mots inconnus* ne peuvent être devinées qu'à partir de leur forme ou de leur contexte. Des informations d'ordre morphologique sont stockées par le système: pour les suffixes sont conservées les trois dernières lettres du mot et pour les préfixes la première. Ces informations sont certes simplistes, mais le système se veut indépendant des langues et des systèmes d'annotation...

## Du jeu d'étiquettes disponible

La version de MBT tagger que nous utilisons en accord avec W. Daelemans n'a été développée que pour l'anglais. Son utilisation en français nécessiterait donc la construction d'un corpus d'entraînement français de taille importante – et donc son étiquetage manuel.

Deux types d'étiquettes sont proposées: des tags de type morphosyntaxique et de type syntaxique. MBT a en effet été entraîné sur le corpus d'arbres syntaxiques développé par le projet Penn Treebank. L'annotation syntaxique du corpus n'a toutefois pas été entièrement exploitée.

### POS Tags

En anglais, les tags sont ceux du Penn Treebank. Cf. Liste seconde possibilité pour traiter l'anglais de Brill tagger.

Un double slash entre le mot et la catégorie signifie que le mot est inconnu à l'étiqueteur.

### Chunks

A l'étiquetage des catégories morphosyntaxiques développées par le projet Penn Treebank s'ajoute celui d'un certain type de syntagmes (les *chunks*), i.e. des syntagmes ne se chevauchant pas.

Type de syntagme	Description	Code	Exemple
Syntagme nominal	Etiquetage des SN du début du syntagme au nom tête, mais exclusion des compléments suivant ce dernier.	[NP...NP]	[P In/IN P] [NP addition/NN NP] PNP} {PNP [P to/TO P]
Préposition	Ne code que la préposition en question, de même que certaines locutions prépositionnelles	[P...P]	[P In/IN P] [NP addition/NN NP] PNP} {PNP [P to/TO P]...
Syntagme prépositionnel	Code une préposition et un ou plusieurs NPs	{PNP...PNP}	{PNP [P on/IN P] [NP the/DT other/JJ hand/NN NP] PNP}
Syntagme adjectival	Etiquetage des SA du début du syntagme au nom tête, mais exclusion des compléments suivant ce dernier. Pas d'étiquetage des SA à l'intérieur d'un SN.	[ADJP...ADJP]	[ADJP more/RBR reliable/JJ ADJP]
Syntagme adverbial	Etiquetage des SAdv du début du syntagme au nom tête, mais exclusion des compléments suivant ce dernier. Pas d'étiquetage des SAdv à l'intérieur d'un SV ou d'un Sadj..	[ADVP...ADVP]	[ADVP when/WRB ADVP]
Syntagme verbal	Contient un verbe principal, ses éventuels auxiliaires ou modaux, de même que les compléments adverbiaux ou verbaux qui le suivent.	[VP...VP]	[VP-3 is/VBZ considered/VBN VP-3]
Introduceur de complétives	A l'instar de la préposition, ne code le plus souvent qu'un mot – l'introduceur de proposition complétive.	[C...C]	[VP-1 suggests/VBZ VP-1] [C that/IN C]

### Annotation de sujets et d'objets

Les syntagmes nominaux sujet et objet (*chunks*, et non SN) sont respectivement étiquetés **[NP-SBJ-i... NP-SBJ-i]** et **[NP-OBJ-i... NP-OBJ-i]**. "i" est un nombre entier indiquant à quel SV – indexé par le même nombre - le SN se rapporte.

Les sujets/objets n'étant pas des SN ne sont pas annotés.

*Exemple:*

[C Because/IN C] **[NP-SBJ-1 the/DT ability/NN NP-SBJ-1]** **[VP-1 to/TO quantify/VB VP-1]** **[NP-OBJ-1 the/DT spatiotemporal//JJ phenomena/NNS NP-OBJ-1]**

Les objets "prédicatifs" (en attributs) sont étiquetés **NP-PRD**

*Exemples:*

[NP-SBJ-1 Nonempirical//JJ studies/NNS NP-SBJ-1] {PNP [P of/IN P] [NP language/NN NP] PNP} [VP-1 have/VBP been/VBN VP-1] **[NP-1NP-PRD extremely/RB successful/JJ NP-1NP-PRD]**

[ADJP available/JJ ADJP] [VP-4 to/TO be/VB VP-4] **[NP-4NP-PRD potentially/RB known/VBN or/CC utilized/VBN NP-4NP-PRD]** ./.

Un SN inclut dans une expression figée est étiqueté **NP-[iNP]-CLR**.

*Exemple:*

[VP-4 take/VBP VP-4] **[NP-4NP-CLR place/NN NP-4NP-CLR]**

### ***Du codage des tags et de leur conversion en balises XML***

La conversion des tags morphosyntaxiques est simple, les tags étant suffixés au mot par un slash (à l'instar de l'annotation proposée par Brill).

Une réflexion sur la récupération de l'annotation syntaxique sera à mener, mais les séparateurs employés sont très identifiables: **{**, **}**, **[** et **]**.

## TNT TAGGER

Sources:

<http://www.coli.uni-saarland.de/~thorsten/tnt/>

- Corpus SUSANNE: <http://www.grsampson.net/RSue.html>

<http://www.grsampson.net/SueDoc.html>

### *Fonctionnement général de l'étiqueteur*

TnT (Trigrams'n'Tags) est un tagger statistique qu'il est – théoriquement – possible d'entraîner sur des langues et des systèmes d'étiquetage différents.

Implémenté sur un algorithme fondé sur les modèles de Markov, l'étiqueteur est constitué d'un ensemble de méthodes de **smoothing** et de traitement des mots inconnus.

Tel qu'il est disponible par convention, TnT ne permet de traiter que l'anglais et l'allemand. Il a en effet été entraîné pour l'allemand sur un corpus journalistique annoté selon les conventions *Stuttgart-Tübingen*, et pour l'anglais sur les corpus SUSANNE<sup>4</sup> (Surface and underlying structural analysis of natural English) et Penn Treebank.

Les concepteurs de TnT ont veillé à optimiser la vitesse de l'étiqueteur; celle-ci dépend bien sûr de l'ambiguïté et du pourcentage de mots inconnus du texte, mais le système est capable d'étiqueter entre 30 000 et 60 000 mots par seconde sur un Pentium 500 sous Linux.

Il est possible d'entraîner le tagger à étiqueter d'autres langues, à condition qu'elles utilisent l'espace comme séparateur de mots et qu'elles puissent être représentées en ASCII. Le corpus d'entraînement doit répondre au format suivant: un élément par ligne, la première colonne correspondant au mot et la seconde au tag.

### *Du jeu d'étiquettes disponible*

La version de TnT telle qu'elle est disponible est fondée sur le système d'annotation SUSANNE (Geoffrey Sampson). Le tagger a en effet été entraîné sur le corpus SUSANNE, sous-corpus du corpus Brown – anglais américain écrit – de 128 000 mots. De taille restreinte (tout spécialement par rapport au corpus développé par le Penn Treebank), Susanne est composé de 4 "types de textes" différents, selon la taxonomie adoptée par le corpus Brown: reportages de presse, belles lettres-biographies-mémoires, écrits "savants" (textes essentiellement techniques et scientifiques) et fiction (aventures et westerns).

Le système d'encodage adopté est fondé sur celui de Lancaster (Garside et al. 1987 appendix B) et comprend 353 codes différents.

TnT reprend deux niveaux d'encodage de SUSANNE: un niveau morphosyntaxique et un niveau supérieur, qui correspondrait globalement à celui des "chunks" (*phrasegroups*) et qui est composé de huit types de segments:

1. **V** verb group
2. **N** noun phrase
3. **J** adjective phrase
4. **R** adverb phrase

---

<sup>4</sup> Extrait du corpus Brown, SUSANNE est un corpus annoté d'anglais américain de 130 000 mots.

5. **P** prepositional phrase
6. **D** determiner phrase
7. **M** numeral phrase
8. **G** genitive phrase

A ces segments sont associés des informations concernant la nature morphosyntaxique des mots:

- **21 catégories pour les verbes:** **Vo** operator section of verb group, when separated from remainder of v e.g. by subject-auxiliary inversion; **Vr** remainder of v from which **Vo** has been separated; **Vm** v beginning with *am*; **Va** v beginning with *are*; **Vs** v beginning with *was*; **Vz** v beginning with other 3rd-singular verb; **Vw** v beginning with *were*; **Vj** v beginning with *be*; **Vd** v beginning with past tense; **Vi** infinitival v; **Vg** v beginning with present participle; **Vn** v beginning with past participle; **Vc** v beginning with modal; **Vk** v containing emphatic *DO*; **Ve** negative v; **Vf** perfective v; **Vu** progressive v; **Vp** passive v; **Vb** v ending with *BE*; **Vx** v lacking main verb; **Vt** catenative v

- **12 catégories pour les noms:** **Nq** *wh-* N; **Nv** *wh...ever* N; **Ne** *I/me* as whole or head; **Ny** *you* as whole or head; **Ni** *it* as whole or head; **Nj** adjectival head; **Nn** proper name; **Nu** unit of measurement as head; **Na** marked as subject; **No** marked as nonsubject; **Ns** marked as singular; **Np** marked as plural

- **5 catégories pour les adjectifs:** **Jq** *wh-* J; **Jv** *wh...ever* J; **Jx** measured absolute J; **Jr** measured comparative J; **Jh** "heavy" (postmodified) J;

- **6 catégories pour les adverbes:** **Rq** *wh-* R; **Rv** *wh...ever* R; **Rx** measured absolute R; **Rr** measured comparative R; **Rs** adverb conducive to asyndeton; **Rw** quasi-nominal adverb

- **4 catégories pour les prépositions:** **Po** *of* phrase; **Pb** *by* phrase; **Pq** *wh-* P; **Pv** *wh...ever* P

- **4 catégories pour les déterminants:** **Dq** *wh-* D; **Dv** *wh...ever* D; **Ds** marked as singular; **Dp** marked as plural

**1 catégorie pour les numéraux:** **Ms** *M* headed by *one*

- **1 catégorie pour les génitifs:** **Gq** *wh-* G *Gv wh...ever* G

### *Du codage des tags et de leur conversion en balises XML*

Le fichier résultat contient un élément par ligne. En "mode basique", le tagger ajoute une seconde colonne contenant le tag associé au mot à chaque ligne. Différentes options sont disponibles: il est par exemple possible de rajouter des tags alternatifs pour chaque mot.

Input	Basic		Optional			
	Output		Extended	Output		
Der	ART		ART	1.000000e+00		
Mandolinen-Club	NN	*	NN	1.000000e+00	*	
Falkenstein	NE	*	NE	8.001280e-01	NN	1.998720e-01 *
und	KON		KON	1.000000e+00		
der	ART		ART	1.000000e+00		
Frauenchor	NN	*	NN	9.828203e-01	NE	1.717975e-02 *



aus	APPR	APPR	1.000000e+00		
dem	ART	ART	1.000000e+00		
sächsischen	ADJA	ADJA	1.000000e+00		
Königstein	NN	NN	7.762892e-01	NE	2.237108e-01
gestalten	VVINF	VVINF	1.000000e+00		
die	ART	ART	9.796126e-01	PRELS	1.443545e-02
PDS 5.951974e-03					
Feier	NN	NN	1.000000e+00		
gemeinsam	ADJD	ADJD	1.000000e+00		
.	\$.	\$.	1.000000e+00		

-----+-----+-----

Les mots marqués par un astérisque (\*) ne sont pas inclus dans le lexique originel de l'étiqueteur; c'est à partir d'une analyse de leurs suffixes qu'ils ont été étiquetés.

## TREETAGGER

Sources:

<http://www.ims.uni-stuttgart.de/projekte/corplex/TreeTagger/DecisionTreeTagger.html>

- Schmid, Helmut. *Probabilistic Part-Of-Speech Tagging Using Decision Trees*. International Conference on New Methods in Language Processing, 1994, Manchester, UK. [revised version]
- Schmid, Helmut. *Improvements in Part-Of-Speech Tagging With an Application to German*. EACL SIGDAT workshop, 1995, Dublin. [revised version]

### *Fonctionnement général de l'étiqueteur*

Développé au sein du projet TC (Textcorpora and Erschliessungswerkzeuge/'textual corpora and tools for their exploration', <http://www.ims.uni-stuttgart.de/projekte/tc/>) à l'institut de linguistique computationnelle de l'université de Stuttgart, TreeTagger est un système d'annotation de catégories morphosyntaxiques permettant d'étiqueter des textes en anglais, français, allemand, italien, grec, et ancien français. Il est possible d'adapter l'étiqueteur à d'autres langues, à condition de disposer d'un lexique et d'un corpus manuellement annoté.

TreeTagger est proche des taggers *ngram* traditionnels. Les deux systèmes modèlent la probabilité de l'annotation d'une séquence de mots.

Cependant, contrairement à la plupart des étiqueteurs qui recourent aux modèles de Markov pour résoudre le problème des "données clairsemées" (*sparse data*), TreeTagger utilise un **arbre de décision** binaire pour calculer la taille du contexte à utiliser pour estimer les probabilités de transition.

Le **lexique** implémenté dans TreeTagger contient la liste des possibilités d'étiquetage pour chaque mot. Il se divise en trois parties:

- un **lexique pleine forme**, contenant les deux millions d'entrées du corpus Penn Treebank. Les mots dont la fréquence relative était inférieure à 1% ont été supprimés: ils étaient le plus souvent dus à des erreurs d'étiquetage;
  - un **lexique de suffixes**, organisé dans une structure arborescente. Chaque nœud de l'arbre, à l'exception du nœud racine, correspond à un caractère. Les feuilles de l'arbre contiennent des vecteurs de probabilité des tags;
- et une **entrée par défaut**, obtenue en soustrayant aux fréquences du nœud racine les fréquences des tags dans les feuilles de l'arbre, et en normalisant les fréquences résultantes.

La recherche d'un mot dans le lexique démarre par une recherche dans le premier fichier (avec changement de la casse du mot si la recherche s'avère infructueuse avec la casse originelle); puis dans le second si le mot n'a pas été trouvé dans le premier.

### *Du jeu d'étiquettes disponible*

#### **Pour l'anglais**

Étiquettes développées par le projet Penn Treebank. Cf. liste, donnée dans le descriptif des étiquettes de Brill.

#### **Pour le français**

**ABR** abréviations  
**ADJ** adjectifs  
**ADV** adverbes  
**DET:ART** articles  
**DET:POS** pronoms possessifs (ma, ta, ...)  
**INT** interjections  
**CON** conjonctions  
**NAM** noms propres  
**NOM** noms  
**NUM** numéraux  
**PRO** pronoms  
**PRO:DEM** pronoms démonstratifs  
**PRO:IND** pronoms indéfinis  
**PRO:PER** pronoms personnels  
**PRO:POS** pronoms possessifs (mien, tien, ...)  
**PRO:REL** pronoms relatifs  
**PRP** prépositions  
**PRP:det** déterminants contractés (au, du, aux, des)  
**PUN** marques de ponctuation  
**PUN:cit** marques de ponctuation marquant les citations  
**SENT** phrases  
**SYM** symboles  
**VER:cond** verbes au conditionnel  
**VER:futu** verbes au futur  
**VER:impe** verbes à l'impératif  
**VER:impf** verbes à l'imparfait  
**VER:infi** verbes à l'infinitif  
**VER:pper** participes passés  
**VER:ppre** participes présents  
**VER:pres** verbes au présent  
**VER:simp** verbes au passé simple  
**VER:subi** verbes au subjonctif imparfait  
**VER:subp** verbes au subjonctif présent

### *Du codage des tags et de leur conversion en balises XML*

Le texte est découpé en mots ou expressions (un mot par ligne) et le fichier résultat est organisé en quatre colonnes: la première colonne reporte le mot étiqueté, la seconde son tag, la troisième sa forme en minuscules et la quatrième (souvent vide) propose une analyse du mot – on y trouve le plus souvent le lemme du mot.

## VISL

Sources:

- Bick, Eckhard. "Danish Symbol Set (CG). Symboles used in the Danish Constraint Grammar and in "free input" mode". <http://visl.sdu.dk/visl/da/info/dansymbolcg.html>

### *Fonctionnement général de l'étiqueteur*

La plupart des modules morphologiques, et tous les modules syntaxiques de VISL sont fondés sur le paradigme des *grammaires de contraintes* (*constraint grammar*, Karlsson et al., 1995).

Fondé sur des représentations en arbres de dépendances, VISL propose une analyse syntaxique de surface (*shallow surface syntactic analysis*). Tous les mots du texte reçoivent une ou plusieurs étiquettes morphosyntaxiques, ainsi qu'un tag indiquant leur relation de dépendance (marquée par le symbole @).

VISL propose différents types de représentations: représentation arborescente horizontale ou verticale, couleurs, représentation plane, etc. L'analyse des textes peut s'effectuer en ligne dans la mesure où une interface d'*uploading* est proposée (<http://beta.visl.sdu.dk/visl/fr/parsing/automatic/upload.php>).

En ce qui nous concerne, et relativement au traitement ultérieur des sorties de l'analyseur, la représentation plane nous semble être la plus adaptée.

### *Du jeu d'étiquettes disponible*

Les étiquettes, de même que les relations de dépendances sont en majuscules. Les étiquettes syntaxiques sont préfixées d'une @ et de flèches de dépendances pointant vers la tête du syntagme (ex.: @SUBJ> pour un sujet à la gauche du prédicat).

VISL propose différents niveaux d'analyse (de la phrase au mot). Au niveau du mot, on retrouve les cinq catégories du discours traditionnelles: nom, adjectif, pronom, verbe et adverbe. Les informations auxquelles elles peuvent être associées ont été synthétisées dans un tableau par Bick (Bick, p.10):

		gender	number	case	degree	person	tense	mode
		M, F, nG	S, P, nN	NOM, DAT, ACC, PIV	COM (POS)	1, 2, 3	PR, IMPF, PS, FUT	IND, SUBJ, COND, IMP
noun	N	+	+					
proper noun	PROP	(+)						
adjective	ADJ	+	+		+	(few)		
pronoun	PRON	+	+	+		+		
personal	PERS							
determiner	PRON DET	+	+					
independent	PRON INDP	+	+					
verb	VFIN		+			+	+	+
finite								
infinitive	INF							

past participle	PCP2	+	+					
present participle	PCP1	+	+					
<b>adverb</b>	ADV				+ (few)			

Une réflexion concernant l'intégration des niveaux d'analyses supérieurs au mot devra être menée; la tâche s'avère difficile, d'autant que les modules VISL sont fondés sur les grammaires de dépendances, ce qui les rend difficiles à intégrer aux modèles fondés sur des représentations en constituants immédiats.

Mentionnons par exemple qu'au niveau de la phrase, VISL propose cinq fonctions de base: un constituant verbal principal (le *predicator* **P**) et quatre constituants satellites non verbaux: le sujet (**S**), l'objet (**O**), l'adverbe (**A**) et le complément (**C**).

### ***Du codage des tags et de leur conversion en balises XML***

Plus complexe car présence de balises syntaxiques. Eventuellement, ne récupérer que les étiquettes morphosyntaxiques.

Tagger	Langues traitées	Théorie linguistique/particularité de l'outil	Corpus d'entraînement	Type (s) d'étiquettes	Nombre de tags
<b>Brill Tagger</b> (Eric Brill)	Français (InaLF)	Travaux structuralistes de Bloomfield et Harris; Inférence de règles d'étiquetage; Analyse distributionnelle pour réduire les erreurs d'étiquetage;	Base Frantext pour le français (InaLF)	Morpho-syntaxiques	17 catégories de mots 49 codes
	Anglais (en ligne)		Corpus Penn Treebank pour l'anglais		36
<b>Cordial Analyseur</b> (Synapse Devt, France)	Français	Grammaire scolaire	X	Morpho-syntaxiques	Plus de 500
<b>MBT Tagger</b> (ILK, W. Daelemans)	Anglais	Techniques d'apprentissage et de classification; Apprentissage à partir d'exemples;	Corpus Penn Treebank	Morpho-syntaxiques et syntaxiques	POS tags (36); Chunks (7); Sujets et objets
<b>TnT Tagger</b> (Univ. Saarlandes, DE, Thorsten Brants)	Anglais	Tagger statistique;	Corpus SUSANNE (BROWN, anglais américain, 130 000 mots) Corpus Penn Treebank	Morpho-syntaxiques	
<b>Tree-Tagger</b> (TC Project, Helmut Schmid)	Anglais (allemand, italien, grec, ancien français)	Utilisation d'arbres de décision binaires; Arbres de suffixes;	Corpus Penn Treebank	Morpho-syntaxiques	36
	Français		?		33
<b>VISL</b> (Syddansk University, Eckhard Bick)	Anglais, français et un grand nombre d'autres langues	Grammaires de contraintes, arbres de dépendances;	?	Morpho-syntaxiques et syntaxiques (fonctionnelles)	5 POS, 11 types, 1 à 7 traits + infos syntaxe
<b>XELDA</b> (Xerox, Grenoble)	Anglais (+ autres langues)	Finite state technology	X	Morpho-syntaxiques	E: 71
	français				F: 45

Tableau synthétisant les caractéristiques externes des logiciels

## Annexe 9 : Systèmes d'annotation

1. Système développé par l'INALF pour Brill à partir de la base Frantext (Lecomte et Paroubek, 1996) ;
2. Système d'étiquetage français de TreeTagger (<http://www.ims.uni-stuttgart.de/~schmid/french-tagset.html>);
3. Annotation développée par le projet PennTreebank (Santorini, 1991) ;
4. Système SUSANNE (Sampson, 1995).

POS ou catégorie	Code	Description
ADJECTIFS	ADJ:sg	Adjectifs singulier
ADJECTIFS	ADJ:pl	Adjectifs pluriel
ADVERBES	ADV	Adverbes et locutions adverbiales (marquées avec _)
CARDINAUX	CAR	Numéraux cardinaux en chiffres/lettres + dates
COORDONNANTS	COO	Conjonctions de coordination + expressions ( <i>soit...soit</i> , etc.)
DETERMINANTS	DTN:sg	Déterminants singulier
DETERMINANTS	DTN:pl	Déterminants pluriel
DETERMINANTS	DTC:sg	Déterminant contracté DU
DETERMINANTS	DTC:pl	Déterminant contracté DES
INTERJECTIONS	INJ	Interjections
PARTICIPES PASSES	VPAR:sg	PP singulier de tous verbes hormis ETRE et AVOIR
PARTICIPES PASSES	VPAR:pl	PP pluriel de tous verbes hormis ETRE et AVOIR
PARTICIPES PASSES	EPAR:sg	PP du verbe ETRE (j'ai été)
PARTICIPES PASSES	APAR:sg	PP du verbe AVOIR au singulier
PARTICIPES PASSES	APAR:pl	PP du verbe AVOIR au pluriel
PARTICIPES PASSES	ADJ2PAR:sg	PP autres contextes (nominaux, ambigus, autres V) singulier
PARTICIPES PASSES	ADJ2PAR:pl	PP autres contextes (nominaux, ambigus, autres V) pluriel
PARTICULES	PUL	Eléments non autonomes/non regroupés ds expressions figées (-t-, <i>quant</i> (à, aux), etc.)
PONCTUATIONS	[signe] / [signe]	Ex.: <i>?!/, ///</i>
PREFIXES	PFX	Préfixes détachés, isolés entre 2 blancs comme <i>entre</i> , <i>ex</i> , <i>micro</i>
PREPOSITIONS	PREP	Prépositions simples (PUL + PREP prépositions complexes)
PRONOMS	PRV:sg	Eléments pronominaux (non relatifs) supportés par le V au singulier. Ex.: JE, ME, TE, SE, LA, LE, LUI, T-IL, T-ON, etc.
PRONOMS	PRV:pl	Idem au pluriel. Ex.: NOUS, VOUS, LEUR, etc.
PRONOMS	PRV:++	Pronoms EN, Y, S' et SE
PRONOMS	PRO:sg	Eléments pronominaux (non relatifs) non supportés par le V au singulier. Ex.: MOI, CECI, SIEN, etc.
PRONOMS	PRO:pl	Idem au pluriel. Ex.: MAINS, PLUSIEURS, etc.
RELATIFS	REL	Pronoms relatifs
SUBORDONNANTS	SUB	Conjonctions de subordination simples ( <i>quand</i> , <i>comme</i> , <i>lorsque</i> , etc.) et locutions ( <i>parce que</i> , <i>afin que</i> , etc.)
SUBORDONNANTS	SUB\$	Subordonnant QUE (codé ainsi dans les cas d'ambiguïté max.)
SUBSTANTIFS	SBC:sg	Noms communs au singulier
SUBSTANTIFS	SBC:pl	Noms communs au pluriel
SUBSTANTIFS	SBP:sg	Noms propres au singulier

<b>SUBSTANTIFS</b>	<b>SBP:pl</b>	Noms propres au pluriel
<b>VERBES – préfixe</b>	<b>A, E ou V</b>	Distinction auxiliaire AVOIR, ETRE ou AUTRE VERBE
<b>VERBES</b>	<b>CJ:sg</b>	Formes conjuguées au singulier
<b>VERBES</b>	<b>CJ:pl</b>	Formes conjuguées au pluriel
<b>VERBES</b>	<b>PAR:sg</b>	Participes passés au singulier
<b>VERBES</b>	<b>PAR:pl</b>	Participes passés au pluriel
<b>VERBES</b>	<b>NCFF</b>	Formes non conjuguées (NC) à l'infinitif (FF)
<b>VERBES</b>	<b>NCNT</b>	Formes non conjuguées (NC) au gérondif ou au participe présent (NT)
<b>RESIDUS</b>	<b>ABR</b>	Abréviations
<b>RESIDUS</b>	<b>FGW</b>	Mots étrangers
<b>RESIDUS</b>	<b>SYM</b>	Symboles

Tableau 1 : [Français] tableau synthétisant les codes adoptés par l'Inalf pour entraîner Brill à partir de la base Frantext (43 codes + codes préfixes verbes)

POS ou catégorie	Code	Description
<b>ADJECTIFS</b>	<b>ADJ</b>	Adjectifs
<b>ADJECTIFS</b>	<b>ADJ:num</b>	Adjectifs numéraux
<b>ADVERBES</b>	<b>ADV</b>	Adverbes
<b>CONJONCTIONS</b>	<b>CON:coo</b>	Conjonctions de coordination
<b>CONJONCTIONS</b>	<b>CON:sub</b>	Conjonctions de subordination
<b>DETERMINANTS</b>	<b>DET:def</b>	Déterminants définis
<b>DETERMINANTS</b>	<b>DET:indef</b>	Déterminants indéfinis
<b>INTERJECTIONS</b>	<b>INT</b>	Interjections
<b>PONCTUATIONS</b>	<b>PON</b>	Marques de ponctuation hors points et virgules
<b>PONCTUATIONS</b>	<b>PON:sep</b>	Points
<b>PONCTUATIONS</b>	<b>PON:comma</b>	Virgules
<b>PREPOSITIONS</b>	<b>PRE</b>	Prépositions
<b>PREPOSITIONS/ DETERMINANTS</b>	<b>PRE:det</b>	Déterminants contractés (au, du, aux, des)
<b>PRONOMS</b>	<b>PRO:demo:attr</b>	Adjectifs démonstratifs CE, CET, etc.
<b>PRONOMS</b>	<b>PRO:demo:pred</b>	Pronoms démonstratifs CELUI, CELLE, CELUI-CI, etc.
<b>PRONOMS</b>	<b>PRO:indef:attr</b>	Adjectifs indéfinis TOUT, PLUSIEURS, CERTAIN, etc.
<b>PRONOMS</b>	<b>PRO:indef:pred</b>	Pronoms indéfinis CHACUN, QUICONQUE, etc.
<b>PRONOMS</b>	<b>PRO:pers:clit</b>	Pronoms personnels clitiques le, la...
<b>PRONOMS</b>	<b>PRO:pers:conj</b>	Pronoms personnels conjugués JE, TU, IL, ON, etc.
<b>PRONOMS</b>	<b>PRO:poss</b>	Adjectifs possessifs SON, SA, SES, etc.
<b>PRONOMS</b>	<b>PRO:rela</b>	Pronoms relatifs
<b>SUBSTANTIFS</b>	<b>NOM</b>	Substantifs
<b>VERBES</b>	<b>VER:cond</b>	Verbes au conditionnel
<b>VERBES</b>	<b>VER:futu</b>	Verbes au futur
<b>VERBES</b>	<b>VER:impe</b>	Verbes à l'impératif (non remarqué dans corpus)
<b>VERBES</b>	<b>VER:impf</b>	Verbes à l'imparfait
<b>VERBES</b>	<b>VER:infi</b>	Verbes à l'infinitif
<b>VERBES</b>	<b>VER:pper</b>	Participes passés
<b>VERBES</b>	<b>VER:ppe</b>	Participes présents
<b>VERBES</b>	<b>VER:pres</b>	Verbes au présent
<b>VERBES</b>	<b>VER:simp</b>	Verbes au passé simple
<b>VERBES</b>	<b>VER:subi</b>	Verbes au subjonctif imparfait
<b>VERBES</b>	<b>VER:subp</b>	Verbes au subjonctif présent
<b>RESIDUS</b>	<b>ABR</b>	Abréviations



RESIDUS	SYM	Symboles
---------	-----	----------

Tableau 2 : [français] tableau synthétisant les codes utilisés par TreeTagger pour le français

POS ou catégorie	Code	Description
ADJECTIFS	JJ	Adjectifs
ADJECTIFS	JJR	Adjectifs comparatifs (suffixés en <i>-er</i> , <i>more</i> et <i>less</i> dans certains cas)
ADJECTIFS	JJS	Adjectifs superlatifs suffixés en <i>-est</i> , <i>most</i> et <i>least</i> dans certains cas
ADVERBES	RB	Adverbes (principalement mots se terminent par <i>-ly</i> ) + adverbes de degré (e.g. <i>quite</i> , <i>too</i> et <i>very</i> ) + modificateurs (e.g. <i>enough</i> ou <i>indeed</i> ) + adverbes de négation (e.g. <i>not</i> , <i>n't</i> ou <i>never</i> ).
ADVERBES	RBR	Adverbes comparatifs (suffixés en <i>-er</i> , pas un sens comparatif strict comme <i>later</i> dans <i>We can always come by later</i> )
ADVERBES	RBS	Adverbes superlatifs
ADVERBES	WRB	Wh-adverbes ( <i>how</i> , <i>where</i> , <i>why</i> , etc.). <i>WHEN</i> est étiqueté WRB dans son acception temporelle (IN dans le sens de IF)
CARDINAUX	CD	Nombres cardinaux
COORDONNANTS	CC	Conjonctions de coordination ( <i>and</i> , <i>but</i> , <i>nor</i> , <i>or</i> , <i>for</i> et <i>yet</i> + opérateurs mathématiques <i>plus</i> , <i>minus</i> , <i>less</i> , <i>times</i> et <i>over</i> )
DETERMINANTS	DT	Déterminants (articles <i>a(n)</i> , <i>every</i> , <i>no</i> et <i>the</i> , indéfinis <i>another</i> , <i>any</i> , <i>some</i> , <i>each</i> , <i>either</i> , <i>neither</i> , <i>that</i> , <i>these</i> , <i>this</i> et <i>those</i> + <i>all</i> et <i>both</i> lorsqu'ils ne précèdent pas un déterminant ou un pronom possessif).
PREDETERMINANTS	PDT	Éléments suivants lorsqu'ils précèdent un article ou un pronom possessif: <i>all</i> , <i>both</i> , <i>half</i> , <i>many</i> , <i>quite</i> , <i>rather</i> , <i>such</i> , etc.
INTERJECTIONS	UH	Interjections ( <i>my</i> , <i>oh</i> , <i>please</i> , <i>see</i> , <i>uh</i> , <i>well</i> , <i>yes</i> , etc.)
PARTICIPES PASSES	VBN	Verbes au participe passé
PARTICULES	RP	Monosyllabiques pour la plupart, susceptibles d'être redoublées en tant qu'adverbes directionnels ou prépositions. Ex.: <i>she told off</i> her friends, particule car non dépendant du SN
POSSESSIFS	POS	Marques de possession en position finale: <i>'s</i> ou <i>'</i> (éléments écartés des mots auquel ils se rattachent comme s'ils étaient eux-mêmes des mots. Ex: <i>John/NP 's/POS</i> idea.)
PREPOSITIONS SUBORDONNANTS	ou IN	Pas de distinction, mais toujours possible de la retrouver, car une préposition est un IN qui précède un SN ou un SP, et une conjonction de subordination un IN qui précède une proposition. A noter que la préposition <i>to</i> a son propre tag.
PRONOMS	PP	Pronoms personnels + pronoms possessifs <i>mine</i> , <i>yours</i> , <i>his</i> , <i>hers</i> , <i>ours</i> et <i>theirs</i>
PRONOMS	PP'S	Pronoms possessifs ( <i>my</i> , <i>your</i> , <i>his</i> , <i>her</i> , <i>its</i> , <i>one's</i> , <i>our</i> et <i>their</i> ).
PRONOMS	WP	Wh-pronoms ( <i>what</i> , <i>who</i> et <i>whom</i> ).
PRONOMS	WP'S	Pronom possessif en wh-: <i>WHOSE</i> .
RELATIFS	WDT	Wh-determiner ( <i>which</i> et <i>that</i> – employé en tant que pronom relatif).
SUBSTANTIFS	NN	Noms au singulier (ou <i>mass nouns</i> ).

<b>SUBSTANTIFS</b>	<b>NNS</b>	Noms au pluriel
<b>SUBSTANTIFS</b>	<b>NP</b>	Noms propres au singulier
<b>SUBSTANTIFS</b>	<b>NPS</b>	Noms propres au pluriel
<b>THERE</b>	<b>EX</b>	"there" existentiel (qui entraîne une inversion du verbe et du sujet logique d'une phrase)
<b>VERBES</b>	<b>MD</b>	Verbes modaux (critère: V qui ne prennent pas un -s à la 3 <sup>e</sup> ps: <i>can, could, (dare), may, might, must, ought, shall, should, will</i> et <i>would</i> ).
<b>VERBES</b>	<b>VB</b>	Verbes, forme "de base" (impératifs, infinitifs et subjonctifs).
<b>VERBES</b>	<b>VBD</b>	Verbes au passé + forme conditionnelle du verbe <i>to be</i> (If I were/VBD rich...)
<b>VERBES</b>	<b>VBG</b>	Verbes au gérondif ou au participe présent
<b>VERBES</b>	<b>VBP</b>	Verbes au présent, formes autres que la troisième personne du singulier
<b>VERBES</b>	<b>VBZ</b>	Verbes au présent, troisième personne du singulier
<b>RESIDUS</b>	<b>FW</b>	Mots d'origine étrangère (pas de liste particulière)
<b>RESIDUS</b>	<b>SYM</b>	Symboles ou expressions mathématiques, scientifiques ou techniques qui ne sont pas des mots (anglais). (usage important. E.g. noms de symboles et d'unités de mesures (abréviations incluses) = noms).
<b>RESIDUS</b>	<b>LS</b>	Marqueurs d'éléments dans une liste (lettres et numéraux)
<b>TO</b>	<b>TO</b>	TO

Tableau 3 : [anglais] tableau synthétisant les codes du corpus Penn Treebank

POS ou catégorie	Code	Description
<b>ADJECTIFS</b>	<b>JJ</b>	Adjectifs
<b>ADJECTIFS</b>	JJQ	Adjectifs en <i>wh-</i>
<b>ADJECTIFS</b>	JJV	Adjectifs en <i>wh...ever</i>
<b>ADJECTIFS</b>	JJX	Adjectifs superlatifs
<b>ADJECTIFS</b>	JJR	Adjectifs comparatifs
<b>ADJECTIFS</b>	JJH	Adjectifs postposés ("heavy")
<b>ADVERBES</b>	RR	Adverbes
<b>ADVERBES</b>	RRQ	Adverbes en <i>wh-</i>
<b>ADVERBES</b>	RRV	Adverbes en <i>wh...ever</i>
<b>ADVERBES</b>	RRX	Adverbes superlatifs
<b>ADVERBES</b>	RRR	Adverbes comparatifs
<b>ADVERBES</b>	RRS	Adverbes marquant une asyndète
<b>ADVERBES</b>	RRW	Adverbes quasi-nominaux

Tableau 4 : [anglais] tableau (incomplet) synthétisant les codes développés par SUSANNE

POS ou catégorie	Français	Anglais
ADJECTIFS	[1] Trait nombre (sg/pl) [2] Numéraux	[3] Adj. Comparatifs/superlatifs [4] Adj. Comparatifs/superlatifs, en <i>wh-</i> , postposés
ADVERBES	CG	[3] Adv. Comparatifs/superlatifs/en <i>wh-</i> [4] Adv. comparatifs/superlatifs, en <i>wh-</i> , interrogatifs, etc.
COORDONNANTS	[1] + expressions ( <i>soit...soit</i> ) [2] CG	[3] + opérateurs mathématiques [4] XX
DETERMINANTS	[1] distinction non contractés/ contractés (DU et DES)	[3] distinction non contractés /contractés (catég. <i>prep.</i> )

	Trait nombre (sg/pl) [2] déterminants définis/indéfinis adjectifs démonstratifs, possessifs (catég. <i>pronom</i> ) adjectifs indéfinis	distinction déterminants/ pré-déterminants (précèdent déterminant ou possessif) adjectifs possessifs [4] distinction déterminants et articles, déterminants sg, pl et indénombrables, en <i>wh-</i> , etc.
INTERJECTIONS	[1] CG [2] XX	[3] CG [4] ???
NUMERAUX/ CARDINAUX	[1] nombres cardinaux (chiffres /lettres), dates [2] XX	[3] nombres cardinaux (chiffres /lettres), dates [4] nombres cardinaux (chiffres /lettres), dates; un tag unique pour ONE
PARTICIPES PASSES	[1] distinction auxiliaires ETRE et AVOIR/autres contextes Trait nombre (sg/pl) [2] CG	CG
PARTICULARITES	[1] catégorie préfixes détachés, isolés entre deux blancs (ex <i>micro...</i> ) [2] XX	[3] étiquetage du 'S ou ' possessif anglais et du THERE existentiel [4] XX
PARTICULES	[1] éléments non autonomes et non regroupés dans des expressions figées [2] XX	[3] particules essentiellement monosyllabiques, souvent prépositions non rattachées à un SN [4] ???
PONCTUATIONS	[1] syntaxe [signe] / [signe] [2] distinction des points/virgules/ autres marques de ponctuation	Idem [1]
PREPOSITIONS	[1] prépositions simples [2] CG	[3] Regroupement des prépositions avec les subordonnants tag unique pour la préposition TO [4] CG
PRONOMS	[1] distinction pronoms relatifs/non relatifs supportés/non supportés par le V Trait nombre (sg/pl) [2] pronoms démonstratifs pronoms indéfinis distinction pronoms personnels clitiques/conjugués pronoms relatifs	[3] Regroupement pronoms personnels/possessifs Pronom possessif WHOSE Pronoms relatifs (wh-) [4] un tag par pronom personnel (sauf HE et SHE regroupés) Pronoms relatifs (en wh-), distinction du THAT
RESIDUS	[1] abréviations, mots étrangers et symboles [2] abréviations et symboles	[3] mots étrangers, symboles et marqueurs d'éléments dans une liste [4] ???
SUBORDONNANTS	[1] distinction des conjonctions et locutions de subordinations/QUE [2] CG	[3] & [4] regroupés avec les prépositions
SUBSTANTIFS	[1] distinction noms communs /propres Trait nombre (sg/pl) [2] regroupement des NC et NP, pas de trait nombre	[3] distinction noms communs /propres Trait nombre (sg/pl) [4] distinction NP (pas de trait nombre)/NC sg/pl + autres distinctions comme "substantifs à tête adjectivale singulier", etc.
VERBES	[1] distinction auxiliaires AVOIR/ ETRE/AUTRE VERBE distinction formes conjuguées/non conjuguées (gérondif & participe présent/infinitif) Trait nombre (sg/pl) [2] étiquetage des temps verbaux simples, mais pas des formes composés (auxiliaires)	[3] modaux "formes de base" (impératifs, infinitifs et subjonctifs) "verbes passés" & forme conditionnelle BE Gérondifs & participes présent Distinction verbes au présent 3PS/Autres personnes [4] granularité élevée, tags individuels pour BE, WAS, WERE, ARE, etc.; temps,

		formes en –ing, etc.
--	--	----------------------

Tableau 5 : Synthèse des quatre systèmes d'étiquetage documentés

*Légende*

Codes: correspondent aux systèmes d'annotation documentés dans les tableaux précédents :

[1] système d'annotation de l'Inalf pour Brill

[2] système d'annotation de TreeTagger

[3] système d'annotation de Penn Treebank

[4] système d'annotation de SUSANNE

CG: catégorie grammair traditionnelle le plus souvent, pas de sous-distinction.

???: catégorie non documentée (ne concerne que le système d'annotation SUSANNE)

XX: catégorie inexistante

## Annexe 10 : Statistiques descriptives (chapitre 3)

### Classes

	Min.	% min..	Max.	Total	Moyenne par texte	Ecart type	Coeff. variation	Variance
<b>Symboles/Sigles /abréviations</b>	0,17	0,45	8,62	444,94	1,99	1,51	0,76	2,28
<b>Adverbes/ connecteurs</b>	3,75	0,45	12,39	1684,56	7,52	1,29	0,17	1,66
<b>Adjectifs</b>	3,45	0,45	10,41	1487,90	6,64	1,33	0,20	1,76
<b>Pronoms</b>	1,86	0,45	14,18	1352,87	6,04	1,86	0,31	3,47
<b>Verbes</b>	6,69	0,45	17,20	2555,18	11,41	1,69	0,15	2,86
<b>Déterminants</b>	6,08	0,45	15,53	2681,28	11,97	1,52	0,13	2,31
<b>Noms</b>	12,44	0,45	25,36	4318,00	19,28	2,28	0,12	5,20
<b>Prépositions</b>	7,52	0,45	17,16	2963,76	13,23	1,83	0,14	3,34
<b>Ponctuations</b>	8,02	0,45	24,91	3091,99	13,80	2,54	0,18	6,46
<b>Subordonnants</b>	0,46	0,45	4,20	318,82	1,42	0,49	0,34	0,24
<b>Numéraux</b>	0,58	0,45	6,71	572,62	2,56	1,17	0,46	1,36
<b>Éléments étrangers</b>	0,00	17,41	33,89	444,69	1,99	4,14	2,09	17,14
<b>Interjections</b>	0,00	70,54	1,58	8,52	0,04	0,15	4,03	0,02
<b>Préfixes</b>	0,00	0,45	2,17	63,59	0,28	0,24	0,84	0,06
<b>PUL</b>	0,00	17,41	0,31	8,63	0,04	0,04	1,00	0,00

Tableau : statistiques descriptives des classes (FT)

	Min.	% min..	Max.	Total	Moyenne par texte	Ecart type	Coeff. variation	Variance
<b>Symboles/Sigles /abréviations</b>	0,11	0,45	9,02	381,12	1,70	1,38	0,81	1,91
<b>Adverbes/ connecteurs</b>	4,41	0,45	13,79	1763,94	7,87	1,26	0,16	1,60
<b>Adjectifs</b>	3,95	0,45	11,25	1585,46	7,08	1,29	0,18	1,68
<b>Pronoms</b>	1,88	0,45	13,04	1300,57	5,81	1,55	0,27	2,41
<b>Verbes</b>	7,51	0,45	16,50	2544,38	11,36	1,43	0,13	2,03
<b>Déterminants</b>	8,45	0,45	15,53	2791,40	12,46	1,25	0,10	1,55
<b>Noms</b>	14,00	0,45	25,40	4439,83	19,82	1,96	0,10	3,86
<b>Prépositions</b>	7,95	0,45	17,45	3131,69	13,98	1,61	0,11	2,58
<b>Ponctuations</b>	8,17	0,45	24,89	2928,33	13,07	2,42	0,19	5,85
<b>Subordonnants</b>	0,46	0,45	4,12	326,81	1,46	0,47	0,32	0,22
<b>Numéraux</b>	0,58	0,45	5,84	498,55	2,23	1,04	0,47	1,08
<b>Éléments étrangers</b>	0,00	21,88	9,40	158,73	0,71	1,42	2,00	2,01
<b>Interjections</b>	0,00	83,93	0,29	1,36	0,01	0,03	4,11	0,00
<b>Préfixes</b>	0,00	0,45	1,89	67,43	0,30	0,24	0,79	0,06
<b>PUL</b>	0,00	19,64	0,22	8,82	0,04	0,04	0,96	0,00

Tableau : statistiques descriptives des classes (TE)

## Marques de personne

	Min.	% min..	Max.	Total	Moyenne par texte	Ecart type	Coeff. variation	Variance
<b>JE</b>	0	16,071	127	3256	14,54	20,26	1,39	410,56
<b>clitique</b>	0	40,625	47	986	4,40	7,14	1,62	50,94
<b>disjoint</b>	0	64,286	11	211	0,94	1,87	1,98	3,49
<b>D. possessif</b>	0	42,857	32	747	3,33	5,60	1,68	31,38
<b>P. possessif</b>	0	98,214	7	12	0,05	0,51	9,47	0,26
<b>TU</b>	0	63,393	325	808	3,61	22,46	6,23	504,27
<b>clitique</b>	0	76,786	16	179	0,80	2,07	2,59	4,28
<b>disjoint</b>	0	81,696	11	96	0,43	1,33	3,10	1,76
<b>D. possessif</b>	0	77,232	29	181	0,81	3,14	3,89	9,87
<b>P. possessif</b>	0	99,107	8	10	0,04	0,55	12,33	0,30
<b>IL impersonnel</b>	1	0,446	301	7424	33,14	27,75	0,84	769,80
<b>IL/ELLE</b>	3	0,893	230	8322	37,15	30,62	0,82	937,61
<b>disjoint</b>	0	11,161	45	1204	5,38	5,72	1,06	32,70
<b>D. possessif</b>	2	1,786	142	5741	25,63	20,32	0,79	413,10
<b>P. possessif</b>	0	88,393	5	40	0,18	0,59	3,33	0,35
<b>ON</b>	3	0,893	154	8131	36,30	23,91	0,66	571,84
<b>NOUS</b>	0	1,786	84	3762	16,79	15,48	0,92	239,78
<b>clitique</b>	0	16,071	31	954	4,26	4,57	1,07	20,87
<b>disjoint</b>	0	68,750	11	139	0,62	1,38	2,22	1,90
<b>D. possessif</b>	0	16,964	30	1048	4,68	5,12	1,09	26,21
<b>P. possessif</b>	0	91,964	2	19	0,08	0,29	3,48	0,09
<b>VOUS</b>	0	64,286	54	479	2,14	5,76	2,70	33,23
<b>clitique</b>	0	76,786	28	176	0,79	2,78	3,54	7,73
<b>disjoint</b>	0	87,054	7	59	0,26	0,88	3,35	0,78
<b>D. possessif</b>	0	80,357	16	121	0,54	1,75	3,23	3,05
<b>P. possessif</b>	0	98,661	1	3	0,01	0,12	8,60	0,01
<b>ILS/ELLES</b>	0	1,786	67	2862	12,78	11,01	0,86	121,27
<b>disjoint</b>	0	16,964	17	592	2,64	2,80	1,06	7,84
<b>D. possessif</b>	0	1,786	62	2304	10,29	9,53	0,93	90,74
<b>P. possessif</b>	0	98,214	1	4	0,02	0,13	7,43	0,02

Tableau : statistiques descriptives des marques de personne (FT)

	Min.	% min..	Max.	Total	Moyenne par texte	Ecart type	Coeff. variation	Variance
<b>JE</b>	0	28,12	101	1732	7,73	12,70	1,64	161,29
<b>clitique</b>	0	50,89	29	520	2,32	4,29	1,85	18,39
<b>disjoint</b>	0	79,91	9	73	0,33	0,90	2,77	0,81
<b>D. possessif</b>	0	54,91	32	395	1,76	3,58	2,03	12,83
<b>P. possessif</b>	0	98,66	7	10	0,04	0,49	10,98	0,24
<b>TU</b>	0	81,25	210	332	1,48	14,13	9,53	199,53
<b>clitique</b>	0	90,62	10	44	0,20	0,93	4,72	0,86
<b>disjoint</b>	0	91,07	5	31	0,14	0,55	3,95	0,30
<b>D. possessif</b>	0	89,73	23	59	0,26	1,65	6,25	2,71
<b>P. possessif</b>	0	99,55	1	1	0,00	0,07	14,97	0,00
<b>IL impersonnel</b>	1	0,45	166	6637	29,63	21,04	0,71	442,58
<b>IL/ELLE</b>	1	0,45	230	6532	29,16	24,38	0,84	594,22
<b>réflexif</b>	3	0,45	177	9634	43	27	0,62	729,22
<b>disjoint</b>	0	15,18	45	1058	4,72	5,37	1,14	28,81
<b>D. possessif</b>	1	0,45	142	4981	22,24	19,47	0,88	379,02
<b>P. possessif</b>	0	93,75	3	18	0,08	0,35	4,29	0,12

<b>ON</b>	2	0,45	154	7576	33,82	23,13	0,68	534,87
<b>NOUS</b>	0	4,46	84	3556	15,88	15,39	0,97	236,92
<b>clitique</b>	0	19,64	31	859	3,83	4,47	1,17	19,99
<b>disjoint</b>	0	75	11	113	0,50	1,31	2,59	1,70
<b>D. possessif</b>	0	20,54	30	985	4,40	5,11	1,16	26,09
<b>P. possessif</b>	0	93,30	1	15	0,07	0,25	3,74	0,06
<b>VOUS</b>	0	82,14	29	176	0,79	2,98	3,80	8,90
<b>clitique</b>	0	90,18	28	59	0,26	1,92	7,30	3,70
<b>disjoint</b>	0	95,54	3	13	0,06	0,30	5,19	0,09
<b>D. possessif</b>	0	91,07	14	45	0,20	1,09	5,44	1,19
<b>P. possessif</b>	0	99,55	1	1	0,00	0,07	14,97	0,00
<b>ILS/ELLES</b>	0	1,79	67	2518	11,24	9,68	0,86	93,67
<b>disjoint</b>	0	19,2	17	550	2,46	2,73	1,11	7,43
<b>D. possessif</b>	0	2,68	60	2139	9,55	9,44	0,99	89,12
<b>P. possessif</b>	0	98,66	1	3	0,01	0,12	8,60	0,01

Tableau : statistiques descriptives des marques de personne (TE)<sup>5</sup>

## Temps verbaux

	Min.	% min..	Max.	Total	Moyenne par texte	Ecart type	Coeff. variation	Variance
<b>Présent</b>	81	0,45	1031	71023	317,07	156,46	0,49	24478,50
<b>Infinitif</b>	19	0,45	503	34470	153,88	86,43	0,56	7470,55
<b>Participes passé</b>	28	0,45	432	33607	150,03	73,28	0,49	5370,52
<b>Aux. présent</b>	9	0,45	272	13468	60,13	34,10	0,57	1162,65
<b>Mod. présent</b>	2	0,45	161	8994	40,15	23,51	0,59	552,63
<b>Participes présent</b>	0	0,45	134	7684	34,30	23,06	0,67	531,55
<b>Futur</b>	0	0,89	77	3069	13,70	10,82	0,79	117
<b>Imparfait</b>	0	4,46	109	2693	12,02	14,44	1,20	208,57
<b>Aux. infinitif</b>	0	0,45	51	2361	10,54	8,17	0,77	66,72
<b>Conditionnel</b>	0	5,36	56	1962	8,76	8,03	0,92	64,48
<b>Subjonctif prés.</b>	0	7,14	78	1139	5,08	6,29	1,24	39,50
<b>Impératif</b>	0	20,98	57	904	4,04	5,98	1,48	35,71
<b>Aux. part. passé</b>	0	10,27	30	1101	4,92	4,88	0,99	23,77
<b>Modaux cond</b>	0	10,71	30	1074	4,79	4,70	0,98	22,11
<b>Passé simple</b>	0	58,48	33	479	2,14	4,69	2,19	22,03
<b>Aux. imparfait</b>	0	32,14	28	607	2,71	3,87	1,43	15,01
<b>Modaux inf.</b>	0	26,79	43	514	2,29	3,66	1,59	13,39
<b>Aux. conditionnel</b>	0	38,84	28	392	1,75	3,12	1,79	9,76
<b>Mod. part. passé</b>	0	29,02	20	401	1,79	2,21	1,24	4,89
<b>Aux. subj. prés.</b>	0	27,68	14	447	2,00	2,19	1,10	4,79
<b>Aux. part. prés.</b>	0	33,48	11	360	1,61	1,92	1,19	3,67
<b>Modaux impf.</b>	0	44,64	13	288	1,29	1,89	1,47	3,57
<b>Mod. subj. prés.</b>	0	37,95	15	314	1,40	1,89	1,35	3,59
<b>Modaux futur</b>	0	46,43	11	278	1,24	1,80	1,45	3,25
<b>Aux. futur</b>	0	36,16	7	355	1,58	1,75	1,10	3,05
<b>Aux. passé simple</b>	0	85,27	5	57	0,25	0,74	2,91	0,55
<b>Subj. imparfait</b>	0	82,59	8	54	0,24	0,72	2,98	0,52
<b>Mod. passé simple</b>	0	95,98	6	18	0,08	0,49	6,16	0,24

Tableau : statistiques descriptives des formes verbales (FT)

<sup>5</sup> Les éléments marginaux du corpus, *i.e.* qui ont un CV > 3, sont surlignés en jaune ; à l'inverse, les éléments a priori caractéristiques du genre de l'article sont en bleu.

	Min.	% min..	Max.	Total	Moyenne par texte	Ecart type	Coeff. variation	Variance
<b>Présent</b>	62	42,41	1031	64071	286,03	142,24	0,50	20232,81
<b>Infinitif</b>	16	39,73	503	30793	137,46	80,10	0,58	6417,48
<b>Participes passé</b>	26	42,86	381	30152	134,60	69,09	0,51	4774,26
<b>Aux. présent</b>	6	0,45	180	11139	49,72	27,64	0,55	764,36
<b>Mod. présent</b>	2	33,48	157	8525	38,05	22,87	0,60	523,05
<b>Participes présent</b>	0	12,95	120	7087	31,63	19,90	0,62	396,02
<b>Futur</b>	0	91,52	77	2685	11,98	9,93	0,82	98,77
<b>Aux. infinitifs</b>	0	7,14	51	2269	10,12	7,89	0,77	62,33
<b>Conditionnel</b>	0	31,25	56	1771	7,90	7,47	0,94	55,79
<b>Imparfait</b>	0	0,45	109	1678	7,49	12,14	1,62	147,48
<b>Subjonctif prés.</b>	0	0,89	48	1021	4,55	4,73	1,03	22,41
<b>Modaux cond.</b>	0	26,34	28	1002	4,47	4,44	0,99	19,73
<b>Aux. part. passé</b>	0	10,27	30	928	4,14	4,38	1,05	19,19
<b>Impératif</b>	0	0,45	23	689	3,07	4,17	1,35	17,45
<b>Aux. imparfait</b>	0	11,61	28	388	1,73	3,10	1,79	9,65
<b>Passé simple</b>	0	47,77	33	142	0,63	2,48	3,91	6,16
<b>Modaux inf.</b>	0	54,02	14	437	1,95	2,19	1,12	4,83
<b>Mod. part. passé</b>	0	27,23	20	338	1,50	2,08	1,37	4,33
<b>Aux. subj. prés.</b>	0	35,71	10	411	1,83	1,96	1,07	3,87
<b>Mod. subj. prés.</b>	0	0,45	14	300	1,33	1,84	1,37	3,40
<b>Aux. conditionnel</b>	0	99,11	12	294	1,31	1,83	1,40	3,38
<b>Aux. part. prés.</b>	0	40,63	9	343	1,53	1,74	1,14	3,05
<b>Aux. futur</b>	0	0,45	7	324	1,44	1,72	1,19	2,96
<b>Modaux futur</b>	0	0,45	11	255	1,13	1,67	1,47	2,81
<b>Modaux impf.</b>	0	0,45	10	225	1,00	1,67	1,66	2,79
<b>Subj. imparfait</b>	0	76,79	8	42	0,18	0,68	3,64	0,46
<b>Aux. passé simple</b>	0	87,05	5	31	0,13	0,54	3,95	0,29
<b>Mod. passé simple</b>	0	8,04	2	3	0,01	0,14	11,14	0,02

Tableau : statistiques descriptives des formes verbales (TE)

## Ponctuations

	Min.	% min..	Max.	Total	Moyenne par texte	Ecart type	Coeff. variation	Variance
<b>Virgules</b>	84	0,45	1431	85678	382,49	214,96	0,56	46206,03
<b>Points</b>	29	0,45	650	41868	186,91	97,00	0,52	9408,62
<b>Parenthèses</b>	14	0,45	474	39278	175,35	97,10	0,55	9429,22
<b>Guillemets</b>	0	0,45	737	21644	96,63	97,95	1,01	9595,17
<b>Deux points</b>	4	0,89	178	13283	59,30	33,68	0,57	1134,13
<b>Slashes</b>	0	6,70	562	4390	19,60	43,16	2,20	1862,39
<b>Crochets</b>	0	37,50	142	4173	18,63	29,91	1,61	894,89
<b>P. virgules</b>	0	3,57	118	3650	16,29	16,63	1,02	276,68
<b>Cotes</b>	0	62,95	523	3344	14,93	48,51	3,25	2352,95
<b>Tirets</b>	0	11,16	251	2796	12,48	20,17	1,62	406,91
<b>P. suspension</b>	0	12,05	107	2680	11,96	15,98	1,34	255,32
<b>P. interrogation</b>	0	12,95	68	1925	8,59	9,59	1,12	91,91
<b>P. exclamation</b>	0	51,34	224	745	3,33	16,35	4,92	267,49
<b>Accolades</b>	0	89,73	176	414	1,85	12,80	6,92	163,8
<b>Antislashes</b>	0	95,98	35	103	0,46	3,33	7,24	11,1

Tableau : statistiques descriptives des ponctuations (FT)



	Min.	% min..	Max.	Total	Moyenne par texte	Ecart type	Coeff. variation	Variance
<b>Virgules</b>	64	0,4	1293	76787	342,80	199,43	0,58	39773,13
<b>Points</b>	27	0,4	601	36759	164,10	90,82	0,55	8248,46
<b>Parenthèses</b>	12	0,4	452	32154	143,54	86,26	0,60	7440,72
<b>Guillemets</b>	0	0,4	737	19602	87,51	88,36	1,01	7807,61
<b>Deux points</b>	1	0,4	172	11893	53,09	29,76	0,56	885,94
<b>P. virgules</b>	0	4,5	101	3183	14,21	14,58	1,03	212,53
<b>Cotes</b>	0	67,0	523	2688	12,00	43,46	3,62	1889,06
<b>Crochets</b>	0	46,4	142	2572	11,48	22,13	1,93	489,65
<b>Slashes</b>	0	12,5	175	2368	10,57	15,97	1,51	255,02
<b>Tirets</b>	0	14,7	76	2361	10,54	12,17	1,15	148,17
<b>P. suspension</b>	0	21,9	48	1403	6,26	8,34	1,33	69,59
<b>P. interrogation</b>	0	18,8	50	1332	5,95	7,38	1,24	54,42
<b>P. exclamation</b>	0	69,6	47	228	1,02	4,51	4,43	20,38
<b>Accolades</b>	0	92,9	18	114	0,51	2,25	4,41	5,04
<b>Antislashes</b>	0	99,6	1	1	0,00	0,07	14,97	0,00

Tableau : statistiques descriptives des ponctuations (TE)

## Numéraux

	Min.	% min..	Max.	Total	Moyenne par texte	Ecart type	Coeff. variation	Variance
<b>Cardinaux</b>	7	0,89	396	14230	63,53	48,78	0,77	2379,36
<b>Dates</b>	0	2,68	153	6508	29,05	26,72	0,92	713,84
<b>Ordinaux</b>	1	0,89	63	3465	15,47	11,47	0,74	131,53

Tableau : statistiques descriptives des numéraux (FT)

	Min.	% min..	Max.	Total	Moyenne par texte	Ecart type	Coeff. variation	Variance
<b>Cardinaux</b>	2	0,45	211	11597	51,77	37,29	0,72	1390,63
<b>Dates</b>	0	4,02	153	5923	26,44	25,92	0,98	671,68
<b>Ordinaux</b>	0	0,45	63	3290	14,69	11,40	0,78	129,92

Tableau : statistiques descriptives des numéraux (TE)

## Marqueurs de structuration textuelle et de renvoi dans le texte

	Min.	% min..	Max.	Total	Moyenne par texte	Ecart type	Coeff. variation	Variance
<b>Indices de liste</b>	0	0,89	233	11551	51,57	39,91	0,77	1593,03
<b>Indices de renvoi</b>	0	27,23	284	4722	21,08	35,04	1,66	1227,72

Tableau : statistiques descriptives des marqueurs de structuration textuelle et de renvoi dans le texte (FT)

	Min.	% min..	Max.	Total	Moyenne par texte	Ecart type	Coeff. variation	Variance
<b>Indices de liste</b>	0	1,34	103	5854	26,13	19,19	0,73	368,34
<b>Indices de renvoi</b>	0	27,68	283	4628	20,66	34,62	1,68	1198,79

Tableau : statistiques descriptives des marqueurs de structuration textuelle et de renvoi dans le texte (TE)

## Marques de formalisation

	Min.	% min..	Max.	Total	Moyenne par texte	Ecart type	Coeff. variation	Variance
<b>Symboles</b>	0	4,91	509	13841	61,79	81,51	1,32	6644,15
<b>Abréviations</b>	0	0,45	409	7146	31,90	39,40	1,24	1552,63
<b>Sigles linguistiques</b>	0	33,93	209	4585	20,47	40,19	1,96	1615,09
<b>Sigles</b>	0	20,54	216	2273	10,15	21,80	2,15	475,39
<b>Symboles gram.</b>	0	55,36	225	2163	9,66	30,72	3,18	943,87
<b>Symboles ling.</b>	0	45,98	85	1233	5,50	11,04	2,00	121,79

Tableau : statistiques descriptives des marques de formalisation (FT)

	Min.	% min..	Max.	Total	Moyenne par texte	Ecart type	Coeff. variation	Variance
<b>Symboles</b>	0	6,70	509	10949	48,88	70,19	1,44	4926,85
<b>Abréviations</b>	0	0,89	195	5613	25,06	28,05	1,12	787,03
<b>Sigles linguistiques</b>	0	37,95	194	3774	16,85	35,34	2,10	1248,62
<b>Sigles</b>	0	26,34	216	1951	8,71	21,44	2,46	459,85
<b>Symboles gram.</b>	0	57,59	225	1445	6,45	22,09	3,42	487,78
<b>Symboles ling.</b>	0	66,07	35	343	1,53	4,32	2,82	18,66

Tableau : statistiques descriptives des marques de formalisation (TE)

## Connecteurs

	Min.	% min..	Max.	Total	Moyenne par texte	Ecart type	Coeff. variation	Variance
<b>Addition</b>	21	0,45	432	28423	126,89	67,13	0,53	4506,49
<b>Disjonction</b>	5	0,45	230	9478	42,31	27,24	0,64	741,85
<b>Opposition</b>	1	0,45	204	7601	33,93	26,34	0,78	693,54
<b>Conséquence</b>	1	2,68	83	3665	16,36	12,57	0,77	157,88
<b>Justification</b>	0	0,45	51	3592	16,04	10,85	0,68	117,63
<b>Espace</b>	0	2,23	53	2359	10,53	7,82	0,74	61,23
<b>Reformulation</b>	0	1,34	40	1855	8,28	6,68	0,81	44,60
<b>Exemplification</b>	0	4,02	46	1828	8,16	7,80	0,96	60,87
<b>Causalité</b>	0	5,36	52	1734	7,74	6,97	0,90	48,63
<b>Temps</b>	0	4,02	34	1329	5,93	4,91	0,83	24,11
<b>Doute</b>	0	18,75	46	767	3,42	4,62	1,35	21,31
<b>Conclusion</b>	0	16,52	16	696	3,11	3,06	0,99	9,38
<b>Concession</b>	0	24,55	37	600	2,68	3,67	1,37	13,44
<b>Présupposition</b>	0	26,34	23	507	2,26	2,77	1,22	7,66
<b>Certitude</b>	0	42,41	16	288	1,29	1,80	1,40	3,23

Tableau : statistiques descriptives des connecteurs (FT)

	Min.	% min..	Max.	Total	Moyenne par texte	Ecart type	Coeff. variation	Variance
<b>Addition</b>	21	0,45	417	26884	120,02	66,67	0,56	4444,87
<b>Disjonction</b>	5	0,45	230	9230	41,21	26,74	0,65	715,22
<b>Opposition</b>	1	0,45	166	7109	31,74	23,87	0,75	569,81
<b>Justification</b>	0	0,89	51	3488	15,57	10,52	0,68	110,65

<b>Conséquence</b>	0	0,45	66	3470	15,49	11,81	0,76	139,45
<b>Espace</b>	0	3,13	53	2224	9,93	7,42	0,75	55,01
<b>Reformulation</b>	0	1,79	40	1833	8,18	6,65	0,81	44,20
<b>Exemplification</b>	0	4,02	46	1780	7,95	7,77	0,98	60,45
<b>Causalité</b>	0	6,25	52	1606	7,17	6,87	0,96	47,23
<b>Temps</b>	0	7,14	34	1170	5,22	4,53	0,87	20,51
<b>Doute</b>	0	19,64	46	724	3,23	4,51	1,40	20,37
<b>Conclusion</b>	0	17,41	16	659	2,94	2,91	0,99	8,49
<b>Certitude</b>	0	26,34	37	587	2,62	3,66	1,40	13,40
<b>Présupposition</b>	0	26,79	23	491	2,19	2,70	1,23	7,28
<b>Concession</b>	0	45,09	16	267	1,19	1,75	1,47	3,06

Tableau : statistiques descriptives des connecteurs (TE)

## *Annexe 11 : Caractéristiques des 15 styles d'auteur (chapitre 7)*

### Valeur test (V. test)

La valeur test est un critère statistique auquel est associé une probabilité. Le principe en est le suivant. Pour évaluer l'ampleur des différences entre proportions ou entre moyennes, on réalise des tests statistiques que l'on exprime finalement en nombre d'écart-types d'une loi normale. La valeur-test est égale à ce nombre d'écart-types. Ainsi lorsque la valeur-test est supérieure à 2 en valeur absolue, un écart est significatif au seuil usuel (5%). En rangeant les items dans l'ordre décroissant des valeurs-tests, on range les items dans l'ordre de leur importance pour caractériser un objet (...) ». De manière complémentaire : « une modalité (ou une catégorie) d'une variable nominale est considérée comme caractéristique de la classe si son abondance dans la classe est jugée significativement supérieure à ce qu'on peut attendre compte tenu de sa présence dans la population ».

## **Authier**

### *Lexique*

Lexique					
Mot	Pourcentage		Fréquence		V. test
	Auteur	Corpus	Auteur	Corpus	
<b>dire</b>	1.53	0.27	814	1708	42.739
<b>méta</b>	0.25	0.02	131	142	23.843
<b>mots</b>	0.53	0.11	280	695	22.810
<b>autonymique</b>	0.19	0.02	103	111	21.195
<b>formes</b>	0.51	0.14	271	853	19.390
	0.00	0.09	0	540	-9.433
<b>politique</b>	0.02	0.15	11	963	-9.978
<b>français</b>	0.01	0.13	5	838	-10.358
<b>les</b>	1.61	2.27	852	14259	-11.203
<b>nous</b>	0.11	0.40	61	2505	-12.552

## *Cordial + lexique épistémique*

Morphosyntaxe et lexique épistémique					
Variable	Moyenne		Ecart-type		V. test
	Auteur	Corpus	Auteur	Corpus	
%tirts/EnsPonct	6.533	1.341	1.776	1.837	9.29
appeler	12.500	2.762	9.190	4.481	7.14
relever	10.300	2.869	5.478	3.744	6.52
%V1ePersSgEnsVConj	8.367	2.773	1.815	2.866	6.42
terme	8.400	2.500	2.458	3.660	5.30
%PronPers1ePersSgEnsPronPers1s	14.911	6.329	2.455	5.448	5.18
%Ptsdesuspension/EnsPonct	3.454	1.375	0.786	1.542	4.43
NbMoyPronparProp	1.112	0.863	0.107	0.188	4.36
commentaire	5.800	1.303	5.510	3.459	4.27
%Pronindéfinis/EnsPron	7.509	4.749	3.161	2.257	4.02
NbtotPonct	1698.800	1130.205	276.691	474.216	3.94
choix	3.400	1.238	2.764	1.838	3.87
exemple	14.900	5.631	7.231	8.640	3.53
viser	2.100	0.746	1.136	1.394	3.19
%CrochAccol/EnsPonct	3.636	1.534	1.584	2.236	3.09
%PronPoss/EnsPron	0.309	0.126	0.338	0.199	3.03
%PropRelEnsProps	4.245	2.550	2.029	1.935	2.88
%VCOIObligEnsV	15.438	13.517	2.019	2.276	2.77
%VTransindirectsEnsV	53.311	50.103	2.037	4.078	2.59
signifier	3.400	1.270	5.903	2.799	2.50
question	6.200	3.213	2.993	4.189	2.34
%V3ePersPIEnsVConj	14.776	18.873	2.588	5.698	-2.36
expérimental	0.100	7.230	0.300	9.050	-2.59
hypothèse	0.400	3.418	0.490	3.806	-2.61
schéma	0.600	5.828	0.490	6.531	-2.63
%deux-points/EnsPonct	4.071	6.068	0.927	2.392	-2.74
présupposé	0.000	1.992	0.000	2.310	-2.83
NbMoyLETparmot	4.685	4.942	0.082	0.190	-4.46
%Pts/EnsPonct	8.901	16.861	1.707	5.095	-5.14

## *TnT + lexique épistémique*

Morphosyntaxe et lexique épistémique					
Variable	Moyenne		Ecart-type		V. test
	Auteur	Corpus	Auteur	Corpus	
PON:tiret	5.661	1.341	1.333	1.621	8.76
%classe11	19.182	13.883	0.935	2.212	7.87
appeler	12.500	2.762	9.190	4.481	7.14
PON:croch	7.720	1.578	3.565	3.041	6.64
relever	10.300	2.869	5.478	3.744	6.52
VER:parpres	8.500	4.668	1.479	2.000	6.30
PRO:pp1sn	6.382	2.253	1.512	2.268	5.98
PREF	0.575	0.248	0.234	0.191	5.61
terme	8.400	2.500	2.458	3.660	5.30
PON:susp	3.573	1.295	0.959	1.447	5.18

PRO:disj:pp3sn	3.028	1.549	0.585	0.980	4.96
ADJ:refl	2.186	1.087	0.771	0.808	4.47
commentaire	5.800	1.303	5.510	3.459	4.27
NUM:par	21.424	7.924	9.879	11.252	3.94
choix	3.400	1.238	2.764	1.838	3.87
DET:poss:pp3sn	5.021	3.320	1.649	1.474	3.79
exemple	14.900	5.631	7.231	8.640	3.53
SUB	94.219	89.305	2.613	4.735	3.41
DTC:sg	13.252	10.741	1.552	2.425	3.40
PRO:clit:pp1sn	1.476	0.654	0.594	0.820	3.30
viser	2.100	0.746	1.136	1.394	3.19
CON:exe:1st	3.875	2.300	1.610	1.625	3.19
PON:guil	15.695	9.841	2.918	6.288	3.06
Total Formes	10628.100	7811.517	2040.698	3391.615	2.73
PRO:poss:pp3sn	0.215	0.075	0.249	0.184	2.51
signifier	3.400	1.270	5.903	2.799	2.50
PON:Pvirg	2.239	1.416	1.043	1.123	2.41
%classe4	7.255	6.237	0.656	1.408	2.38
question	6.200	3.213	2.993	4.189	2.34
<b>NUM:dat</b>					
NUM:dat	9.018	18.398	4.458	13.229	-2.33
<b>ADJ:pl</b>					
ADJ:pl	24.539	29.949	3.578	7.627	-2.33
<b>DTC:pl</b>					
DTC:pl	6.552	8.536	1.333	2.763	-2.36
<b>%classe2</b>					
%classe2	6.787	7.898	0.266	1.427	-2.56
<b>expérimental</b>					
expérimental	0.100	7.230	0.300	9.050	-2.59
<b>hypothèse</b>					
hypothèse	0.400	3.418	0.490	3.806	-2.61
<b>schéma</b>					
schéma	0.600	5.828	0.490	6.531	-2.63
<b>NUM:ord</b>					
NUM:ord	4.988	13.265	3.421	10.109	-2.69
<b>présupposé</b>					
présupposé	0.000	1.992	0.000	2.310	-2.83
<b>%classe7</b>					
%classe7	18.566	20.393	0.565	2.109	-2.85
<b>%classe3</b>					
%classe3	5.950	6.905	0.457	1.004	-3.12
<b>PON:colon</b>					
PON:colon	3.472	5.921	0.904	2.501	-3.22
<b>PRO:pp3isn</b>					
PRO:pp3isn	4.232	6.386	1.290	2.168	-3.27
<b>SUB:1st</b>					
SUB:1st	5.781	10.695	2.613	4.735	-3.41
<b>PON:par</b>					
PON:par	8.584	16.175	2.121	6.790	-3.67
<b>CON:csq:1st</b>					
CON:csq:1st	1.899	4.247	0.690	2.042	-3.78
<b>PON:dot</b>					
PON:dot	8.082	15.710	1.349	5.151	-4.87

## Barberis

### Lexique

Lexique					
Mot	Pourcentage		Fréquence		V. test
	Auteur	Corpus	Auteur	Corpus	
rue	0.41	0.03	147	209	24.282
i	0.33	0.06	120	376	15.589
actualisation	0.18	0.02	67	114	15.202
tu	0.19	0.03	70	198	12.487
là	0.33	0.10	121	643	11.394
<b>contexte</b>					
contexte	0.49	0.79	179	4962	-6.969
<b>b</b>					
b	0.00	0.08	1	516	-6.974
<b>sujet</b>					
sujet	0.00	0.09	0	540	-7.639
<b>sémiotique</b>					
sémiotique	0.01	0.15	4	963	-9.049

discours	0.01	0.16	2	1026	-10.017
----------	------	------	---	------	---------

### *Cordial + lexique épistémique*

Morphosyntaxe et lexique épistémique					
Variable	Moyenne		Ecart-type		V. test
	Auteur	Corpus	Auteur	Corpus	
description	6.000	1.098	6.394	2.556	5.95
%deux-points/EnsPonct	10.320	6.068	1.331	2.392	5.52
explication	1.556	0.311	1.499	0.790	4.89
%PronPersls2ePersSgEnsPronPersls	4.944	1.132	7.163	2.656	4.46
proposer	5.222	1.770	4.076	2.445	4.38
%VauFut/EnsVConj	4.872	2.905	1.274	1.484	4.12
%V2ePersSgEnsVConj	3.006	1.194	4.275	1.528	3.68
montrer	4.222	1.672	3.083	2.474	3.20
savoir	7.222	2.590	13.105	4.643	3.10
%PropIndepEnsProps	89.342	84.755	2.569	5.257	2.71
%PronPersls2ePersPlEnsPronPersls	4.059	1.658	4.201	2.764	2.70
%V2ePersPlEnsVConj	2.246	0.928	2.252	1.589	2.57
Phénomène	4.000	1.656	2.981	2.868	2.54
%Ptsd'interrogation/EnsPonct	2.304	1.270	0.700	1.295	2.48
%PronPersls/EnsPron	60.261	55.384	3.909	6.151	2.46
Comprendre	2.333	1.098	2.404	1.575	2.43
questionnement	0.000	4.369	0.000	5.739	-2.36
%Vauconditionnel/EnsVConj	1.684	2.944	0.944	1.631	-2.40
expérimental	0.000	7.230	0.000	9.050	-2.48
présupposé	0.000	1.992	0.000	2.310	-2.68
NbMoymotsparphrase	18.009	22.432	3.001	5.096	-2.69
%virgules/EnsPonct	38.212	47.615	2.280	10.651	-2.74
%Pronrelatifs/EnsPron	22.053	26.846	2.767	4.986	-2.98
%VNegEnsV	5.584	8.264	0.969	2.729	-3.05
%desPh UNE Sub/EnsPh	5.053	8.686	1.530	3.348	-3.37
%PropsSubEnsProps	6.377	10.854	1.994	3.781	-3.68

### *TnT + lexique épistémique*

Morphosyntaxe et lexique épistémique					
Variable	Moyenne		Ecart-type		V. test
	Auteur	Corpus	Auteur	Corpus	
INT	0.351	0.038	0.353	0.135	7.18
description	6.000	1.098	6.394	2.556	5.95
PON:slash	4.873	1.329	3.065	1.913	5.75
PRO:disj:pp2sn	0.238	0.037	0.289	0.122	5.13
CON:spc:1st	10.129	4.104	9.203	3.660	5.11
explication	1.556	0.311	1.499	0.790	4.89
PRO:pp2sn	2.370	0.345	4.037	1.305	4.82
PON:int	2.039	0.882	0.580	0.789	4.55
PRO:pp2pl	2.017	0.532	2.164	1.049	4.39
proposer	5.222	1.770	4.076	2.445	4.38
PON:excl	0.867	0.202	1.256	0.498	4.15
PON:colon	9.188	5.921	1.013	2.501	4.05
VER:aux:fut	0.357	0.145	0.243	0.166	3.95

VER:fut	2.411	1.448	0.679	0.769	3.89
PRO:disj:pp1sn	0.643	0.195	0.559	0.362	3.85
PRO:clit:pp2pl	0.494	0.135	0.663	0.301	3.71
PRO:clit:pp2sn	0.353	0.086	0.522	0.234	3.54
montrer	4.222	1.672	3.083	2.474	3.20
PRO:poss:pp2pl	0.028	0.003	0.079	0.025	3.14
savoir	7.222	2.590	13.105	4.643	3.10
PRO:pp1sn	4.434	2.253	2.246	2.268	2.99
SYM	56.216	33.111	18.953	24.911	2.88
PREP:1st	0.934	0.586	0.405	0.414	2.62
phénomène	4.000	1.656	2.981	2.868	2.54
PRO:disj:pp2pl	0.119	0.035	0.189	0.106	2.45
comprendre	2.333	1.098	2.404	1.575	2.43
LS	34.613	23.572	9.310	14.335	2.39
ADV	80.062	76.185	3.399	5.111	2.36
CON:csq:1st	5.783	4.247	2.688	2.042	2.34
questionnement	0.000	4.369	0.000	5.739	-2.36
CON:ccl:1st	0.471	1.188	0.414	0.938	-2.37
CON:dis	9.438	13.784	1.312	5.648	-2.39
%classe12	0.999	1.367	0.136	0.463	-2.47
expérimental	0.000	7.230	0.000	9.050	-2.48
CON:ref:1st	1.963	3.195	0.754	1.514	-2.53
présupposé	0.000	1.992	0.000	2.310	-2.68
CON:cau:1st	1.282	2.909	1.050	1.878	-2.69
CON:exe:1st	0.837	2.300	0.477	1.625	-2.80
PRO:rel	21.014	25.587	2.617	4.939	-2.87
PON:comma	34.097	43.905	2.864	10.088	-3.02
ADV:neg:1st	12.928	18.297	3.176	4.789	-3.48

## Bergounioux

### Lexique

Lexique					
Mot	Pourcentage		Fréquence		V. test
	Auteur	Corpus	Auteur	Corpus	
psychologie	0.24	0.02	73	112	17.284
linguistique	0.70	0.19	212	1174	16.448
française	0.31	0.04	94	261	15.602
parole	0.40	0.09	123	546	14.268
la	6.15	4.43	1872	27842	14.232
nous	0.16	0.40	49	2505	-7.537
c	0.23	0.54	70	3387	-8.399
sémiotique	0.01	0.16	3	1026	-8.684
politique	0.00	0.15	1	963	-9.029
peut	0.10	0.37	30	2308	-9.274

### Cordial + lexique épistémique

Morphosyntaxe et lexique épistémique			
Variable	Moyenne	Ecart-type	V. test

	Auteur	Corpus	Auteur	Corpus	
%Vauconditionnelpassé/EnsVConj	0.958	0.172	0.463	0.283	6.95
démontrer	2.167	0.230	3.184	0.921	5.26
%Vauplus-que-parfait/EnsVConj	2.108	0.546	1.566	0.762	5.13
méthodologie	3.833	0.746	2.115	1.725	4.48
%Pronrelatifs/EnsPron	34.747	26.846	1.830	4.986	3.96
étudier	10.000	3.008	8.660	4.555	3.84
expérimentation	1.000	0.189	0.816	0.563	3.61
matériau	1.500	0.328	1.708	0.854	3.44
impliquer	6.500	2.598	5.795	3.034	3.22
%VauFutantérieur/EnsVConj	0.247	0.074	0.191	0.134	3.21
typologie	7.167	2.525	4.879	3.620	3.21
%VàImpft/EnsVConj	9.555	4.155	5.803	4.240	3.19
NbMoyNparProp	3.683	3.016	0.469	0.534	3.12
%N/totalmotsSfiants	61.103	55.753	2.670	4.338	3.09
%PropsCOI	37.432	31.121	3.998	5.471	2.89
%VauSubjprésent/EnsVConj	3.077	1.974	1.419	1.007	2.74
présupposé	4.500	1.992	2.363	2.310	2.72
NbMoymotsparProp	12.482	10.812	1.316	1.572	2.66
%VinTransEnsV	10.960	8.346	1.352	2.473	2.64
réflexion	5.000	1.738	4.163	3.208	2.54
NbMoymotsparparagraphe	75.650	53.743	20.242	21.869	2.51
%VauSubjpassé/EnsVConj	0.272	0.113	0.227	0.161	2.46
%Vaupassésimple/EnsVConj	4.083	1.637	3.168	2.525	2.42
NbMoymotsparphrase	27.355	22.432	4.687	5.096	2.42
%VpronominiauxEnsV	12.512	9.528	2.056	3.148	2.37
%d'adV/totalmotsSfiants	9.398	12.150	1.249	2.709	-2.54
%Vauprésent/EnsVConj	63.593	78.326	8.666	7.886	-4.67
%AdjDem/EnsDET	2.182	4.951	0.573	1.449	-4.78

### *TnT + lexique épistémique*

Morphosyntaxe et lexique épistémique					
Variable	Moyenne		Ecart-type		V. test
	Auteur	Corpus	Auteur	Corpus	
VER:aux:cond	0.652	0.161	0.280	0.207	5.93
CON:add:1st	66.142	45.226	8.087	8.931	5.86
démontrer	2.167	0.230	3.184	0.921	5.26
VER:aux:impf	1.678	0.401	1.749	0.609	5.24
SYM:ling	6.063	0.994	5.432	2.500	5.07
VER:mod:pper	0.812	0.254	0.452	0.290	4.81
méthodologie	3.833	0.746	2.115	1.725	4.48
PRO:rel	34.318	25.587	1.578	4.939	4.42
%classe7	24.042	20.393	0.970	2.109	4.33
VER:pper	26.613	18.752	2.452	4.620	4.26
NUM:dat	40.633	18.398	12.043	13.229	4.20
VER:impf	5.498	1.914	3.317	2.162	4.15
NP	13.773	6.306	6.114	4.546	4.11
VER:mod:simp	0.105	0.013	0.151	0.059	3.90
étudier	10.000	3.008	8.660	4.555	3.84
VER:aux:pper	0.952	0.432	0.397	0.345	3.77
%classe8	16.655	14.462	0.791	1.476	3.72
VER:simp	1.623	0.427	1.411	0.820	3.65
expérimentation	1.000	0.189	0.816	0.563	3.61



matériau	1.500	0.328	1.708	0.854	3.44
VER:aux:simp	0.172	0.039	0.161	0.099	3.36
VER:cond	1.902	1.063	0.714	0.645	3.25
impliquer	6.500	2.598	5.795	3.034	3.22
typologie	7.167	2.525	4.879	3.620	3.21
PRO:pp3msn	11.598	7.917	2.976	2.910	3.17
VER:aux:subp	0.530	0.258	0.281	0.228	2.98
PRO:clit	11.715	9.204	0.884	2.234	2.81
présupposé	4.500	1.992	2.363	2.310	2.72
DET:poss:pp3sn	4.895	3.320	1.164	1.474	2.67
réflexion	5.000	1.738	4.163	3.208	2.54
%classe6	13.872	12.487	1.355	1.472	2.35
PRO:dem	10.270	14.056	1.247	4.061	-2.33
CON:spc:1st	0.647	4.104	0.624	3.660	-2.36
ADV	71.252	76.185	5.538	5.111	-2.41
NC:sg	66.325	72.420	7.201	6.153	-2.48
PUL	0.008	0.052	0.019	0.043	-2.50
VER:fut	0.495	1.448	0.298	0.769	-3.10
VER:inf	12.420	17.553	1.668	4.100	-3.13
%classe2	6.063	7.898	0.577	1.427	-3.22
PRO:pp3sn	3.362	8.145	1.281	3.557	-3.36
CON:csq:1st	1.380	4.247	1.051	2.042	-3.51
VER:mod:pres	1.705	4.207	0.464	1.644	-3.81
VER:pres	28.907	39.221	5.593	6.279	-4.11
CON:jus:1st	1.388	5.907	0.845	2.739	-4.13
DET:dem	2.210	6.269	0.726	2.044	-4.97

## Calvet

### Lexique

Lexique					
Mot	Pourcentage		Fréquence		V. test
	Auteur	Corpus	Auteur	Corpus	
langues	0.97	0.10	366	638	35.008
langue	1.08	0.21	408	1316	27.975
français	0.79	0.13	299	838	25.672
anglais	0.34	0.03	127	171	22.874
pays	0.41	0.04	154	268	22.679
contexte	0.01	0.12	3	743	-8.123
b	0.00	0.10	0	639	-8.532
sujet	0.01	0.15	3	970	-9.598
sémiotique	0.00	0.16	0	1026	-10.959
discours	0.02	0.35	7	2193	-14.545

### *Cordial + lexique épistémique*

Morphosyntaxe et lexique épistémique					
Variable	Moyenne		Ecart-type		V. test
	Auteur	Corpus	Auteur	Corpus	
tendance	3.125	0.500	2.147	1.196	6.40
montrer	6.625	1.672	3.120	2.474	5.83
principe	3.625	0.648	4.029	1.509	5.75
%V_AVOIR_EnsV	7.531	4.493	1.240	1.836	4.82
constituer	8.125	2.934	5.441	3.724	4.06
travail	5.125	1.992	4.649	2.407	3.79
proposer	4.750	1.770	2.947	2.445	3.55
NbMoymotsparphrase	27.865	22.432	1.897	5.096	3.11
problème	6.250	2.680	3.419	3.474	2.99
choix	3.000	1.238	3.279	1.838	2.79
%VàImpft/EnsVConj	8.108	4.155	3.091	4.240	2.72
recherche	1.750	0.582	1.785	1.266	2.69
terme	5.875	2.500	3.295	3.660	2.69
%VauSubjpassé/EnsVConj	0.259	0.113	0.239	0.161	2.63
%PronPersls1ePersPIEnsPronPersls	12.540	6.776	3.778	6.434	2.61
NbMoyPhparparagraphe	3.094	1.622	2.386	1.679	2.55
exemple	12.875	5.631	5.904	8.640	2.44
%V3ePersPIEnsVConj	23.570	18.873	7.798	5.698	2.40
système	4.375	1.574	3.426	3.407	2.40
présupposé	0.000	1.992	0.000	2.310	-2.51
%Ptsvirgules/EnsPonct	0.303	1.675	0.251	1.221	-3.28

### *TnT + lexique épistémique*

Morphosyntaxe et lexique épistémique					
Variable	Moyenne		Ecart-type		V. test
	Auteur	Corpus	Auteur	Corpus	
tendance	3.125	0.500	2.147	1.196	6.40
montrer	6.625	1.672	3.120	2.474	5.83
principe	3.625	0.648	4.029	1.509	5.75
CON:ces:1st	2.996	1.259	1.272	1.058	4.78
VER:impf	5.119	1.914	1.808	2.162	4.32
constituer	8.125	2.934	5.441	3.724	4.06
VER:mod:impf	0.448	0.147	0.460	0.219	3.99
NP	12.518	6.306	10.163	4.546	3.98
travail	5.125	1.992	4.649	2.407	3.79
PRO:clit:pp1pl	1.749	0.775	0.603	0.778	3.65
proposer	4.750	1.770	2.947	2.445	3.55
NUM:car	54.094	36.840	8.229	14.228	3.53
problème	6.250	2.680	3.419	3.474	2.99
PRO:poss:pp2sn	0.045	0.005	0.119	0.039	2.96
VER:aux:pper	0.777	0.432	0.217	0.345	2.92
PRO:disj:pp3pl	1.019	0.540	0.770	0.482	2.90
SIG	21.693	9.415	21.799	12.737	2.81
choix	3.000	1.238	3.279	1.838	2.79
recherche	1.750	0.582	1.785	1.266	2.69
terme	5.875	2.500	3.295	3.660	2.69
CON:tem:1st	3.050	1.859	1.119	1.324	2.62

VER:mod:cond	1.102	0.644	0.544	0.535	2.50
exemple	12.875	5.631	5.904	8.640	2.44
système	4.375	1.574	3.426	3.407	2.40
CON:pre:1st	0.056	0.582	0.100	0.610	-2.51
présupposé	0.000	1.992	0.000	2.310	-2.51
LS	9.282	23.572	6.403	14.335	-2.90
PON:Pvirg	0.276	1.416	0.212	1.123	-2.96
NC:sg	65.475	72.420	8.394	6.153	-3.29

## Combettes

### Lexique

Lexique					
Mot	Pourcentage		Fréquence		V. test
	Auteur	Corpus	Auteur	Corpus	
syntagme	0.30	0.04	160	235	22.429
proposition	0.28	0.04	152	262	20.205
grammaticalisation	0.18	0.02	98	112	19.901
verbe	0.32	0.06	175	386	19.094
topique	0.18	0.02	95	141	17.163
a	0.27	0.64	147	4031	-12.496
politique	0.00	0.15	1	963	-12.527
sémiotique	0.00	0.16	0	1026	-13.312
discours	0.07	0.35	39	2193	-13.589
et	1.41	2.72	761	17112	-21.315

### Cordial + lexique épistémique

Morphosyntaxe et lexique épistémique					
Variable	Moyenne		Ecart-type		V. test
	Auteur	Corpus	Auteur	Corpus	
considérer	10.417	2.418	3.402	3.421	8.50
examiner	3.917	0.926	1.847	1.543	7.04
exemple	21.667	5.631	9.885	8.640	6.74
constater	3.750	0.943	1.785	1.522	6.70
%Vauconditionnel/EnsVConj	5.934	2.944	1.611	1.631	6.66
unité	7.583	1.533	6.639	3.720	5.91
système	7.083	1.574	4.663	3.407	5.88
phénomène	6.000	1.656	5.099	2.868	5.50
%Vaupasséantérieur/EnsVConj	0.268	0.037	0.430	0.166	5.05
constituer	7.750	2.934	3.677	3.724	4.70
%Vaupassésimple/EnsVConj	4.718	1.637	3.693	2.525	4.43
%Ptsvirgules/EnsPonct	3.123	1.675	0.967	1.221	4.31
%VinTransEnsV	11.234	8.346	1.296	2.473	4.24
commentaire	5.250	1.303	6.685	3.459	4.15
tendance	1.833	0.500	0.986	1.196	4.05
relever	6.917	2.869	3.095	3.744	3.93
%desPh_UNE_Sub/EnsPh	11.949	8.686	2.545	3.348	3.54
%Pronrelatifs/EnsPron	31.656	26.846	3.040	4.986	3.50
%PropRelEnsProps	4.398	2.550	2.707	1.935	3.47

fonction	7.750	3.779	6.043	4.232	3.41
%AdjIndef/EnsDET	3.865	2.654	2.362	1.293	3.40
%VauSubjplus-que-parfait/EnsVCon	0.134	0.037	0.161	0.105	3.34
%VCOIObligEnsV	15.462	13.517	2.209	2.276	3.10
%AdjDem/EnsDET	6.163	4.951	0.980	1.449	3.04
montrer	3.667	1.672	1.599	2.474	2.93
%PropsSubEnsProps	13.895	10.854	3.050	3.781	2.92
valeur	14.917	5.623	18.652	11.664	2.89
%d'adV/totalmotsSfiants	14.283	12.150	0.954	2.709	2.86
étude	5.000	2.467	2.828	3.270	2.81
%VauSubjImpft/EnsVConj	0.193	0.073	0.212	0.156	2.77
présenter	6.417	3.705	2.691	3.748	2.63
%Ptsdesuspension/EnsPonct	2.445	1.375	1.769	1.542	2.52
%AdjPoss3ePersSg/EnsAdjPoss	38.889	16.797	46.812	34.343	2.34
<b>-----</b>					
%Pronindéfinis/EnsPron	3.301	4.749	1.448	2.257	-2.33
%dePhInt/EnsPh	1.189	2.815	0.798	2.532	-2.33
définir	0.417	2.016	0.640	2.479	-2.34
Nbmoyen_AdjparProp	0.738	0.867	0.143	0.200	-2.35
NbMoymotsparProp	9.794	10.812	0.696	1.572	-2.35
%AdjPoss/EnsDET	3.346	4.765	1.005	2.185	-2.36
NbMoyNparProp	2.655	3.016	0.267	0.534	-2.45
%Ptsd'interrogation/EnsPonct	0.389	1.270	0.279	1.295	-2.47
%Pts/EnsPonct	13.367	16.861	2.980	5.095	-2.49
constat	0.000	2.156	0.000	2.903	-2.70
présupposé	0.250	1.992	0.433	2.310	-2.74
questionnement	0.000	4.369	0.000	5.739	-2.77
%AdjCard/EnsDET	0.989	2.321	0.638	1.729	-2.80
%PropsCOD	28.641	31.964	2.070	4.300	-2.81
théorie	0.417	6.025	0.640	7.249	-2.81
expérimental	0.000	7.230	0.000	9.050	-2.90
%VCODObligEnsV	32.030	35.402	2.288	4.066	-3.01
%PropIndepEnsProps	79.516	84.755	4.481	5.257	-3.62
%V_ETRE_EnsV	16.834	21.550	2.893	4.635	-3.70
%VTransindirectsEnsV	45.835	50.103	3.377	4.078	-3.80

### *TnT + lexique épistémique*

Morphosyntaxe et lexique épistémique					
Variable	Moyenne		Ecart-type		V. test
	Auteur	Corpus	Auteur	Corpus	
considérer	10.417	2.418	3.402	3.421	8.50
examiner	3.917	0.926	1.847	1.543	7.04
FGW	5.878	0.978	4.717	2.636	6.75
exemple	21.667	5.631	9.885	8.640	6.74
constater	3.750	0.943	1.785	1.522	6.70
CON:jus:1st	10.827	5.907	1.865	2.739	6.53
VER:aux:inf	2.156	1.047	0.560	0.630	6.39
VER:mod:cond	1.567	0.644	0.691	0.535	6.28
unité	7.583	1.533	6.639	3.720	5.91
système	7.083	1.574	4.663	3.407	5.88
VER:simp	1.716	0.427	1.374	0.820	5.71
VER:cond	2.046	1.063	0.563	0.645	5.53
phénomène	6.000	1.656	5.099	2.868	5.50
VER:mod:pres	6.607	4.207	1.022	1.644	5.30
PON:susp	3.351	1.295	1.389	1.447	5.16

CON:dou:1st	3.275	1.535	1.256	1.273	4.97
constituer	7.750	2.934	3.677	3.724	4.70
PRO:pp3isn	9.093	6.386	0.718	2.168	4.54
PON:Pvirg	2.756	1.416	0.573	1.123	4.34
commentaire	5.250	1.303	6.685	3.459	4.15
tendance	1.833	0.500	0.986	1.196	4.05
relever	6.917	2.869	3.095	3.744	3.93
LS	38.249	23.572	6.580	14.335	3.72
%classe12	1.838	1.367	0.468	0.463	3.70
VER:mod:inf	0.377	0.195	0.213	0.185	3.59
fonction	7.750	3.779	6.043	4.232	3.41
VER:aux:simp	0.128	0.039	0.164	0.099	3.26
DET:indef	26.282	22.917	3.008	4.002	3.05
DET:dem	7.970	6.269	1.350	2.044	3.02
VER:subp	0.987	0.629	0.701	0.430	3.02
montrer	3.667	1.672	1.599	2.474	2.93
valeur	14.917	5.623	18.652	11.664	2.89
PRO:poss:pp3pl	0.033	0.007	0.076	0.033	2.89
étude	5.000	2.467	2.828	3.270	2.81
VER:aux:parpres	0.316	0.160	0.235	0.205	2.76
PRO:rel	29.324	25.587	3.047	4.939	2.75
PRO:poss:pp1sn	0.151	0.028	0.454	0.164	2.72
présenter	6.417	3.705	2.691	3.748	2.63
PRO:poss:pp3sn	0.205	0.075	0.336	0.184	2.58
VER:fut	1.993	1.448	0.599	0.769	2.58
DET:poss:pp2pl	0.118	0.044	0.163	0.105	2.58
VER:aux:cond	0.297	0.161	0.298	0.207	2.38
<b>-----</b>					
DET:def	62.438	65.345	2.850	4.514	-2.34
définir	0.417	2.016	0.640	2.479	-2.34
PON:int	0.355	0.882	0.245	0.789	-2.43
CON:ref:1st	2.165	3.195	1.062	1.514	-2.47
NUM:car	27.059	36.840	10.159	14.228	-2.50
VER:pres	34.883	39.221	3.634	6.279	-2.51
constat	0.000	2.156	0.000	2.903	-2.70
présupposé	0.250	1.992	0.433	2.310	-2.74
DET:poss:pp3pl	0.689	1.258	0.337	0.753	-2.74
PRO:clit	7.508	9.204	1.715	2.234	-2.76
questionnement	0.000	4.369	0.000	5.739	-2.77
théorie	0.417	6.025	0.640	7.249	-2.81
expérimental	0.000	7.230	0.000	9.050	-2.90
PRO:disj:pp3sn	0.751	1.549	0.252	0.980	-2.96
PON:dot	11.084	15.710	2.247	5.151	-3.26
CON:cau:1st	1.197	2.909	0.916	1.878	-3.31
DET:poss:pp3sn	1.958	3.320	0.728	1.474	-3.36
PREF	0.067	0.248	0.048	0.191	-3.45
%classe6	11.003	12.487	0.508	1.472	-3.66
%classe7	18.002	20.393	1.365	2.109	-4.12

## François

### *Lexique*

Lexique					
Mot	Pourcentage		Fréquence		V. test
	Auteur	Corpus	Auteur	Corpus	
0	0.46	0.03	158	179	28.068
x	0.80	0.14	275	884	23.890
prédications	0.35	0.02	121	155	23.260
ex	0.44	0.04	151	259	23.152
prédication	0.51	0.06	177	396	22.530
formes	0.00	0.14	1	853	-9.054
qu	0.34	0.72	118	4516	-9.335
discours	0.08	0.35	28	2193	-10.243
sémiotique	0.00	0.16	0	1026	-10.445
qui	0.71	1.36	243	8568	-11.796

### *Cordial + lexique épistémique*

Morphosyntaxe et lexique épistémique					
Variable	Moyenne		Ecart-type		V. test
	Auteur	Corpus	Auteur	Corpus	
Nbttotalparagraphe	557.250	169.820	399.024	157.865	7.15
propriété	19.750	3.230	21.862	7.468	6.45
donnée	3.250	0.533	3.072	1.269	6.24
%CrochAccol/EnsPonct	6.102	1.534	3.402	2.236	5.95
%Vaupassécomposé/EnsVConj	12.320	5.503	8.729	3.623	5.48
%motsSfiant/Ensmots	54.588	50.011	4.182	2.517	5.30
%N/totalmotsSfiant	63.609	55.753	4.378	4.338	5.28
NbttotalPh	658.000	345.033	430.872	178.560	5.11
méthode	2.000	0.377	2.121	0.978	4.84
%Ptsd'interrogation/EnsPonct	3.369	1.270	2.551	1.295	4.72
interprétation	8.250	2.156	9.614	3.822	4.65
%AdjCard/EnsDET	5.003	2.321	2.361	1.729	4.52
%PropIndepEnsProps	92.590	84.755	2.962	5.257	4.34
viser	2.625	0.746	3.533	1.394	3.93
présenter	8.125	3.705	5.441	3.748	3.44
voir	27.875	6.328	56.197	18.375	3.42
valeur	18.125	5.623	17.077	11.664	3.12
%parenthèses/EnsPonct	23.574	17.078	6.173	6.187	3.06
recherche	1.875	0.582	2.147	1.266	2.98
désigner	6.250	2.820	4.710	3.483	2.87
%Pronindéfinis/EnsPron	6.824	4.749	3.907	2.257	2.68
%V3ePersSgEnsVConj	79.736	74.143	5.693	6.246	2.61
comprendre	2.500	1.098	2.291	1.575	2.59
conceptuel	5.625	1.943	11.554	4.249	2.53
fonction	7.375	3.779	6.999	4.232	2.48
condition	3.125	1.344	4.136	2.164	2.40
observation	2.375	0.852	2.781	1.854	2.39
présupposé	0.125	1.992	0.331	2.310	-2.35
%AdjDem/EnsDET	3.682	4.951	0.754	1.449	-2.55
Nbmoyen_AdjparProp	0.686	0.867	0.209	0.200	-2.63
%PropRelEnsProps	0.666	2.550	0.333	1.935	-2.84
%Vauconditionnel/EnsVConj	1.237	2.944	0.736	1.631	-3.05
%PropAttSuj	14.821	18.866	3.894	3.164	-3.73
%d'adV/totalmotsSfiant	8.634	12.150	1.247	2.709	-3.78

%deV/totalmotsSfiants	13.155	16.162	2.114	2.262	-3.87
%PropsSubEnsProps	5.715	10.854	3.093	3.781	-3.96
NbMoymotsparProp	8.600	10.812	1.573	1.572	-4.10
NbMoyPropsparphrase	1.240	2.378	0.185	0.788	-4.21
%desPh UNE Sub/EnsPh	3.552	8.686	1.894	3.348	-4.47
%PropsCOD	24.974	31.964	6.060	4.300	-4.74
NbMoymotsparparagraphe	15.768	53.743	4.974	21.869	-5.06
%mots-outils/Ensmots	45.412	49.989	4.182	2.517	-5.30
Nbmoyen_adVparProp	0.400	0.647	0.075	0.129	-5.56
NbMoymotsparphrase	12.538	22.432	3.307	5.096	-5.66
NbMoyPronparProp	0.403	0.863	0.163	0.188	-7.13
%virgules/EnsPonct	20.884	47.615	4.676	10.651	-7.31
%PropsCOI	17.319	31.121	4.454	5.471	-7.35

### *TnT + lexique épistémique*

Morphosyntaxe et lexique épistémique					
Variable	Moyenne		Ecart-type		V. test
	Auteur	Corpus	Auteur	Corpus	
%classe1	5.550	1.307	1.727	1.397	8.85
PON:accol	1.691	0.119	1.718	0.608	7.54
VER:aux:pres	10.705	5.477	4.558	2.330	6.54
%classe14	4.820	2.081	2.261	1.230	6.49
propriété	19.750	3.230	21.862	7.468	6.45
donnée	3.250	0.533	3.072	1.269	6.24
PON:croch	7.042	1.578	4.171	3.041	5.24
méthode	2.000	0.377	2.121	0.978	4.84
PON:par	27.379	16.175	4.119	6.790	4.81
interprétation	8.250	2.156	9.614	3.822	4.65
PRO:refl	19.190	12.476	12.973	4.699	4.16
PON:slash	3.955	1.329	3.136	1.913	4.00
viser	2.625	0.746	3.533	1.394	3.93
SYM	63.109	33.111	18.007	24.911	3.51
PRO:indef	5.141	2.913	2.844	1.888	3.44
présenter	8.125	3.705	5.441	3.748	3.44
voir	27.875	6.328	56.197	18.375	3.42
VER:inf	22.006	17.553	10.943	4.100	3.17
valeur	18.125	5.623	17.077	11.664	3.12
ADJ:sg	76.959	68.964	4.999	7.688	3.03
recherche	1.875	0.582	2.147	1.266	2.98
CON:tem:1st	3.207	1.859	2.322	1.324	2.97
VER:aux:parpres	0.364	0.160	0.401	0.205	2.90
désigner	6.250	2.820	4.710	3.483	2.87
comprendre	2.500	1.098	2.291	1.575	2.59
conceptuel	5.625	1.943	11.554	4.249	2.53
PREP	83.143	80.138	3.371	3.502	2.50
fonction	7.375	3.779	6.999	4.232	2.48
condition	3.125	1.344	4.136	2.164	2.40
observation	2.375	0.852	2.781	1.854	2.39
ADV:int	2.018	3.715	0.801	2.105	-2.35
présupposé	0.125	1.992	0.331	2.310	-2.35
NUM:dat	7.603	18.398	7.245	13.229	-2.38
%classe6	11.279	12.487	1.250	1.472	-2.39
VER:subp	0.275	0.629	0.168	0.430	-2.40
ADJ:pl	22.534	29.949	4.859	7.627	-2.83

VER:mod:pres	2.582	4.207	0.952	1.644	-2.88
ABR	16.450	42.104	10.051	25.740	-2.90
CON:ces:1st	0.168	1.259	0.304	1.058	-3.01
PRO:rel	20.470	25.587	4.461	4.939	-3.02
VER:cond	0.377	1.063	0.104	0.645	-3.10
%classe2	6.173	7.898	0.952	1.427	-3.52
CON:jus:1st	2.074	5.907	0.935	2.739	-4.08
VER:pres	29.505	39.221	3.354	6.279	-4.51
%classe4	3.799	6.237	1.501	1.408	-5.05
PON:comma	25.935	43.905	6.147	10.088	-5.19

## Kerbrat

### Lexique

Lexique					
Mot	Pourcentage		Fréquence		V. test
	Auteur	Corpus	Auteur	Corpus	
<b>politesse</b>	0.35	0.02	144	149	27.147
<b>vous</b>	0.54	0.08	223	488	24.031
	0.52	0.09	213	540	21.910
<b>femmes</b>	0.18	0.02	74	118	15.846
<b>interaction</b>	0.17	0.03	72	167	13.241
<b>nous</b>	0.15	0.40	60	2505	-9.547
<b>du</b>	1.15	1.74	476	10956	-9.973
<b>politique</b>	0.00	0.15	2	963	-10.423
<b>sémiotique</b>	0.00	0.16	2	1026	-10.811
<b>la</b>	3.35	4.43	1380	27842	-11.452

### Cordial + lexique épistémique

Morphosyntaxe et lexique épistémique					
Variable	Moyenne		Ecart-type		V. test
	Auteur	Corpus	Auteur	Corpus	
%PronPersls2ePersPIEnsPronPersls	6.536	1.658	4.765	2.764	5.14
%deux-points/EnsPonct	10.189	6.068	1.498	2.392	5.02
%V2ePersPIEnsVConj	3.382	0.928	3.079	1.589	4.50
questionnement	12.375	4.369	8.731	5.739	4.07
%AdjPoss1ePersPI/EnsAdjPoss	37.500	6.250	48.412	23.590	3.86
expérimental	19.000	7.230	7.746	9.050	3.79
%d'adV/totalmotsSfians	15.499	12.150	1.749	2.709	3.60
%AdjIndef/EnsDET	4.214	2.654	1.351	1.293	3.52
%Vàimpératif/EnsVConj	2.709	1.503	0.803	1.036	3.39
problématique	6.750	2.574	6.339	3.720	3.27
%AdjDem/EnsDET	6.551	4.951	0.959	1.449	3.22
%V1ePersSgEnsVConj	5.863	2.773	3.564	2.866	3.14
schéma	12.500	5.828	9.474	6.531	2.98
%Ptsvirgules/EnsPonct	2.904	1.675	1.413	1.221	2.93
%parenthèses/EnsPonct	23.280	17.078	4.935	6.187	2.92
%dePhExcl/EnsPh	2.615	0.854	3.000	1.769	2.90
%Ptsd'exclamation/EnsPonct	1.315	0.423	1.485	0.912	2.85
%PronPersls1ePersSgEnsPronPersls	11.604	6.329	5.547	5.448	2.82
indiquer	2.875	1.361	2.027	1.650	2.67



%V_ETRE_EnsV	25.565	21.550	3.123	4.635	2.52
%PropsCoordEnsProps	0.246	0.075	0.353	0.199	2.51
présupposé	3.875	1.992	1.269	2.310	2.38
NbMoymotsparProp	9.456	10.812	0.950	1.572	-2.51
%N/totalmotsSfiants	51.829	55.753	1.576	4.338	-2.64
NbMoyNparProp	2.506	3.016	0.347	0.534	-2.78
%Pts/EnsPonct	11.809	16.861	3.356	5.095	-2.89

### *TnT + lexique épistémique*

Morphosyntaxe et lexique épistémique					
Variable	Moyenne		Ecart-type		V. test
	Auteur	Corpus	Auteur	Corpus	
PON:guil	22.045	9.841	3.925	6.288	5.66
questionnement	12.375	4.369	8.731	5.739	4.07
expérimental	19.000	7.230	7.746	9.050	3.79
PON:colon	9.155	5.921	2.143	2.501	3.77
DET:poss:pp2pl	0.165	0.044	0.187	0.105	3.36
%classe2	9.504	7.898	0.476	1.427	3.28
problématique	6.750	2.574	6.339	3.720	3.27
DET:dem	8.480	6.269	1.590	2.044	3.15
LS	39.070	23.572	12.479	14.335	3.15
PON:Pvirg	2.616	1.416	0.887	1.123	3.12
VER:imp	1.096	0.484	0.519	0.591	3.02
schéma	12.500	5.828	9.474	6.531	2.98
CON:ref:1st	4.735	3.195	1.363	1.514	2.96
PREP	83.511	80.138	2.990	3.502	2.81
%classe12	1.804	1.367	0.254	0.463	2.75
VER:mod:pres	5.744	4.207	1.018	1.644	2.72
indiquer	2.875	1.361	2.027	1.650	2.67
ADV	80.806	76.185	2.500	5.111	2.64
présupposé	3.875	1.992	1.269	2.310	2.38
PREP:1st	0.237	0.586	0.166	0.414	-2.45
%classe8	13.161	14.462	0.709	1.476	-2.57
PON:comma	34.301	43.905	2.430	10.088	-2.77
VER:pper	14.245	18.752	1.487	4.620	-2.84
PON:dot	9.327	15.710	2.021	5.151	-3.61

## Kleiber

### *Lexique*

Lexique					
Mot	Pourcentage		Fréquence		V. test
	Auteur	Corpus	Auteur	Corpus	
noms	0.37	0.04	171	278	23.179
ça	0.44	0.07	203	416	22.700
sens	0.87	0.27	403	1708	20.942
dénomination	0.33	0.04	151	271	20.815
nom	0.47	0.10	218	612	20.027

<b>politique</b>	0.00	0.15	0	963	-11.843
<b>discours</b>	0.08	0.35	35	2193	-12.297
<b>dans</b>	1.03	1.73	475	10869	-12.844
<b>de</b>	5.08	6.65	2345	41831	-14.593
<b>la</b>	3.03	4.43	1401	27842	-15.899

### *Cordial + lexique épistémique*

Morphosyntaxe et lexique épistémique					
Variable	Moyenne		Ecart-type		V. test
	Auteur	Corpus	Auteur	Corpus	
%V AVOIR EnsV	7.881	4.493	1.716	1.836	5.38
%d'adV/totalmotsSfiants	16.833	12.150	2.188	2.709	5.04
%VNegEnsV	12.846	8.264	2.820	2.729	4.89
%desPh UNE Sub/EnsPh	13.913	8.686	2.052	3.348	4.55
%PropsSubEnsProps	16.607	10.854	2.963	3.781	4.43
question	9.125	3.213	6.864	4.189	4.11
%AdjInterrog/EnsDET	0.916	0.344	0.210	0.409	4.08
notion	9.375	3.664	4.923	4.160	4.00
%deV/totalmotsSfiants	19.157	16.162	1.275	2.262	3.86
idée	5.625	2.123	5.171	2.757	3.70
discussion	4.125	1.115	7.149	2.389	3.67
%ArtIndef/EnsDET	21.137	16.591	3.871	3.664	3.62
caractérisation	3.500	0.992	3.640	2.078	3.52
%V ETRE EnsV	26.950	21.550	2.136	4.635	3.40
Nbmoyen_adVparProp	0.796	0.647	0.121	0.129	3.36
%PropAttSuj	22.305	18.866	2.231	3.164	3.17
observable	5.875	1.344	14.040	4.292	3.08
choisir	2.750	0.795	4.054	1.855	3.07
approche	12.250	4.361	20.903	7.524	3.06
conséquence	5.250	2.156	6.741	2.989	3.02
%PropPPIEnsProps	0.044	0.007	0.077	0.036	2.99
%V1ePersPIEnsVConj	5.929	3.137	1.768	2.918	2.79
%deProndémonstratifs/EnsPron	16.285	12.896	6.040	3.751	2.63
%AdjIndef/EnsDET	3.691	2.654	1.157	1.293	2.34
NbMoymotsparProp	9.521	10.812	0.576	1.572	-2.39
%Ptsvirgules/EnsPonct	0.608	1.675	0.427	1.221	-2.55
%ArtDef/EnsART	64.601	69.928	3.586	4.485	-3.46
%ArtDef/EnsDET	78.863	83.409	3.871	3.664	-3.62
NbMoyNparProp	2.287	3.016	0.231	0.534	-3.97
%PropIndepEnsProps	76.870	84.755	3.444	5.257	-4.37
%N/totalmotsSfiants	48.313	55.753	2.380	4.338	-5.00

### *TnT + lexique épistémique*

Morphosyntaxe et lexique épistémique					
Variable	Moyenne		Ecart-type		V. test
	Auteur	Corpus	Auteur	Corpus	
%classe2	10.741	7.898	1.088	1.427	5.81
%classe5	12.954	10.839	0.522	1.123	5.49
%classe12	2.205	1.367	0.210	0.463	5.28
CON:cau:1st	6.039	2.909	2.015	1.878	4.86

PON:cote	1.906	0.285	2.323	1.019	4.63
question	9.125	3.213	6.864	4.189	4.11
DET:poss:pp1pl	1.390	0.498	0.577	0.639	4.07
%classe4	8.191	6.237	0.913	1.408	4.04
notion	9.375	3.664	4.923	4.160	4.00
idée	5.625	2.123	5.171	2.757	3.70
discussion	4.125	1.115	7.149	2.389	3.67
CON:dis	20.754	13.784	7.760	5.648	3.60
caractérisation	3.500	0.992	3.640	2.078	3.52
CON:pre:1st	1.316	0.582	0.693	0.610	3.51
CON:cer:1st	0.701	0.264	0.286	0.391	3.26
observable	5.875	1.344	14.040	4.292	3.08
choisir	2.750	0.795	4.054	1.855	3.07
ADV:neg:1st	23.321	18.297	3.635	4.789	3.06
approche	12.250	4.361	20.903	7.524	3.06
VER:subp	1.079	0.629	0.319	0.430	3.04
conséquence	5.250	2.156	6.741	2.989	3.02
DET:indef	26.956	22.917	4.343	4.002	2.94
PRO:clit:pp1pl	1.512	0.775	0.990	0.778	2.76
VER:mod:inf	0.366	0.195	0.259	0.185	2.70
PRO:clit	11.160	9.204	1.666	2.234	2.55
PRO:pp3isn	8.257	6.386	2.052	2.168	2.51
VER:fut	2.075	1.448	0.608	0.769	2.38
<b> </b>					
PON:Pvirg	0.411	1.416	0.421	1.123	-2.61
VER:parpres	2.786	4.668	0.681	2.000	-2.74
PRO:indef	1.106	2.913	0.588	1.888	-2.79
%classe11	11.714	13.883	0.684	2.212	-2.86
ADV	70.724	76.185	3.893	5.111	-3.11
CON:add:1st	34.870	45.226	5.938	8.931	-3.38
%classe8	12.689	14.462	0.295	1.476	-3.50
%classe7	17.591	20.393	1.175	2.109	-3.87

## Lamizet

### Lexique

Lexique					
Mot	Pourcentage		Fréquence		V. test
	Auteur	Corpus	Auteur	Corpus	
sémiotique	1.11	0.16	715	1026	45.481
politique	1.04	0.15	672	963	44.119
symbolique	0.46	0.05	294	330	33.417
espace	0.73	0.13	468	837	32.607
public	0.45	0.06	290	386	30.113
<b> </b>					
plus	0.19	0.50	120	3118	-13.334
x	0.00	0.14	0	884	-13.585
1	0.01	0.19	5	1172	-14.399
pas	0.27	0.68	177	4281	-14.838
2	0.01	0.20	5	1276	-15.131

## *Cordial + lexique épistémique*

Morphosyntaxe et lexique épistémique					
Variable	Moyenne		Ecart-type		V. test
	Auteur	Corpus	Auteur	Corpus	
résultat	19.625	3.213	14.238	6.231	7.68
critère	13.750	3.492	9.404	4.914	6.08
investigation	18.750	3.705	16.238	7.445	5.89
définition	14.125	2.811	14.426	5.692	5.79
%virgules/EnsPonct	68.695	47.615	1.887	10.651	5.77
%mots-outils/Ensmots	54.803	49.989	1.159	2.517	5.57
systématique	6.000	1.393	7.314	2.616	5.13
NbMoymotsparparagraphe	88.974	53.743	11.263	21.869	4.69
conception	13.625	3.189	14.916	6.563	4.63
%Vauprésent/EnsVConj	90.802	78.326	3.141	7.886	4.61
NbMoymotsparphrase	30.337	22.432	2.299	5.096	4.52
classification	3.500	1.000	4.183	1.779	4.10
%PropsCOI	38.572	31.121	2.644	5.471	3.97
concept	8.625	2.803	5.499	4.360	3.89
approche	13.750	4.361	10.733	7.524	3.64
matériau	1.375	0.328	1.798	0.854	3.57
NbMoymotsparProp	12.650	10.812	0.962	1.572	3.41
%ArtDef/EnsART	75.077	69.928	2.664	4.485	3.35
questionnement	10.750	4.369	9.614	5.739	3.24
%VCODObligEnsV	39.851	35.402	5.580	4.066	3.19
NbMoyPronparProp	1.064	0.863	0.042	0.188	3.12
%Pts/EnsPonct	22.052	16.861	2.016	5.095	2.97
%AdjPoss3ePersPl/EnsAdjPoss	16.666	2.732	33.333	14.530	2.79
taxinomie	1.875	0.393	4.594	1.592	2.71
NbMoyPropsparphrase	3.075	2.378	0.462	0.788	2.58
%PropIndepEnsProps	89.259	84.755	2.045	5.257	2.50
conséquence	4.625	2.156	2.446	2.989	2.41
déterminer	5.125	2.484	4.314	3.217	2.39
%PropsCOD	35.447	31.964	4.540	4.300	2.36
%Pronrelatifs/EnsPron	30.871	26.846	1.530	4.986	2.35
%AdjDem/EnsDET	3.756	4.951	0.807	1.449	-2.40
travail	0.000	1.992	0.000	2.407	-2.41
%V1ePersSgEnsVConj	0.339	2.773	0.403	2.866	-2.47
%Ptsdesuspension/EnsPonct	0.019	1.375	0.033	1.542	-2.56
NbMoyPhparparagraphe	0.125	1.622	0.331	1.679	-2.60
%PropsSubEnsProps	7.435	10.854	1.727	3.781	-2.64
%Ptsvirgules/EnsPonct	0.515	1.675	0.404	1.221	-2.77
présenter	0.000	3.705	0.000	3.748	-2.88
%PronPersls1ePersSgEnsPronPersls	0.737	6.329	0.967	5.448	-2.99
%AdjCard/EnsDET	0.471	2.321	0.153	1.729	-3.12
%Vàimpératif/EnsVConj	0.373	1.503	0.256	1.036	-3.18
%PronPersls/EnsPron	48.450	55.384	3.159	6.151	-3.28
%deux-points/EnsPonct	3.128	6.068	1.397	2.392	-3.58
%VNégEnsV	4.498	8.264	1.298	2.729	-4.02
Nbmoyen_adVparProp	0.454	0.647	0.026	0.129	-4.35
%d'adV/totalmotsSfiants	7.991	12.150	0.741	2.709	-4.47
%VauFut/EnsVConj	0.623	2.905	0.737	1.484	-4.48
%VinTransEnsV	4.256	8.346	0.912	2.473	-4.82
%motsSfiants/Ensmots	45.197	50.011	1.159	2.517	-5.57
%parenthèses/EnsPonct	3.646	17.078	1.544	6.187	-6.33

## *TnT + lexique épistémique*

Morphosyntaxe et lexique épistémique					
Variable	Moyenne		Ecart-type		V. test
	Auteur	Corpus	Auteur	Corpus	
résultat	19.625	3.213	14.238	6.231	7.68
NUM:ord	36.537	13.265	13.030	10.109	6.71
%classe6	15.826	12.487	0.958	1.472	6.61
PON:comma	66.043	43.905	5.102	10.088	6.39
critère	13.750	3.492	9.404	4.914	6.08
CON:ccl:1st	3.125	1.188	0.966	0.938	6.02
investigation	18.750	3.705	16.238	7.445	5.89
définition	14.125	2.811	14.426	5.692	5.79
PUL	0.132	0.052	0.062	0.043	5.43
CON:dou:1st	3.781	1.535	0.801	1.273	5.14
systématique	6.000	1.393	7.314	2.616	5.13
DET:poss:pp3pl	2.520	1.258	1.204	0.753	4.89
VER:pres	49.311	39.221	4.784	6.279	4.68
conception	13.625	3.189	14.916	6.563	4.63
%classe8	16.758	14.462	0.891	1.476	4.53
CON:add:1st	58.343	45.226	3.922	8.931	4.28
classification	3.500	1.000	4.183	1.779	4.10
concept	8.625	2.803	5.499	4.360	3.89
approche	13.750	4.361	10.733	7.524	3.64
matériau	1.375	0.328	1.798	0.854	3.57
SYM:gram	12.813	3.520	17.022	8.279	3.27
PON:dot	21.456	15.710	1.890	5.151	3.25
questionnement	10.750	4.369	9.614	5.739	3.24
ABR	70.185	42.104	23.403	25.740	3.18
DET:def	69.956	65.345	3.211	4.514	2.98
PRO:rel	30.382	25.587	1.665	4.939	2.83
taxinomie	1.875	0.393	4.594	1.592	2.71
%classe7	22.341	20.393	0.592	2.109	2.69
conséquence	4.625	2.156	2.446	2.989	2.41
déterminer	5.125	2.484	4.314	3.217	2.39
CON:csq:1st	2.600	4.247	0.819	2.042	-2.35
travail	0.000	1.992	0.000	2.407	-2.41
PON:guil	4.512	9.841	6.336	6.288	-2.47
PON:Pvirg	0.456	1.416	0.403	1.123	-2.49
%classe1	0.090	1.307	0.059	1.397	-2.54
PRO:pp1sn	0.244	2.253	0.340	2.268	-2.58
PON:susp	0.010	1.295	0.018	1.447	-2.59
SYM	10.938	33.111	17.045	24.911	-2.59
NP	1.934	6.306	0.765	4.546	-2.80
DET:dem	4.276	6.269	1.018	2.044	-2.84
présenter	0.000	3.705	0.000	3.748	-2.88
LS	8.747	23.572	8.329	14.335	-3.01
CON:ces:1st	0.097	1.259	0.126	1.058	-3.20
%classe11	11.407	13.883	1.194	2.212	-3.26
PREF	0.034	0.248	0.021	0.191	-3.27
DET:indef	18.413	22.917	4.252	4.002	-3.28
PON:colon	3.055	5.921	1.472	2.501	-3.34
CON:dis	7.189	13.784	1.221	5.648	-3.40
%classe12	0.825	1.367	0.104	0.463	-3.41
CON:opp:1st	4.686	11.642	1.414	5.496	-3.69
%classe14	0.512	2.081	0.156	1.230	-3.72

VER:pper	12.662	18.752	2.101	4.620	-3.84
VER:fut	0.275	1.448	0.256	0.769	-4.45
PON:par	3.591	16.175	1.511	6.790	-5.40

## Moirand

### Lexique

Lexique					
Mot	Pourcentage		Fréquence		V. test
	Auteur	Corpus	Auteur	Corpus	
folle	0.34	0.02	129	131	26.392
vache	0.33	0.02	127	130	26.039
scientifiques	0.31	0.02	120	151	22.799
libération	0.25	0.02	96	111	21.170
science	0.34	0.04	130	237	20.335
x	0.02	0.14	7	884	-8.066
nous	0.16	0.40	61	2505	-8.568
dire	0.07	0.27	28	1708	-8.957
sémiotique	0.01	0.16	3	1026	-10.031
est	1.07	1.80	410	11329	-11.857

### Cordial + lexique épistémique

Morphosyntaxe et lexique épistémique					
Variable	Moyenne		Ecart-type		V. test
	Auteur	Corpus	Auteur	Corpus	
expliquer	9.571	1.787	6.543	2.937	7.19
%Ptsdesuspension/EnsPonct	4.786	1.375	1.209	1.542	6.00
observer	5.286	1.705	2.763	2.360	4.12
observation	3.571	0.852	5.095	1.854	3.98
%V3ePersPIEnsVConj	27.199	18.873	4.031	5.698	3.96
expérimentation	1.000	0.189	1.069	0.563	3.91
notion	9.143	3.664	4.940	4.160	3.57
réflexion	5.714	1.738	4.891	3.208	3.36
typologie	7.000	2.525	2.828	3.620	3.36
démonstration	1.714	0.459	1.030	1.033	3.30
%ArtDef/EnsART	75.293	69.928	1.566	4.485	3.25
fonction	8.571	3.779	5.288	4.232	3.07
NbMoyNparProp	3.619	3.016	0.261	0.534	3.06
NbMoyLETparmot	5.149	4.942	0.116	0.190	2.95
paradigme	2.571	0.959	1.917	1.506	2.91
examen	2.286	0.910	1.829	1.293	2.89
%ArtDef/EnsDET	87.261	83.409	1.287	3.664	2.85
NbMoymotsparProp	12.433	10.812	0.785	1.572	2.80
étude	5.714	2.467	3.954	3.270	2.70
étudier	7.429	3.008	5.447	4.555	2.63
Nbmoyen_AdjparProp	1.059	0.867	0.123	0.200	2.60
schéma	12.000	5.828	7.111	6.531	2.56
%VinTransEnsV	10.669	8.346	2.593	2.473	2.55
%PropRelEnsProps	4.343	2.550	1.575	1.935	2.52
analyser	11.429	5.254	7.384	6.776	2.47

%VCOIObligEnsV	15.553	13.517	1.801	2.276	2.43
%AdjPoss/EnsDET	2.849	4.765	0.900	2.185	-2.38
%VlePersPIEnsVConj	0.349	3.137	0.300	2.918	-2.59
%ArtIndef/EnsDET	12.739	16.591	1.287	3.664	-2.85
%V_ETRE_EnsV	15.954	21.550	2.258	4.635	-3.28
%PropAttSuj	14.937	18.866	1.548	3.164	-3.37
%deV/totalmotsSfiants	13.337	16.162	0.857	2.262	-3.39
%Pts/EnsPonct	9.957	16.861	1.098	5.095	-3.68

### *TnT + lexique épistémique*

Morphosyntaxe et lexique épistémique					
Variable	Moyenne		Ecart-type		V. test
	Auteur	Corpus	Auteur	Corpus	
expliquer	9.571	1.787	6.543	2.937	7.19
PREP:1st	1.543	0.586	0.391	0.414	6.28
NC:pl	32.297	21.274	3.989	5.103	5.86
ADJ:pl	45.713	29.949	4.402	7.627	5.61
DTC:pl	13.690	8.536	2.961	2.763	5.06
SIG	31.750	9.415	16.745	12.737	4.76
NUM:dat	39.643	18.398	7.832	13.229	4.36
VER:mod:pper	0.700	0.254	0.383	0.290	4.18
observer	5.286	1.705	2.763	2.360	4.12
observation	3.571	0.852	5.095	1.854	3.98
expérimentation	1.000	0.189	1.069	0.563	3.91
VER:pper	25.186	18.752	2.572	4.620	3.78
CON:dis	21.491	13.784	4.207	5.648	3.70
notion	9.143	3.664	4.940	4.160	3.57
PRO:pp3sn	12.803	8.145	1.919	3.557	3.55
réflexion	5.714	1.738	4.891	3.208	3.36
typologie	7.000	2.525	2.828	3.620	3.36
PON:cote	1.537	0.285	2.394	1.019	3.33
démonstration	1.714	0.459	1.030	1.033	3.30
fonction	8.571	3.779	5.288	4.232	3.07
%classe8	16.089	14.462	0.364	1.476	2.99
paradigme	2.571	0.959	1.917	1.506	2.91
examen	2.286	0.910	1.829	1.293	2.89
étude	5.714	2.467	3.954	3.270	2.70
étudier	7.429	3.008	5.447	4.555	2.63
PRO:disj:pp3pl	1.000	0.540	0.647	0.482	2.59
schéma	12.000	5.828	7.111	6.531	2.56
PON:susp	2.631	1.295	1.071	1.447	2.51
analyser	11.429	5.254	7.384	6.776	2.47
VER:aux:pper	0.129	0.432	0.160	0.345	-2.39
PRO:pp3isn	4.450	6.386	0.941	2.168	-2.42
LS	10.759	23.572	3.168	14.335	-2.43
ADV:neg:1st	13.916	18.297	3.464	4.789	-2.48
%classe5	9.763	10.839	0.688	1.123	-2.60
PRO:pp3msn	4.890	7.917	1.701	2.910	-2.82
PON:dot	8.561	15.710	1.094	5.151	-3.77
NC:sg	62.533	72.420	2.606	6.153	-4.36
PREP	73.817	80.138	3.012	3.502	-4.90
ADJ:sg	52.627	68.964	4.457	7.688	-5.77

# Neveu

## Lexique

Lexique					
Mot	Pourcentage		Fréquence		V. test
	Auteur	Corpus	Auteur	Corpus	
apposition	0.40	0.02	141	155	26.807
support	0.35	0.03	122	179	22.026
segment	0.31	0.03	109	170	20.314
détaché	0.25	0.02	87	117	19.239
référentielle	0.25	0.03	86	176	16.172
x	0.01	0.14	4	884	-8.350
les	1.64	2.27	575	14259	-8.418
politique	0.01	0.15	2	963	-9.441
sémiotique	0.01	0.16	3	1026	-9.489
nous	0.07	0.40	25	2505	-12.100

## Cordial + lexique épistémique

Morphosyntaxe et lexique épistémique					
Variable	Moyenne		Ecart-type		V. test
	Auteur	Corpus	Auteur	Corpus	
objectif	16.000	3.156	9.124	5.677	6.59
hypothèse	10.250	3.418	4.841	3.806	5.23
Nbmoyen_AdjparProp	1.203	0.867	0.206	0.200	4.88
argument	5.875	1.615	2.571	2.549	4.87
paradigme	3.375	0.959	1.867	1.506	4.67
NbMoyLETparmot	5.204	4.942	0.148	0.190	4.02
présupposé	5.125	1.992	1.364	2.310	3.95
%Pronrelatifs/EnsPron	33.478	26.846	3.149	4.986	3.88
analyser	13.875	5.254	8.433	6.776	3.71
définir	5.125	2.016	2.619	2.479	3.65
%d'Adj/totalmotsSfiants	18.521	15.935	1.696	2.101	3.59
NbMoymotsparProp	12.665	10.812	1.346	1.572	3.44
tableau	8.000	2.549	8.930	4.775	3.33
%V3ePersSgEnsVConj	81.270	74.143	2.245	6.246	3.32
NbMoyNparProp	3.620	3.016	0.577	0.534	3.29
Nbmoyen_adVparProp	0.769	0.647	0.050	0.129	2.74
concevoir	2.375	0.738	2.446	1.759	2.71
désigner	5.750	2.820	3.527	3.483	2.45
théorie	12.000	6.025	9.233	7.249	2.40
%PronPersls1ePersPIEnsPronPersls	1.467	6.776	1.536	6.434	-2.40
%V1ePersPIEnsVConj	0.604	3.137	0.610	2.918	-2.53
%PronPersls/EnsPron	49.128	55.384	2.904	6.151	-2.96
%deV/totalmotsSfiants	13.713	16.162	1.761	2.262	-3.16
%deux-points/EnsPonct	3.322	6.068	1.158	2.392	-3.34



## *TnT + lexique épistémique*

Morphosyntaxe et lexique épistémique					
Variable	Moyenne		Ecart-type		V. test
	Auteur	Corpus	Auteur	Corpus	
objectif	16.000	3.156	9.124	5.677	6.59
hypothèse	10.250	3.418	4.841	3.806	5.23
argument	5.875	1.615	2.571	2.549	4.87
paradigme	3.375	0.959	1.867	1.506	4.67
PRO:rel	32.416	25.587	2.051	4.939	4.03
présupposé	5.125	1.992	1.364	2.310	3.95
analyser	13.875	5.254	8.433	6.776	3.71
définir	5.125	2.016	2.619	2.479	3.65
%classe3	8.156	6.905	0.660	1.004	3.63
CON:csq:1st	6.734	4.247	1.978	2.042	3.55
tableau	8.000	2.549	8.930	4.775	3.33
VER:pper	23.775	18.752	3.078	4.620	3.17
ADJ:sg	76.891	68.964	4.379	7.688	3.00
CON:ref:1st	4.615	3.195	1.710	1.514	2.73
CON:cau:1st	4.665	2.909	2.281	1.878	2.72
concevoir	2.375	0.738	2.446	1.759	2.71
NC:sg	78.146	72.420	4.575	6.153	2.71
CON:ces:1st	2.176	1.259	1.272	1.058	2.52
PON:comma	52.554	43.905	3.401	10.088	2.50
désigner	5.750	2.820	3.527	3.483	2.45
théorie	12.000	6.025	9.233	7.249	2.40
SUB	93.124	89.305	2.552	4.735	2.35
SUB:1st	6.876	10.695	2.552	4.735	-2.35
PON:susp	0.120	1.295	0.146	1.447	-2.37
PRO:pp1pl	0.311	3.177	0.445	3.526	-2.37
PRO:clit:pp1pl	0.096	0.775	0.184	0.778	-2.54
NC:pl	16.653	21.274	2.102	5.103	-2.64
ADJ:pl	22.619	29.949	4.471	7.627	-2.80
PON:colon	3.304	5.921	1.160	2.501	-3.05

## **Rabatel**

### *Lexique*

Lexique					
Mot	Pourcentage		Fréquence		V. test
	Auteur	Corpus	Auteur	Corpus	
pdv	0.88	0.06	377	378	44.807
narrateur	0.55	0.05	235	298	31.020
personnage	0.40	0.04	171	221	26.213
perceptions	0.26	0.02	111	121	22.911
mi	0.23	0.02	99	109	21.510
langues	0.00	0.10	0	638	-9.156
politique	0.01	0.15	5	963	-9.819
français	0.00	0.13	2	838	-9.835
sémiotique	0.00	0.16	1	1026	-11.380
langue	0.01	0.21	6	1316	-11.717

## *Cordial + lexique épistémique*

Morphosyntaxe et lexique épistémique					
Variable	Moyenne		Ecart-type		V. test
	Auteur	Corpus	Auteur	Corpus	
vérifier	21.625	3.197	25.470	8.576	6.26
%VàImpft/EnsVConj	12.722	4.155	4.911	4.240	5.89
comparaison	2.750	0.844	1.984	1.312	4.23
%VauSubjplus-que-parfait/EnsVCon	0.170	0.037	0.142	0.105	3.68
%V3ePersSgEnsVConj	81.997	74.143	2.759	6.246	3.66
%Vaupassésimple/EnsVConj	4.569	1.637	2.320	2.525	3.38
interprétation	6.500	2.156	2.693	3.822	3.31
%AdjDem/EnsDET	6.596	4.951	0.782	1.449	3.31
condition	3.750	1.344	4.603	2.164	3.24
désigner	6.625	2.820	5.894	3.483	3.18
%Vauplus-que-parfait/EnsVConj	1.336	0.546	0.734	0.762	3.03
classification	2.625	1.000	2.643	1.779	2.66
%VTransindirectsEnsV	53.743	50.103	4.017	4.078	2.60
%deV/totalmotsSfiants	18.153	16.162	1.255	2.262	2.56
%V ETRE EnsV	25.275	21.550	4.096	4.635	2.34
<b>-----</b>					
%Vauprésent/EnsVConj	71.545	78.326	5.280	7.886	-2.51
%Vaupassécomposé/EnsVConj	2.284	5.503	0.899	3.623	-2.59
%VCODObligEnsV	31.695	35.402	1.800	4.066	-2.66
%VauFut/EnsVConj	1.397	2.905	0.917	1.484	-2.96
%VTransdirectsEnsV	48.671	53.185	2.526	3.783	-3.48

## *TnT + lexique épistémique*

Morphosyntaxe et lexique épistémique					
Variable	Moyenne		Ecart-type		V. test
	Auteur	Corpus	Auteur	Corpus	
SIG:ling	57.374	10.443	17.002	16.962	8.06
vérifier	21.625	3.197	25.470	8.576	6.26
NUM:par	28.374	7.924	11.341	11.252	5.30
DET:dem	9.550	6.269	1.423	2.044	4.68
comparaison	2.750	0.844	1.984	1.312	4.23
DTC:sg	13.930	10.741	2.064	2.425	3.83
interprétation	6.500	2.156	2.693	3.822	3.31
condition	3.750	1.344	4.603	2.164	3.24
désigner	6.625	2.820	5.894	3.483	3.18
PRO:dem	18.454	14.056	6.627	4.061	3.16
CON:opp:1st	17.028	11.642	11.095	5.496	2.86
%classe12	1.813	1.367	0.260	0.463	2.81
classification	2.625	1.000	2.643	1.779	2.66
VER:pres	44.894	39.221	6.226	6.279	2.63
VER:subi	0.106	0.037	0.119	0.077	2.63
%classe2	9.038	7.898	1.119	1.427	2.33
<b>-----</b>					
VER:mod:inf	0.046	0.195	0.065	0.185	-2.34
VER:mod:pres	2.859	4.207	0.852	1.644	-2.39
VER:aux:pper	0.141	0.432	0.130	0.345	-2.46
DET:poss:pp3pl	0.565	1.258	0.330	0.753	-2.68
ABR	9.871	42.104	3.136	25.740	-3.65

## Rastier

### Lexique

Lexique					
Mot	Pourcentage		Fréquence		V. test
	Auteur	Corpus	Auteur	Corpus	
ontologie	0.20	0.02	111	119	21.774
textes	0.44	0.09	251	590	21.652
corpus	0.41	0.10	233	602	19.718
genres	0.27	0.04	153	271	19.679
texte	0.47	0.13	266	787	19.326
dans	1.22	1.73	687	10869	-10.210
politique	0.02	0.15	10	963	-10.655
ce	0.45	0.81	253	5125	-10.963
c	0.24	0.54	138	3387	-11.045
x	0.00	0.14	1	884	-12.266

### Cordial + lexique épistémique

Morphosyntaxe et lexique épistémique					
Variable	Moyenne		Ecart-type		V. test
	Auteur	Corpus	Auteur	Corpus	
théorique	16.111	3.115	12.529	6.820	5.92
méthodologique	1.333	0.148	1.700	0.623	5.91
examen	3.222	0.910	1.931	1.293	5.55
définir	6.333	2.016	2.211	2.479	5.41
constat	7.111	2.156	3.247	2.903	5.30
expérimental	21.333	7.230	9.357	9.050	4.84
problématique	7.333	2.574	5.598	3.720	3.97
problème	7.111	2.680	6.437	3.474	3.96
procéder	7.333	2.549	3.621	3.785	3.92
définition	9.778	2.811	4.289	5.692	3.80
catégorisation	2.889	0.844	3.178	1.674	3.79
représenter	6.556	3.107	4.058	2.916	3.67
méthodologie	2.778	0.746	2.657	1.725	3.66
réflexion	5.444	1.738	6.379	3.208	3.59
tableau	8.000	2.549	6.633	4.775	3.54
NbMoyLEtparmot	5.151	4.942	0.083	0.190	3.42
conclusion	2.889	1.107	2.470	1.619	3.42
%ArtDef/EnsART	74.762	69.928	1.381	4.485	3.35
%Pts/EnsPonct	22.167	16.861	1.757	5.095	3.23
conceptuel	6.333	1.943	3.916	4.249	3.21
%V3ePersPIEnsVConj	24.661	18.873	4.788	5.698	3.15
décrire	4.889	1.738	9.279	3.289	2.97
%VinTransEnsV	10.634	8.346	2.091	2.473	2.87
étudier	7.222	3.008	4.709	4.555	2.87
démarche	4.667	2.008	3.742	2.884	2.86
NbMoyPropsparphrase	3.074	2.378	0.446	0.788	2.75
conception	8.889	3.189	9.792	6.563	2.70
traiter	3.000	1.221	4.497	2.067	2.67
NbtotalPhverbales	351.667	262.656	80.944	105.241	2.63
investigation	9.889	3.705	10.493	7.445	2.58

discussion	3.000	1.115	1.633	2.389	2.45
%VNegEnsV	10.344	8.264	2.740	2.729	2.37
<b>%deProndémonstratifs/EnsPron</b>					
NbMoyPhparparagraphe	0.278	1.622	0.533	1.679	-2.49
indiquer	0.000	1.361	0.000	1.650	-2.56
%Ptsdesuspension/EnsPonct	0.050	1.375	0.062	1.542	-2.67
%PronPersls lePersSgEnsPronPersls	1.608	6.329	1.182	5.448	-2.69
NbMoyPronparProp	0.669	0.863	0.052	0.188	-3.20
%AdjDem/EnsDET	3.352	4.951	0.714	1.449	-3.42

### *TnT + lexique épistémique*

Morphosyntaxe et lexique épistémique					
Variable	Moyenne		Ecart-type		V. test
	Auteur	Corpus	Auteur	Corpus	
théorique	16.111	3.115	12.529	6.820	5.92
méthodologique	1.333	0.148	1.700	0.623	5.91
examen	3.222	0.910	1.931	1.293	5.55
définir	6.333	2.016	2.211	2.479	5.41
constat	7.111	2.156	3.247	2.903	5.30
CON:exe:1st	4.923	2.300	1.385	1.625	5.01
expérimental	21.333	7.230	9.357	9.050	4.84
PRO:pp3msn	12.350	7.917	2.186	2.910	4.73
ABR	78.386	42.104	15.591	25.740	4.38
problématique	7.333	2.574	5.598	3.720	3.97
problème	7.111	2.680	6.437	3.474	3.96
procéder	7.333	2.549	3.621	3.785	3.92
PON:dot	22.046	15.710	1.827	5.151	3.82
définition	9.778	2.811	4.289	5.692	3.80
catégorisation	2.889	0.844	3.178	1.674	3.79
DTC:pl	11.811	8.536	2.525	2.763	3.68
représenter	6.556	3.107	4.058	2.916	3.67
méthodologie	2.778	0.746	2.657	1.725	3.66
NC:pl	27.272	21.274	5.078	5.103	3.65
PRO:pp3pl	4.350	2.674	1.357	1.431	3.64
réflexion	5.444	1.738	6.379	3.208	3.59
tableau	8.000	2.549	6.633	4.775	3.54
conclusion	2.889	1.107	2.470	1.619	3.42
VER:mod:fut	0.220	0.085	0.113	0.130	3.22
conceptuel	6.333	1.943	3.916	4.249	3.21
CON:ccl:1st	2.093	1.188	0.893	0.938	3.00
décrire	4.889	1.738	9.279	3.289	2.97
%classe6	13.853	12.487	0.712	1.472	2.88
étudier	7.222	3.008	4.709	4.555	2.87
démarche	4.667	2.008	3.742	2.884	2.86
CON:cau:1st	4.633	2.909	0.922	1.878	2.85
ADV:neg:1st	22.554	18.297	4.359	4.789	2.76
DET:poss:pp3pl	1.919	1.258	0.598	0.753	2.73
PRO:clit	11.157	9.204	2.266	2.234	2.71
conception	8.889	3.189	9.792	6.563	2.70
traiter	3.000	1.221	4.497	2.067	2.67
ADJ:pl	36.427	29.949	6.968	7.627	2.64
investigation	9.889	3.705	10.493	7.445	2.58
VER:mod:pres	5.536	4.207	0.839	1.644	2.51
discussion	3.000	1.115	1.633	2.389	2.45

SUB:1st	14.322	10.695	4.850	4.735	2.38
PRO:dem	10.949	14.056	0.845	4.061	-2.38
SUB	85.678	89.305	4.850	4.735	-2.38
ADJ:sg	62.816	68.964	6.932	7.688	-2.48
indiquer	0.000	1.361	0.000	1.650	-2.56
PON:susp	0.051	1.295	0.063	1.447	-2.67
%classe4	5.022	6.237	0.332	1.408	-2.68
SYM	10.223	33.111	12.500	24.911	-2.85
PREP:1st	0.206	0.586	0.090	0.414	-2.85
PREP	76.877	80.138	2.136	3.502	-2.89
VER:parpres	2.592	4.668	0.589	2.000	-3.22
DET:dem	4.091	6.269	0.799	2.044	-3.31

## Rosier

### Lexique

Lexique					
Mot	Pourcentage		Fréquence		V. test
	Auteur	Corpus	Auteur	Corpus	
<b>p</b>	0.61	0.11	200	691	20.070
<b>nous</b>	1.23	0.40	401	2505	19.838
<b>discours</b>	1.10	0.35	359	2193	19.086
<b>espèce</b>	0.26	0.02	84	122	18.619
<b>mon</b>	0.35	0.04	116	263	18.411
<b>deux</b>	0.16	0.33	53	2046	-5.799
<b>sciences</b>	0.00	0.06	0	374	-5.885
<b>sens</b>	0.11	0.27	35	1708	-6.571
<b>espace</b>	0.02	0.13	6	837	-7.120
<b>sémiotique</b>	0.02	0.16	7	1026	-8.006

### Cordial + lexique épistémique

Morphosyntaxe et lexique épistémique					
Variable	Moyenne		Ecart-type		V. test
	Auteur	Corpus	Auteur	Corpus	
%V1ePersPIEnsVConj	9.263	3.137	5.127	2.918	6.12
%PronPersls1ePersPIEnsPronPersls	17.035	6.776	12.546	6.434	4.65
%dePhExcl/EnsPh	3.501	0.854	3.716	1.769	4.36
%Ptsd'exclamation/EnsPonct	1.646	0.423	1.823	0.912	3.91
%AdjPoss/EnsDET	7.650	4.765	4.540	2.185	3.85
%Vàimpératif/EnsVConj	2.803	1.503	1.032	1.036	3.65
%dePhInt/EnsPh	5.870	2.815	1.874	2.532	3.52
%PronPersls/EnsPron	62.340	55.384	6.385	6.151	3.30
expérimental	17.250	7.230	10.022	9.050	3.23
théorie	14.000	6.025	10.452	7.249	3.21
présupposé	4.125	1.992	3.219	2.310	2.69
%AdjPoss1ePersSg/EnsAdjPoss	25.000	5.952	43.301	21.449	2.59
%parenthèses/EnsPonct	22.471	17.078	2.577	6.187	2.54
%Ptsd'interrogation/EnsPonct	2.305	1.270	0.869	1.295	2.33
présenter	0.375	3.705	0.696	3.748	-2.59

%V3ePersSgEnsVConj	68.533	74.143	5.368	6.246	-2.62
%ArtDef/EnsART	65.760	69.928	6.261	4.485	-2.71
%Pronrelatifs/EnsPron	21.417	26.846	4.551	4.986	-3.17

### *TnT + lexique épistémique*

Morphosyntaxe et lexique épistémique					
Variable	Moyenne		Ecart-type		V. test
	Auteur	Corpus	Auteur	Corpus	
PRO:pp1pl	10.514	3.177	6.672	3.526	6.06
DET:poss:pp1pl	1.755	0.498	1.067	0.639	5.73
PON:int	2.150	0.882	0.847	0.789	4.68
VER:imp	1.370	0.484	1.006	0.591	4.37
PRO:disj:pp1pl	1.005	0.200	2.021	0.599	3.91
expérimental	17.250	7.230	10.022	9.050	3.23
théorie	14.000	6.025	10.452	7.249	3.21
PON:colon	8.536	5.921	2.427	2.501	3.05
DET:poss:pp1sn	1.143	0.326	2.524	0.796	2.99
DET:poss:pp2sn	0.099	0.024	0.199	0.075	2.91
PON:par	22.903	16.175	2.789	6.790	2.89
présupposé	4.125	1.992	3.219	2.310	2.69
PRO:clit:pp1pl	1.450	0.775	0.905	0.778	2.53
présenter	0.375	3.705	0.696	3.748	-2.59
PON:comma	34.740	43.905	3.951	10.088	-2.65
PRO:dem	10.319	14.056	2.617	4.061	-2.68

## Siblot

### *Lexique*

Lexique					
Mot	Pourcentage		Fréquence		V. test
	Auteur	Corpus	Auteur	Corpus	
congo	0.26	0.02	68	96	17.778
prédicat	0.18	0.03	48	174	10.471
thème	0.20	0.04	52	222	10.112
nom	0.34	0.10	88	612	9.971
dénomination	0.21	0.04	55	271	9.677
exemple	0.05	0.18	14	1138	-5.662
ou	0.50	0.79	130	4962	-5.781
sémiotique	0.03	0.16	9	1026	-6.211
	0.00	0.09	0	540	-6.356
espace	0.01	0.13	3	837	-6.877

### *Cordial + lexique épistémique*

Morphosyntaxe et lexique épistémique				
Variable	Moyenne		Ecart-type	V. test

	<b>Auteur</b>	<b>Corpus</b>	<b>Auteur</b>	<b>Corpus</b>	
NbMoyPropsparphrase	4.024	2.378	1.324	0.788	4.75
démonstration	2.400	0.459	2.577	1.033	4.27
%Pts/EnsPonct	26.282	16.861	0.765	5.095	4.20
démarche	7.000	2.008	4.858	2.884	3.94
analyser	15.800	5.254	10.186	6.776	3.54
conséquence	6.400	2.156	3.499	2.989	3.23
choisir	3.400	0.795	4.079	1.855	3.19
catégorisation	3.000	0.844	1.789	1.674	2.93
procéder	7.400	2.549	3.499	3.785	2.91
problème	7.000	2.680	3.950	3.474	2.83
constater	2.800	0.943	1.720	1.522	2.78
indiquer	3.200	1.361	1.166	1.650	2.53
systématiser	0.800	0.221	0.748	0.551	2.39
conclusion	2.800	1.107	2.786	1.619	2.38

### *TnT + lexique épistémique*

<b>Morphosyntaxe et lexique épistémique</b>					
<b>Variable</b>	<b>Moyenne</b>		<b>Ecart-type</b>		<b>V. test</b>
	<b>Auteur</b>	<b>Corpus</b>	<b>Auteur</b>	<b>Corpus</b>	
démonstration	2.400	0.459	2.577	1.033	4.27
démarche	7.000	2.008	4.858	2.884	3.94
PON:dot	24.150	15.710	2.139	5.151	3.73
analyser	15.800	5.254	10.186	6.776	3.54
PRO:poss:pp1pl	0.090	0.012	0.112	0.050	3.52
PON:guil	18.942	9.841	2.625	6.288	3.29
conséquence	6.400	2.156	3.499	2.989	3.23
choisir	3.400	0.795	4.079	1.855	3.19
DTC:sg	14.126	10.741	1.343	2.425	3.17
catégorisation	3.000	0.844	1.789	1.674	2.93
procéder	7.400	2.549	3.499	3.785	2.91
problème	7.000	2.680	3.950	3.474	2.83
constater	2.800	0.943	1.720	1.522	2.78
CON:pre:1st	1.320	0.582	0.686	0.610	2.75
ADV:neg:1st	23.738	18.297	2.334	4.789	2.58
indiquer	3.200	1.361	1.166	1.650	2.53
systématiser	0.800	0.221	0.748	0.551	2.39
conclusion	2.800	1.107	2.786	1.619	2.38
ADV	69.964	76.185	4.283	5.111	-2.77
CON:ref:1st	0.878	3.195	0.730	1.514	-3.48