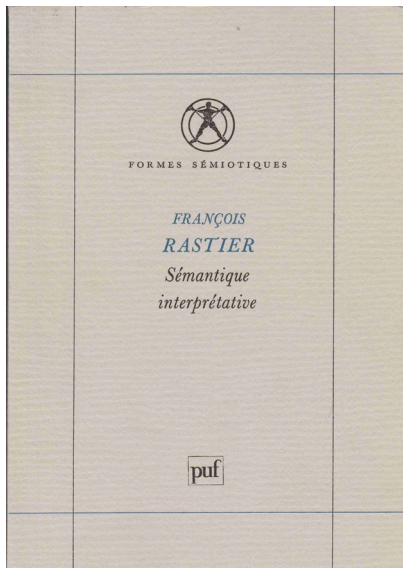


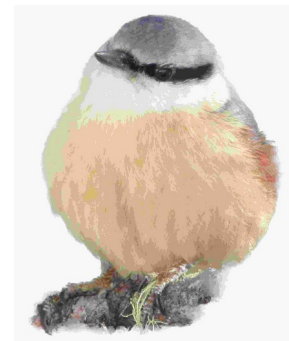
SEMI CULT

*II Congrès international Sémiotique et culture
De la linguistique à l'anthropologie sémiotique
Axe 2 – Corpus, texte et méthodes, 5 octobre 2022*

Un corpus pour la Sémantique interprétative L'initiative du projet Sittelle



Bénédicte PINCEMIN
IHRIM, université de Lyon, CNRS



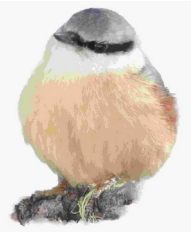
This work is licensed under the Creative Commons Attribution 4.0 International License.
<http://creativecommons.org/licenses/by/4.0/>



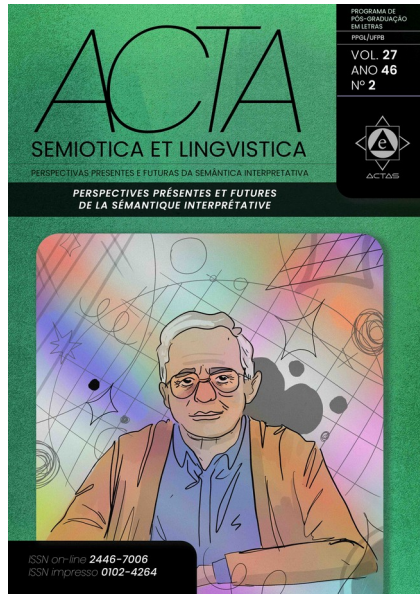
Plan

- I. Définition du projet
- II. Approche textométrique – Illustrations
- III. Usages envisagés – Exemple de parcours
- IV. Réflexions méthodologiques sur le corpus

Projet Sittelle : définition



- Sémantique Interprétative eT TextuELLE



- Constituer un corpus de textes de François Rastier en sémantique
- Initiative de Bénédicte Pincemin et François Rastier, lancée en 2021
- Présentée dans l'article de B. Pincemin du numéro spécial « Perspectives présentes et futures de la Sémantique interprétative » de *Acta Semiotica et Lingvistica* (v. 27 n. 2, 2022)

Projet Sittelle : principes



- Réalisation progressive
 - Jusqu'à une centaine de textes (une dizaine début 2023 ?)
- Science ouverte
 - Format standard XML **TEI**
 - Contributions collaboratives (Guide d'encodage,...)
 - Diffusion à l'ensemble de la communauté scientifique, autant que possible publique sous licence CC BY
- Édition numérique en vue de donner accès à des lectures outillées, notamment textométriques
 - Corpus aux formats de logiciels
 - En ligne

Textométrie ?

- Par comparaison à une recherche dans un fichier PDF :
 - Des **objets de recherche** plus précis, des **sorties** synthétiques
 - Infralexical et morphologique
 - Grammatical
 - Supralexical et syntagmatique
 - Typographique et structurel
 - Une **vision intertextuelle**
 - Recherches transversales
 - Explorations contrastives
- Concordance
- Liste de mots avec fréquence
- Cooccurrences...
- Spécificités
- Cartographie AFC...

Recherche

Infralexicale

Grammaticale

Supralexicale

Typo. & struct.

Intertextualité

Transversale

Contrastive

Sorties

Concordance

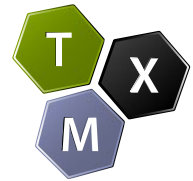
Liste de mots

Spécificités

Cartographie

Illustration sur le corpus « Zéro »

- 28 articles publiés en ligne sur *Texto!*
 - De François Rastier, et en sémantique (sélection sans FR – voir détail dans l'[article ASEL](#))
 - Au format HTML (*Texto!* [ancienne série](#))
 - Datant de 1992 à 2006
 - Importé dans le logiciel open-source TXM : 320 000 mots
- Corpus simplifié
 - Dans son contenu (sélection pragmatique)
 - Dans sa forme (codage texte brut TXT)



Recherche

Infralexicale

Grammaticale

Supralexicale

Typo. & struct.

Intertextualité

Transversale

Contrastive

Sorties

Concordance

Liste de mots

Spécificités

Cartographie

Requête: `iso.*"%c>@word` Propriétés: word Éditer

word	Fréquence
isotopies	125
isotopie	115
isosémies	29
isolé	16
isoler	14
isosémie	11
isonomie	9
isolément	7
isolés	6

Requête: `iso.*(ie|isme|ique)s?"%c>...` Propriétés: frlemma Éditer

frlemma	Fréquence
isotopie	241
isosémies	29
isosémie	11
isonomie	9
isomorphisme	3
isophonie	3
isotopique	3
isophonies	1

Console

Sortie standard

Index de `<"iso.*"%c>`, propriété `@word`, dans le corpus TEXT0211026...
31 items pour 379 occurrences.

Index de `<"iso.*(ie|isme|ique)s?"%c>`, propriété `@frlemma`, dans le corpus TEXT0211026...
8 items pour 300 occurrences.

*Les mots qui commencent par **iso-** ?*

*Et en précisant qu'ils se terminent par **-ie, -isme, -ique** ?* Partie II - 7



Recherche

Infralexicale

Grammaticale

Supralexicale

Typo. & struct.

Intertextualité

Transversale

Contrastive

Sorties

Concordance

Liste de mots

Spécificités

Cartographie

Requête Propriétés : frl

frlemma	Fréquence
texte	1420
exemple	814
genre	655
sen sens	596
forme	594
langue	586
discours	524
mot	516
p	497
langage	459
sémantique	457

Requête Propriétés : frl

frlemma	Fréquence
sémantique	794
même	729
autre	658
linguistique	495
textuel	259
cognitif	235
scientifique	233
interprétatif	230
générique	218
général	214
divers	202

Requête Propriétés : frler

frlemma	Fréquence
être	5469
avoir	2118
pouvoir	1404
faire	572
permettre	464
rester	427
définir	421
devoir	350
relever	288
dire	280
décrire	262

Console

Sortie standard

Index de <[frpos = "NOM"]>, propriété @frlemma, dans le corpus TEXT0211026...
6 047 items pour 67 557 occurrences.

Index de <[frpos = "ADJ"]>, propriété @frlemma, dans le corpus TEXT0211026...
2 934 items pour 24 414 occurrences.

Index de <[frpos = "VER.*"]>, propriété @frlemma, dans le corpus TEXT0211026...
1 933 items pour 35 035 occurrences.

Les **noms** communs les plus utilisés ? Les **adjectifs** qualificatifs, les **verbes** ?

Recherche

Infralexicale

Grammaticale

Supralexicale

Typo. & struct.

Intertextualité

Transversale

Contrastive

Sorties

Concordance

Liste de mots

Spécificités

Cartographie

Requête: ET.*"?[frl

Propriétés: word

word	Fréquence
sémantique cognitive	78
sémantique des textes	51
linguistique de corpus	45
sémantique interprétative	20
linguistique historique	17
linguistique générale	15
sémantique linguistique	14
sémantique structurale	14
sémantique lexicale	13
linguistique cognitive	10
sémantique formelle	10
linguistique textuelle	9
sémantique différentielle	9

100 / 134 t 523, v 134, fmin 1, fmax 78

Requête: "N|OM"][f

Propriétés: word

word	Fréquence
formes sémantiques	79
sémantique cognitive	78
parcours interprétatifs	56
champs génériques	51
molécules sémiques	50
positivisme logique	44
sciences sociales	41
pratiques sociales	40
molécule sémique	39
parcours interprétatif	38
isotopies génériques	36
fonds sémantiques	33
traits sémantiques	32

1 100 / 8556 t 14377, v 8556, fmin 1, fmax 79

Console

Sortie standard

```
Index de <[frlemma="(séma|linguis)tique"%c & frpos="NOM"]([frpos="ADJ"]|[frpos="PRP.*"]|[frpos="DET.*"]?[frpos="NOM"]>,
134 items pour 523 occurrences.
Index de <[frpos = "NOM"]|[frpos = "ADJ"]>, propriété @word, dans le corpus TEXT0211026...
8 556 items pour 14 377 occurrences.
```

Quelles sémantiques (ou linguistiques) ? Et plus généralement les NOM + ADJ ?

Recherche

Infralexicale

Grammaticale

Supralexicale

Typo. & struct.

Intertextualité

Transversale

Contrastive

Sorties

Concordance

Liste de mots

Spécificités

Cartographie

Requête: `(([word="local.*"%c]*[word="global.*"%c]) | ([word="global.*"%c]*[word="local.*"%c]) within 4`

text_id	Contexte gauche	Pivot	Contexte droit
2005Enjeux	ulièrement révélatrice, car ils multiplient les rapports	global / local	(par des structures en abyme) et les rapports entre pl
2003Semiotique-ontologie	la partie au tout. L'accès du	global au local	, dans la mémorisation par exemple - et toute interpré
2006Formes-semantiques	la partie au tout. L'accès du	global au local	, dans la mémorisation par exemple, reste en effet mé
2005Doxa	l'externe à l'interne, mais du	global au local	. Les perturbations locales, comme les paradoxes, peu
2006Formes-semantiques	. Elle engage à concevoir le rapport du	global au local	d'une façon moins simpliste et moins statique que cel
2003Semiotique-ontologie	intermédiaires permettent de concevoir le rapport du	global au local	d'une façon moins simpliste et moins statique que cel
2001Action	de l'évolution biologique. Le rapport du	global au local	se pense alors en termes d'appartenance au monde d
2006Textes	ou religieux. 1. 5. Du	global au local	Comme tout usage modifie et configure potentiellem
2005Herm-et-ling	ou de l'intertextualité; elle conçoit le	global comme extension du local	; elle utilise des codes qui transforment en tout ou par
2005Classes-lexicales	[27]. Cependant, comme le	global détermine le local	, même si une perturbation locale importante peut mc
1996PourSdT	façon nouvelle, dans la mesure où le	global détermine le local	. On ne peut réduire un texte à une suite de phrases
2005Microsemantique	ication élémentaire du principe herméneutique que le	global détermine le local	. À un niveau de complexité supérieur, l'ordre herméne
1996Themes	se règle sur les principes herméneutiques que le	global détermine le local	, et que l'hypothèse gouverne l'objectivation. Ainsi, les
2003Semiotique-ontologie	du sens, conformément au principe que le	global détermine le local	. Cela conduit à réexaminer l'engagement ontologique
1996Concepts	à reformuler dans ce cadre le rapport entre	global et local	. Enfin à reformuler le lien entre les textes et leur ento
2002Acces	globaux. Le paramétrage de ces rapports entre	global et local	reste un problème ouvert: il diffère vraisemblableme
2006Textes	xtualité globale La compréhension des rapports entre	global et local	est une des difficultés majeures que rencontre la lingu
2005Saussure	et global; mieux, les rapports entre	global et local	sur chaque plan du langage ne peuvent être établis sar
2005Microsemantique	aut encore décrire l'incidence du contexte linguistique	global sur le contexte local	. Nous distinguons quatre paliers de contextualité: le
2002Genres	ailleurs, en raison de la détermination du	global sur le local	, les variations sont déterminées par les discours, les
2004Ontologies	, car il témoigne de la détermination du	global sur le local	. En voici un exemple, qui intéresse la distinction entre
1997Defigements	sur la lexie, et plus généralement du	global sur le local	. La typologie que nous allons présenter se fonde sur
2005Microsemantique	qu'ils rencontrent témoignent de l'incidence du	global sur le local	. Il faudrait donc leur adjoindre un composant qui met
2001Semiotique	mais surtout empêche de penser l'incidence du	global sur le local	, en l'occurrence du texte sur chacun des signes qui le
2002Macrosemantique	clairement ses lacunes: en effet, le	global y détermine le local	et le recompose. C'est pourquoi une phrase et a fortio
2002Genres	re constitue une variable qui véhicule des contraintes	globales et locales	, négliger cette variable peut obscurcir la compréhens
2001Ling-de-style	propre aux normes linguistiques, qu'elles soient	globales ou locales	. Pour le plan du contenu, cela exige une théorie capab
2005Semantique-cognitive	du signe linguistique au texte est celui du	local au global	. e) Le sens des signes linguistiques peut varier indéfin
2005Enjeux	texte, et plus généralement le rapport du	local au global	médiatisé par plusieurs paliers d'organisation de cor

1 - 47 / 47

*Relevé des contextes associant « local » et « global »
(28/47, tri pivot puis contexte gauche)*

Recherche

Infralexicale

Grammaticale

Supralexicale

Typo. & struct.

Intertextualité

Transversale

Contrastive

Sorties

Concordance

Liste de mots

Spécificités

Cartographie

Objet de recherche typographique et structurel

- Par exemple, recherche de trois mentions soulignées dans un même paragraphe (ou dans une même liste), comme on a dans cet extrait :

C. Enfin, aux discours théoriques sont associées des *instances auxiliaires* qui permettent diverses transpositions. (i) Une instance d'*interdisciplinarité* permet le dialogue avec les disciplines voisines. (ii) Une instance *applicative* dirige les mises en œuvre pratiques, qui imposent des traductions simplifiantes à l'usage des disciplines techniques : une théorie doit pouvoir donner des versions partielles et affaiblies d'elle-même, voire partager son champ avec d'autres théories. (iii) Enfin, une instance *didactique* règle la transmission des savoirs acquis. Soit :

F. Rastier, « Pour une sémantique des textes théoriques », *RSP*, 2005.

(perspective – pas illustrable dans le corpus Zéro)

Recherche

Infralexicale

Grammaticale

Supralexicale

Typo. & struct.

Intertextualité

Transversale

Contrastive

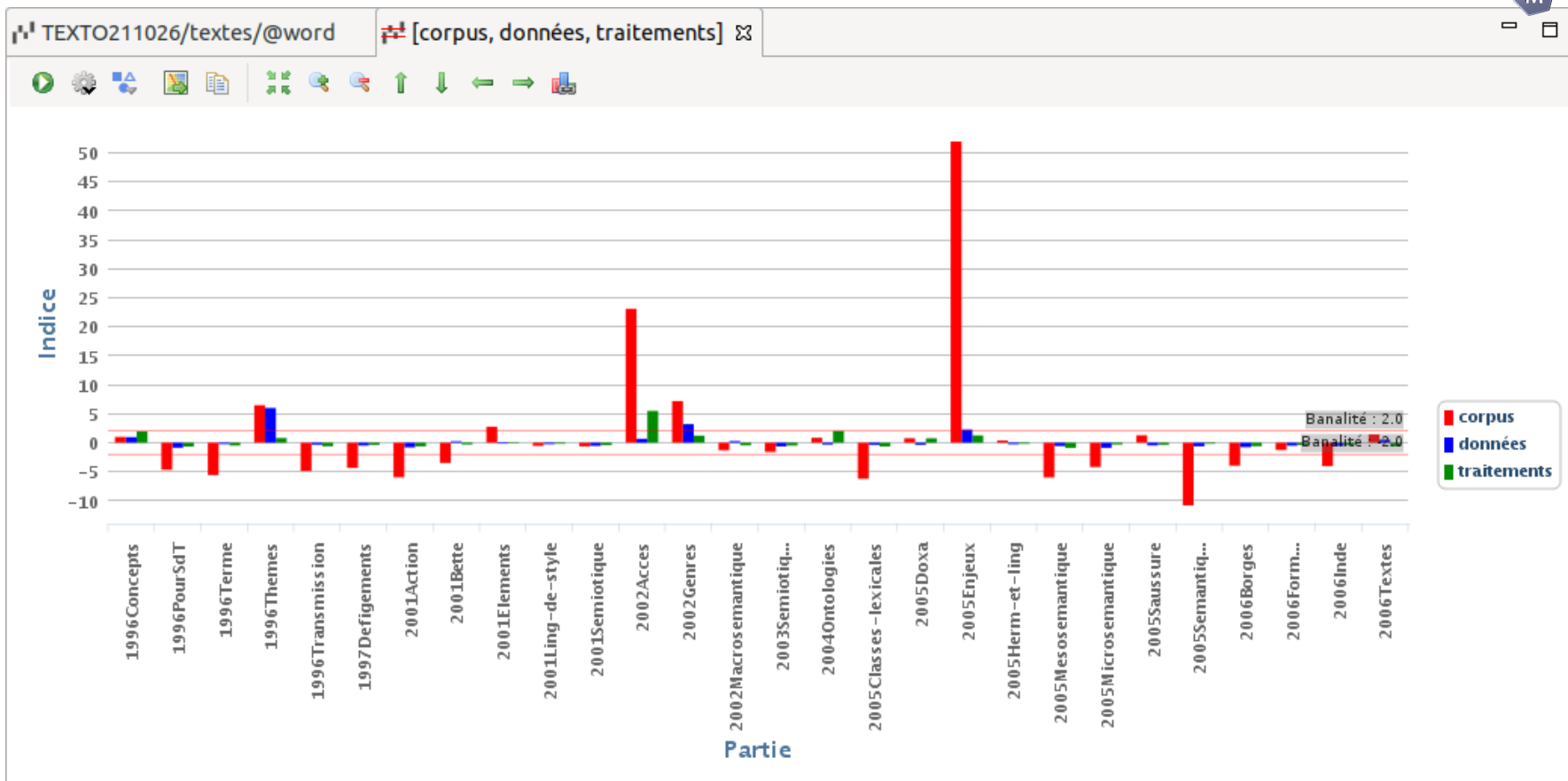
Sorties

Concordance

Liste de mots

Spécificités

Cartographie



Quels textes utilisent particulièrement « corpus », « données », « traitements » ?

Recherche

Infralexicale

Grammaticale

Supralexicale

Typo. & struct.

Intertextualité

Transversale

Contrastive

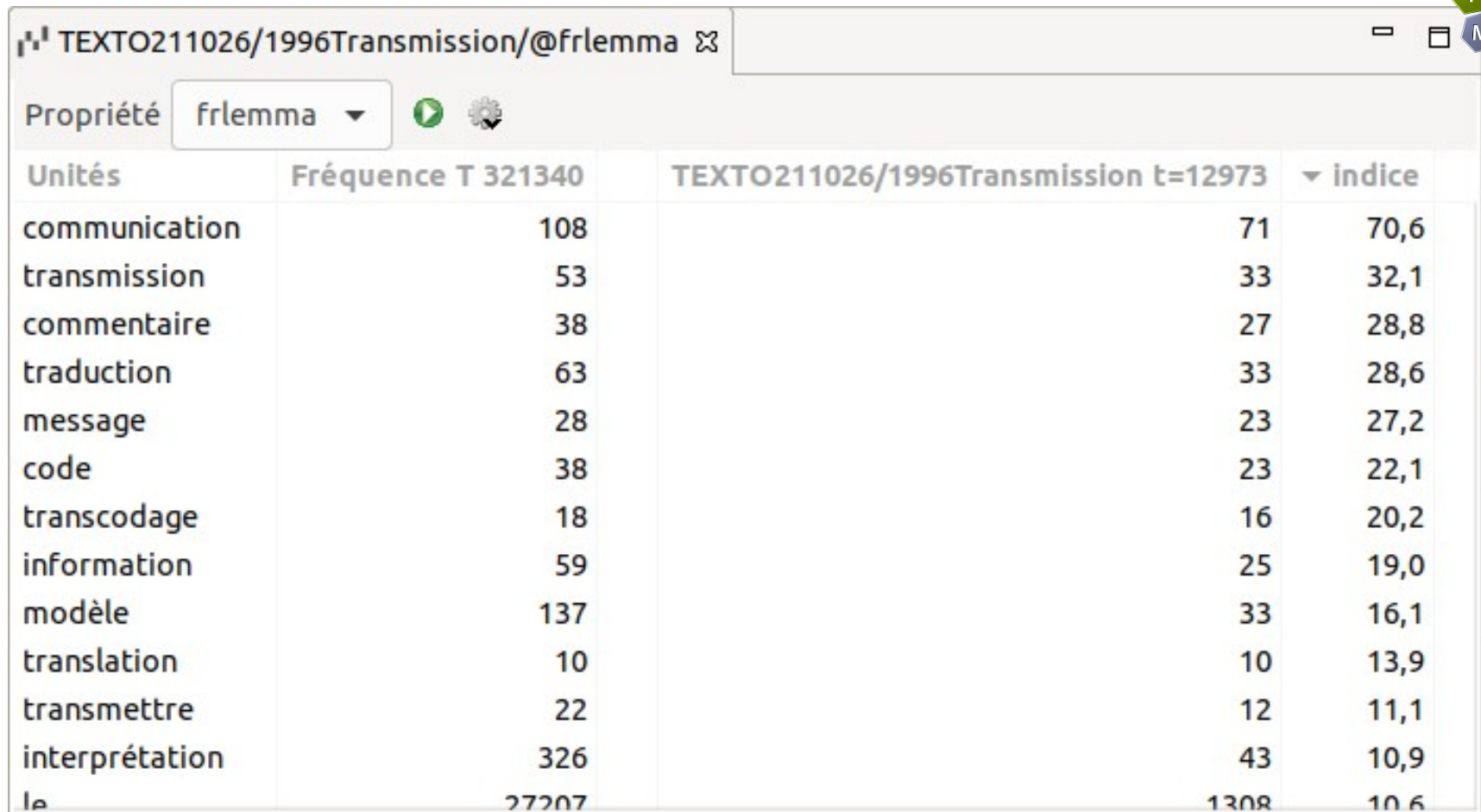
Sorties

Concordance

Liste de mots

Spécificités

Cartographie



Unités	Fréquence T 321340	TEXTO211026/1996Transmission t=12973	indice
communication	108	71	70,6
transmission	53	33	32,1
commentaire	38	27	28,8
traduction	63	33	28,6
message	28	23	27,2
code	38	23	22,1
transcodage	18	16	20,2
information	59	25	19,0
modèle	137	33	16,1
translation	10	10	13,9
transmettre	22	12	11,1
interprétation	326	43	10,9
le	27207	1308	10,6

*Ou réciproquement, quel vocabulaire est particulièrement utilisé dans le texte
« Communication ou transmission » (1996) ?*



Recherche

Infralexicale

Grammaticale

Supralexicale

Typo. & struct.

Intertextualité

Transversale

Contrastive

Sorties

Concordance

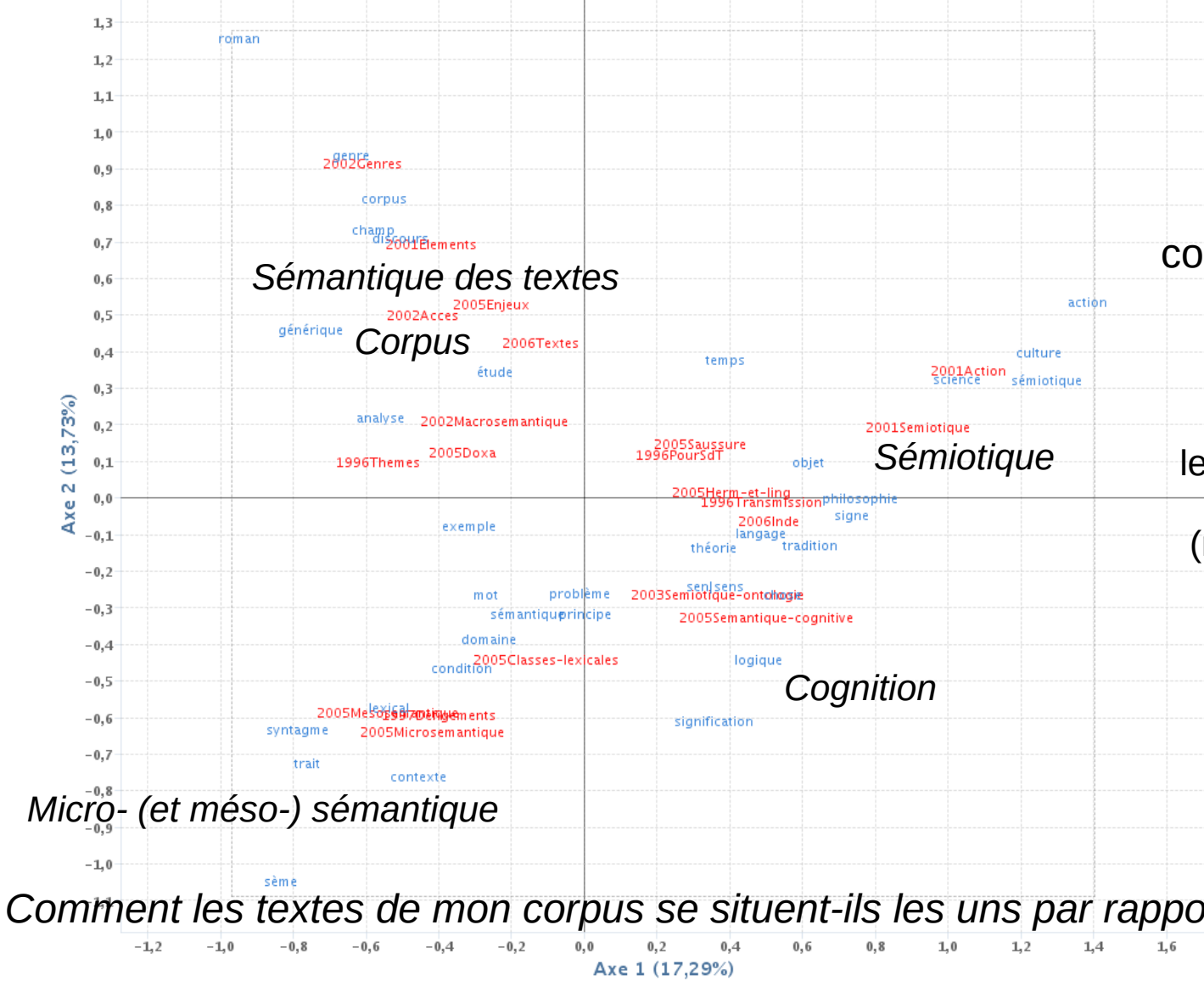
Liste de mots

Spécificités

Cartographie

Analyse factorielle des correspondances.

Les textes sont positionnés en fonction de leur profil d'emploi des mots (noms et adjectifs avec $f \geq 150$)



Comment les textes de mon corpus se situent-ils les uns par rapport aux autres ?

Usages

- Pédagogiques
 - rapprocher des « passages parallèles » qui s'éclairent mutuellement, se complètent
 - s'orienter dans ses lectures
 - des vues globales pour contextualiser les premières lectures : notions fréquentes
- Savants
 - retrouver un passage pour citation (l'édition Sittelle peut être un intermédiaire vers une édition de référence)
 - procéder à une lecture systématique (selon une requête explicite) : approfondissement d'un concept, vérification terminologique...
 - vue d'ensemble, nouvelles entrées de lecture, nouveaux observables
- Recherche « méta » à l'occasion de la conception de l'édition numérique et de l'affinement de la méthodologie
 - *sur la textualité*, notamment dans les écrits rastiériens : adéquation et limites de la représentation formelle, modélisation et fonctionnement des structures textuelles
 - *sur la sémantique* : peut-on traduire des isotopies en requêtes, est-ce qu'une classification type Reinert peut approcher empiriquement le concept d'isotopie (comment, pourquoi) ; approfondissement des affinités entre sémantique interprétative et textométrie, etc.

Illustration d'un parcours

- Pédagogique, mais pas forcément très loin non plus d'un parcours savant ?
- Point de départ : la famille de mots autour du radical *isotop*–



Requête [frlemma="isotope"]

text_id	Contexte gauche	Pivot	Contexte droit
2005Classes-lexicales	fiques : il est donc compatible voire	isotope	avec un grand nombre de sé
2005Classes-lexicales]. A palier lexical, les contextes	isotopes	d'un sémème conservent un
2005Classes-lexicales	ouver un grand nombre de contextes	isotopes	. C'est pourquoi l'on rencon
2005Mesosemantique	n / perfectif / cinq minutes seraient	isotopes	, à la différence de ? Il a mar

Requête [frlemma="isotopique"]

text_id	Contexte gauche	Pivot	Contexte droit
2005Mesosemantique	es sèmes, et qu'ainsi l'analyse	isotopique	permet dans son principe un t
2005Mesosemantique	hors contexte. 4. 3. Typologie	isotopique	des énoncés Les isotopies gé
2006Formes-semantiques	diffuses (isotopies, faisceaux	isotopiques) ou rhapsodiques (thèmes). U

On a essentiellement le nom, et pas isotopant par ex. (dans notre corpus).

isotope(s) : en relation d'isotopie (être isotope, les contextes isotopes d'un sémème)

isotopique(s) : relatif à l'isotopie (analyse ~, typologie ~, faisceaux ~)

```
Sortie standard
Index de <"isotop.*"%c>, propriété @frlemma, dans le corpus TEXT0211026...
3 items pour 248 occurrences.
Sous-corpus "isotop.*"%c de TEXT0211026/<"isotop.*"%c>@frlemma ≤321 340 /321 340 <"isotop.*"%c>...
Terminé.
Concordance de <[frlemma="isotope"]> dans le corpus TEXT0211026/"isotop.*"%c...
4 occurrences.
Concordance de <[frlemma="isotopique"]> dans le corpus TEXT0211026/"isotop.*"%c...
3 occurrences.
```

Mots de la famille d'isotopie et examen des contextes des deux formes adjectivales.

Tri alphabétique



Requête [frlemma="isotopie"]

text_id	Contexte gauche	Pivot	Contexte droit
2006Borges	communs, pour renforcer l'	isotopie	dominante et la plus valori
2006Borges	mais hiérarchiquement inf	isotopie	dominée mais hiérarchique
2001Bette	on ne doit pas masquer les	isotopies	dominées, car la poly-isoto
2001Bette	euls valides. En général, les	isotopies	dominées et peu denses di
2006Borges	nœud syntaxique ; soit une	isotopie	entre deux nœuds syntaxiq
2001Semiotique	a phrase. Par exemple, une	isotopie	est instituée par la récurre
2005Mesosemantiq	rces, puisque le concept d'	isotopie	est issu, indirectement, du
1997Defigements	e traits par présomption d'	isotopie	est un facteur de reséman
2006Formes-semanti	ne se réduit pas à une seule	isotopie	et consiste en un faisceau c
1996Themes	end de la dominance entre	isotopies	et du parcours interpréti
2005Mesosemantiq	me ordre de phénomènes (isotopies	et faisceaux d'isotopies), n
1996Concepts	que lexicale, la théorie des	isotopies	et l'analyse narrative. Dans
2005Microsemantiq	eutique, la présomption d'	isotopie	et l'effet de pertinence qui

Page 5 / 9 Texte 2001Semiotique

2001Semiotique - 5

une dimension herméneutique de la connaissance.

La sémiotique générale et comparée et les sciences herméneutiques.- Les sémiotiques globales sont des philosophies (cf. Locke, Peirce, Apel) ; sans doute doivent elles d'ailleurs à leur nature philosophique leur caractère global. En revanche, les sémiotiques scientifiques ont des objets régionaux : langues, images, musique, etc.

Une sémiotique générale ne peut être que fédérative ; elle définit le champ où la linguistique, l'iconologie, la musicologie et les autres sciences sémiotiques procèdent à leurs échanges pluridisciplinaires.

Si la linguistique se définit comme la sémiotique des langues, la sémiotique discursive se confond avec la linguistique du texte. Cette évidence a été longtemps obscurcie parce que diverses linguistiques restreintes se cantonnaient au palier de la phrase, et déléguaient l'étude de la textualité à diverses disciplines, comme la pragmatique conversationnelle, la poétique ou la sémiotique discursive. Mais une sémiotique discursive autonome ne serait sans doute que " l'envers complice " d'une linguistique restreinte. Outre qu'il faut rappeler les sciences du langage, on ne peut maintenir une frontière disciplinaire entre le texte et la phrase. Par exemple, une isotopie est instituée par la récurrence de sèmes (unités microsemantiques du palier inférieur au sémème) ; et pourtant elle peut s'étendre

Le retour au texte pour isotopie(s) comporte beaucoup plus d'occurrences (241). Je trie la concordance par ses contextes droits, ce qui me permet de repérer un contexte type définitoire (une isotopie est instituée par...). Un double-clic sur la ligne permet d'accéder à l'occurrence dans le contexte du texte entier.



Requête "isosémies?"%c

text_id	Contexte gauche	Pivot	Contexte droit
2005Mesosemantique	t perfectifs finaux, et il y a	isosémie	entre le relateur et le mor
2005Mesosemantique	(5) Étudions à présent les	isosémies	entre relateurs et grammé
2005Mesosemantique	mparfait (/ imperfectif /) l'	isosémie	sera rétablie. À l'inverse, u
2005Mesosemantique	Michèle cousit ... ; mais en	isosémie	avec l'imparfait : ex. Pend
2005Mesosemantique	heuristique de l'étude des	isosémies	. Ils devraient être complé
2005Mesosemantique) ni isotopie facultative, ni	isosémie	; ex. : Que inutilement Au r
2005Mesosemantique	, ni un énoncé. b) des	isosémies	, mais pas d'isotopie facul
2005Mesosemantique	acultative, mais rupture d'	isosémies	; ex. : Le train disparu, la g
2005Mesosemantique	isotopie facultative et des	isosémies	; ex. : Le signal vert indique
2005Mesosemantique	es isotopies obligatoires (isosémies), si l'on admet du moins q
2005Mesosemantique	b. En d'autres termes, les	isosémies	ont un rôle secondaire dai
2005Mesosemantique	tiques qui participent aux	isosémies	sont rapidement oubliés.
2006Borges	admire chez Lugones, les	isosémies	ou accords syntaxiques er

1 - 40 / 40

2005Mesosemantique - 5

de la présente limite de la phrase une isotopie peut s'étendre sur deux morphèmes, sur deux mots, sur un paragraphe, sur tout un texte (cf. infra). On peut distinguer entre les isotopies induites par la récurrence d'un trait spécifique (ex. : / inchoatif / dans L'aube allume la source (Éluard), où ce trait est récurrent dans 'aube', 'allume' et 'source') ; ou d'un trait générique (comme / navigation / dans L'amiral ordonna de carguer les voiles). On distingue aussi celles qui sont prescrites par le système fonctionnel de la langue (les **isosémies** selon Pottier) et celles qui sont facultatives parce ce qu'elles relèvent d'autres systèmes de normes (et pour lesquelles nous conserverons le terme général d'isotopies).

Les accords, de la dépendance à la concordance

Les accords [13] présentent les exemples d'**isosémie** les plus simples, mais aussi les plus trompeurs. En apparence, la description des phénomènes d'accord comme des **isosémies** semble un retour aux sources, puisque le concept d'isotopie est issu, indirectement, du traitement que Togeby propose de l'accord (et de la rection, cf. 1951, pp. 143 sq.). Cependant, si ces relations morpho-syntaxiques relèvent bien, selon Togeby, de l'étude du contenu, elles ne concernent que sa forme et non sa substance, et ne sont donc pas sémantiques, à proprement parler.

S'il est vrai que les accords peuvent manifester des **isosémies**, c'est indépendamment de toute considération morphologique. Pour montrer le caractère pro

Page 5 / Texte 2005Mesosemantique

Quant aux isosémie(s), la concordance fait apparaître que toutes les occurrences sont concentrées dans un seul texte (« Mésosémantique et syntaxe », publié 2005 sur Texto!) (et elle permet d'accéder à quelques passages denses). À une exception près : une occurrence dans le texte « L'hypallage et Borges » (2006), isolée et dotée d'une glose éclairante.

Requête Proprié

word	Fréq
interprétation	294
interprétatifs	72
interprétative	65
interprétatif	60
interprétatives	33
interprétations	30
Interprétation	2

t 556, v 7, fmin 2, fmax 294

Requête "interprétati.*"%c

Cooccurrent	Fréq	CoFréq	Indice	Distance moy
parcours	189	112	113	,6
production	55	30	29	2,3
l'	5948	339	21	2,6
suspens	16	10	11	4,0
stratégies	35	12	9	3,3
opérations	64	14	7	1,4
herméneutique	211	25	7	5,4
résultat	25	9	7	3,1
interprétation	294	30	7	5,7
moments	36	10	6	3,6
conditions	128	18	6	3,3
identification	38	10	6	5,8
foyers	7	5	6	2,0

t pivot 556, v cooc 159, t cooc 2812, T corpus 321340

Requête [frlemma = "interprétation"]

Cooccurrent	Fréq	CoFréq	Indice	Distance moy
l'	5948	272	36	2,0
production	55	30	36	2,3
suspens	16	10	13	4,0
interprétation	294	25	8	5,5
d'	3706	125	8	2,6
identification	38	9	7	5,8
immanent	12	6	7	7,7
et	6340	180	6	3,7
ses	357	23	6	2,7
assistance	6	4	5	2,0
départ	21	6	5	5,0
résultat	25	6	5	2,2
autonomes	3	3	5	8,0

t pivot 326, v cooc 115, t cooc 1704, T corpus 321340

Console

Sortie standard

Index de <"interprétati.*"%c>, propriété @word, dans le corpus TEXT0211026...

7 items pour 556 occurrences.

Cooccurrents de <"interprétati.*"%c>, propriété @word 9 9, $\geq 2 \geq 2 \geq 2.0$, dans le corpus TEXT0211026...

159 cooccurrents pour 556 occurrences du pivot.

Cooccurrents de <[frlemma = "interprétation"]>, propriété @word 9 9, $\geq 2 \geq 2 \geq 2.0$, dans le corpus TEXT0211026...

115 cooccurrents pour 326 occurrences du pivot.

Exploration autour de l'interprétation : plusieurs centaines d'occurrences (326 noms et 230 adj.), la cooccurrence (mots attirés au voisinage) opère une synthèse des contextes.



Requête `[frlemma = "interprétation"]`

Cooccurrent	Fréqu	CoFréqu	Indice	Distance moy
l'	5948	272	36	2,0
production	55	30	36	2,3
suspens	16	10	13	4,0
interprétation	294	25	8	5,5
d'	3706	125	8	2,6
identification	38	0	7	5,8

t pivot 326, v cooc 115, t cooc 1704, T corpus 321340

2005Semantique-cognitive - 3

(cf. Saranskis, 1991 pour l'herméneutique formelle ; pour l'herméneutique matérielle, l'auteur, 2001, ch. 4).

Le régime herméneutique du symbole logique est celui du **suspens** : le **suspens de l'interprétation** est le moyen de déployer l'effectivité du calcul. Mais ce **suspens advient entre deux phases où l'interprétation** est possible sinon prescrite. En revanche, le régime herméneutique du signe linguistique est ordinairement celui de l'interprétation compulsive : on ne peut l'isoler et l'identifier que par un faisceau d'hypothèses [11] qu'il est impossible de suspendre, tant dans la description sémantique que dans le traitement psychique. Alors que dans le calcul l'interprétation du symbole est momentanément exclue, celle du signe linguistique

Page 3 / Texte 2005Semantique-cognitive

Requête `(([frlemma = "interprétation"] []* @[word="suspens"]) | (@[word="suspens"] []* [frlemma = "interprétation"])) within 10`

text_id	Contexte gauche	Pivot	Contexte droit
1996PourSdT	que et la théorie sémantique de l'	interprétation. Une question cruciale et restée en [suspens]	commande cependant l'unificati
1996Terme	tiquent chacune à leur manière le	[suspens] de l'interprétation	. La première la réifie, en norma
1996Terme	cusées. La solitude du signe et le	[suspens] de l'interprétation	vont de pair. Or nous estimons qu
1996Transmission	des langages formels est celui du	[suspens], car leur interprétation	peut se déployer après le calcul, l
1996Transmission	s textes ne connaissent jamais le	[suspens] de l'interprétation	. Elle est compulsive et incoercibl
2004Ontologies	le sont suscitées par le refus de l'	interprétation, qui se traduit en premier lieu par le [suspens]	du contexte. Ce refus renvoie la s
2005Semantique-cognitive	e du symbole logique est celui du	[suspens] : le suspens de l'interprétation	est le moyen de déployer l'effecti
2005Semantique-cognitive	yer l'effectivité du calcul. Mais ce	[suspens] advient entre deux phases où l'interprétation	est possible sinon prescrite. En r
2005Semantique-cognitive	le sémiotique caractérisée par le	[suspens] de l'interprétation	. Il a certes eu le mérite d'insister

1 - 9 / 9

Étude de l'attirance entre suspens et interprétation en revenant aux contextes (double-clic).

Pertinence attendue de cette approche pour ce corpus

- Les caractéristiques du corpus
 - Volumétrie (une centaine de textes)
 - L'unité d'écriture et de pensée de l'auteur
 - Style d'écriture par évocation, rapprochements, reprises : relations intertextuelles denses, sémantique diffuse
- Un instrument d'observation
 - Ne fournit pas des réponses, mais des faits à observer de façon critique (occurrences, répétitions, variations, répartition...)
 - Complémentarité à la lecture, aux travaux de synthèse (A/R textes ↔ synthèses)
 - Des humanités numériques qui ne sont ni coquetterie, ni concurrence à la réflexion de recherche.

Les données ne sont pas données

« un corpus n'est pas un ensemble de données : comme toujours dans les sciences de la culture, **les données sont faites de ce que l'on se donne** (cf. n. 2 p. 96) et le point de vue qui préside à la constitution d'un corpus conditionne naturellement les recherches ultérieures. »

F. Rastier, *Arts et sciences du texte* (2001), chapitre III Philologie numérique, p. 86

=> Pas d'évidence ni de neutralité.

Les données ne sont pas données

2. Rien ne nous est donné

[...] Trois hypothèses ont [...] cours.

(i) **Les données sont ce qu'on vous donne** : [...] les fournisseurs d'accès [...] mettent à disposition des données : mais [...] ce dont on ne connaît pas la source et le mode d'acquisition est ininterprétable.

(ii) **Les données sont ce qu'on vous prend**. [...] Mais en général nous l'ignorons et elles sont données à d'autres (pour des profilages marchands, policiers ou politiques).

(iii) **Les données sont ce qu'on se donne** – ce devrait du moins être le cas en linguistique de corpus et dans l'ensemble des sciences de la culture. En effet, le travail d'analyse commence toujours par la qualification des données comme telles, ne serait-ce que par la qualification et la délimitation du corpus d'étude.

Le choix des textes

- **Un continuum des sujets** : la linguistique s'élargit aux sciences de la culture – aux arts, à la littérature ; elle croise des questions de société (éducation, racisme, déconstruction, etc.)
- **Les limites artéfactuelles** (disponibilité matérielle ou juridique) : à expliciter (objectif d'interprétabilité plutôt que de perfection)
- **L'engagement interprétatif de l'utilisateur du corpus** : que le projet propose une collection de textes, avec la possibilité de la redéfinir et l'ajuster. L'intertexte étant déterminant, le chercheur doit pouvoir le construire en fonction de sa problématique.

Des textes « vivants »

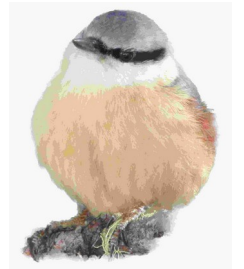
- L'idée (la sémantique interprétative) plutôt que le document (l'écriture de François Rastier)
- Une date de première publication, sans exclure des révisions ultérieures : une chronologie complexe
- Correction des coquilles

Un codage philologique équilibré ?

- Restituer la textualité par-delà les mots :
 - organisation du texte dans sa structuration (paragrapes, sections, listes), sa typographie (italiques, variation de police, retraits), ses hors-texte (schémas), son paratexte (bibliographie, notes, exergue...)
 - **non pas sophistication technique éperdue, mais pertinence pour l'analyse**
- **Éditer c'est interpréter** → s'arrêter avant la sur-interprétation.
Un critère : **s'en tenir aux marques visibles ?**
 - Ex. renvois et références bibliographiques
 - Ex. hésitation à utiliser la balise TEI <soCa11ed> qui induit une valeur dysphorique /rejet/ (« contains a word or phrase for which the author or narrator indicates a disclaiming of responsibility ») ou à s'en tenir à un codage plus générique <q> (citation, discours rapporté).
- Homogénéiser des hétérogénéités « invisibles »
 - variantes d'apostrophes, de guillemets, de tirets

Récapitulatif de l'exposé en quelques propositions

- Une **collection numérique de textes de Sémantique interprétative**
 - Diffusée de façon **ouverte**
 - Conçue et outillée pour des lectures **textométriques**
 - En début de construction (l'aventure commence !)
 - Avec des limites et des choix explicités
- Pour répondre à différents enjeux
 - **Pédagogiques** : « entrer » dans l'écrit et les idées de François Rastier
 - Connaissance et approfondissement de la **théorie** sémantique interprétative : lire autrement, re-lire, observer les faits, approche systématique, vue d'ensemble
 - Recherche sur la **modélisation** de la textualité et du fonctionnement sémantique



Sittelle