



HAL
open science

Évolution des emplois des marqueurs de topiques de discours dans Le Figaro de la fin du XIXe et du début du XXIe siècles

Michel Charolles, Sascha Diwersy, Denis Vigier

► **To cite this version:**

Michel Charolles, Sascha Diwersy, Denis Vigier. Évolution des emplois des marqueurs de topiques de discours dans Le Figaro de la fin du XIXe et du début du XXIe siècles. *Langages*, 2017, Du quantitatif au qualitatif en diachronie: prépositions françaises, 206, pp.85-104. 10.3917/lang.206.0085. hal-01907073

HAL Id: hal-01907073

<https://hal.science/hal-01907073>

Submitted on 21 Apr 2020

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

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

Evolution des emplois des marqueurs de topiques de discours dans *Le Figaro* de la fin du XIX^e et du début du XXI^e siècles

Evolution of the use of topic of discourse markers in *Le Figaro* of the end of the XIXth and of the beginning of the XXIst century

Michel Charolles
Université de Paris III
UMR-CNRS LATTICE ENS Paris

Sascha Diwersy
Université Paul Valéry Montpellier 3
UMR-CNRS PRAXILING

Denis Vigier
Université Lumière Lyon 2
UMR-CNRS ICAR

Résumé

L'article porte sur l'usage des expressions *à propos de*, *en matière (de)*, *au sujet de*, *sur le plan (de)*, *au niveau (de)*, *dans le domaine (de)*, *dans le secteur (de)* et *quant à* comme marqueurs de topique de discours. Ces locutions présentent des différences aux niveaux morphosyntaxique et sémantique qui suggèrent qu'elles ne sont pas toutes grammaticalisées au même degré. C'est cette hypothèse que nous testons sur leurs emplois dans *Le Figaro* de la fin du XIX^e siècle et du début du XXI^e siècle. La comparaison et l'étude statistique de leurs évolutions fréquentielles montre que ces expressions peuvent être réparties dans des sous-ensembles distincts de marqueurs de topique de discours, qui connaissent des tendances d'enracinement divergentes.

Mots clés : Marqueurs de topique de discours, grammaticalisation, lexicométrie

Abstract

The present paper is concerned with the use of the prepositional phrases *à propos de*, *en matière (de)*, *au sujet de*, *sur le plan (de)*, *au niveau (de)*, *dans le domaine (de)*, *dans le secteur (de)* and *quant à* as markers of topics in ongoing discourse. The differences which hold between these expressions at the syntactic, semantic and discourse level suggest that they are not grammaticalized to the same extent. This hypothesis is tested against two samples of newspaper texts from the end of the 19th and the beginning of the 21st century. Our analysis, which is underpinned by the results of frequency comparison, shows that the expressions under investigation form distinct sets of markers, which are involved in divergent trends of entrenchment.

Key words: discourse topic markers, grammaticalization, lexicometrics

Introduction

L'article qui suit porte sur une sélection de marqueurs pragmatiques du français du type de *au sujet de*, *pour ce qui est de*, *quant à*, *à propos de*, *au niveau (de)*...., qui ont pour fonction de guider les

auditeurs/lecteurs dans la mise à jour des topiques des discours¹. Les appellations utilisées pour désigner ces marqueurs varient d'un auteur à l'autre : Porhiel (1998, 2003) parle d'indicateurs d'intérêt, Charolles (1997) et Porhiel (2004 et 2005) d'introducteurs de cadres thématiques, Anscombe (2006), Velghe (2015), ainsi que Velghe & Lahousse (2015) de marqueurs de thématisation. Dans la suite nous parlerons de marqueurs de topiques de discours, ce qui incite à les rapprocher des topiques de phrase dont ils diffèrent cependant (cf. Reinhart (1980 et 1981) et Van Dijk (1977) pour des études princeps).

Sans entrer dans les détails, nous considérerons que, si l'on peut en général sans difficultés affecter aux phrases (au moins catégoriques) un topique dont le prédicat dit quelque chose, il n'en va pas de même au niveau du discours (cf. Marandin, 1988, pour une étude pionnière). Dans le prolongement de Kuppevelt (1995), nous admettons que dans une suite de deux phrases (P1 et P2), P1 n'a aucun topique de discours. P2 apporte un contenu dont on s'attend à ce qu'il réponde à une question que l'on s'est posée ou que l'on aurait pu se poser à la suite de P1². Défini de cette façon, le topique de discours local³ correspond à l'orientation que prend le discours à chacune de ses avancées. Pour rendre compte de ces avancées, on a besoin de notions comme celles d'*introduction*, de *continuation*, de *changement* de topique de discours et de *digression*, qui correspondent à des enchaînements particuliers (sur ces notions cf. notamment Kuppevelt, 1995 ; Goutsos, 1996 ; Fraser, 2009a et 2009b).

Dans la première partie nous présentons la liste des marqueurs sur laquelle nous avons travaillé. En nous appuyant sur différents traits qui peuvent rapprocher ou différencier ces marqueurs aux niveaux morphosyntaxique et sémantique, nous avançons une hypothèse pour l'étude de leur grammaticalisation⁴, aspect qui est au centre de notre travail. Dans la deuxième partie nous présentons les résultats de l'annotation de ces marqueurs sur un échantillon de huit cents occurrences, prélevé aléatoirement sur l'un des corpus « complémentaires » de Presto (voir Diwersy, Falaise, Lay & Souvay dans ce numéro). Cet échantillon réunit des numéros du quotidien *Le Figaro* de la fin du XIX^e et du début du XXI^e siècle et compte presque 70 millions de mots. Dans la troisième partie nous rapportons les résultats d'un calcul des spécificités (Lafon, 1980 ; Lebart & Salem, 1994) mettant en évidence, pour chacune des expressions qui nous intéressent, des différences fréquentielles (en termes de sur-emploi / sous-emploi) entre les échantillons comparés. Dans la conclusion, nous confrontons nos résultats avec ceux présentés dans les analyses (non diachroniques) parues sur les marqueurs étudiés dans cet article et nous envisageons quelques pistes pour aller plus loin sur le sujet.

1. Quelques marqueurs de topique de discours et hypothèses sur leur formation et leur évolution

Les données sur corpus que nous avons rassemblées pour cette étude portent sur les expressions suivantes :

¹ *Stricto sensu*, il convient de distinguer les expressions plus ou moins figées comme *à propos de*, *en matière (de)*, ... des Marqueurs de Topique de Discours proprement dits qui sont des SP : *à propos de Paul*, *en matière linguistique*, ... et qu'il faudrait noter *à propos de SN/SAdj* etc. Pour éviter le recours à des formulations fastidieuses, nous confondrons à dessein la notation de l'expression avec celle du Marqueur, sauf *distinguo* nécessaire dans l'argumentation.

² Sur ces questions dites « Question Under Discussion » (QUD), cf. notamment Roberts (2004).

³ Ce topique de discours local ne correspond pas au topique de discours plus global (macrostructurel) de Van Dijk (1977).

⁴ Par grammaticalisation on entend l'évolution linguistique qui soit change des lexèmes pleins en mots grammaticaux, soit rend des éléments grammaticaux encore plus grammaticaux. Depuis les travaux de Traugott (Traugott, 1989, 1995 ; Traugott & König, 1991 ; Traugott & Dasher 2002), la notion de grammaticalisation est proche de la *pragmaticalisation* (Lamiroy & Charolles 2005): les mots passant à un sens plus abstrait, renvoient moins au monde concret (p.ex. relations spatiales) qu'à l'organisation de celui-ci par les locuteurs (p.ex. connecteurs et marqueurs de discours) et à leurs croyances (p.ex. modaux épistémiques).

- (i) *à propos de, en matière (de), au sujet de, sur le plan (de), au niveau (de), dans le domaine (de), dans le secteur (de), quant à*

Toutes ces expressions se prêtent, en français contemporain, à des emplois comme marqueurs de topique de discours illustrés par les exemples suivants (tirés du corpus *Le Figaro*). Le critère sur lequel nous nous sommes fondés pour ranger ces emplois parmi les marqueurs de topique de discours est le suivant : l'expression (en gras) peut commuter⁵ avec les syntagmes (*en + pour*) *ce qui concerne, concernant*.

- (1) Le président palestinien implore l'Union européenne de prendre le relais des Etats-Unis pour relancer la négociation. (...) **A propos du** cargo chargé d'armes arraisonné en mer Rouge, Arafat s'élève contre la version israélienne selon laquelle le chargement était destiné à l'Autorité palestinienne.
- (2) Bernard Thibault réclame une sortie par le haut du système des multiples smic. Il faut, explique-t-il, arrêter de penser que le 'coût' social est la seule et unique variable d'ajustement de toute politique économique. **En matière de** retraite, il propose de cumuler les droits dès l'âge de 18 ans.
- (3) La grande vision hyperbolique à la manière de Virgile et de Dante n'est pas son domaine. **A ce sujet**, le quatrième volume contient un texte étrange, énigmatique à la date du 8 août 1988.
- (4) Les marchés anticipent ensuite une période de prudence tant qu'une image claire ne se dégagera pas des résultats et prévisions des sociétés. **Sur le plan** macro-économique, les nouvelles plus rassurantes sur le front du chômage et de la consommation devraient soutenir l'économie.
- (5) [II] assure que les programmes européens essaient d'encourager la distribution et la promotion des films. **Au niveau de** l'exploitation, une aide pour toutes les salles diffusant des films européens est demandée par les professionnels.
- (6) Les ministres du Tourisme successifs ont répandu l'idée que la France était la première destination mondiale (...). Pourtant, le chiffre d'affaires ne suit pas. **Dans ce domaine**, la France n'est que troisième derrière les Etats-Unis et l'Espagne.
- (7) Dans l'automobile, sa préférence va à Volkswagen (...) **Dans le secteur** bancaire, il cite UBS en raison de la solidité de son 'private banking' (banque privée) (...)
- (8) Le classique triptyque 'informer, éduquer, divertir' ne suffit plus à identifier clairement la télévision publique. Si l'on en croit les mesures d'audience, les téléspectateurs français font plus confiance à la télévision commerciale lorsqu'ils souhaitent être informés ou divertis. **Quant à** la mission éducative, n'est-elle pas autant assurée par certaines chaînes thématiques privées que par celles créées dans le giron du secteur public ?

Les locutions prépositionnelles regroupées *supra* sous (i) présentent des différences *syntaxiques, morphologiques et sémantico-référentielles* qui autorisent à les regrouper en sous-classes.

- Parmi les expressions où figure un N, seules celles construites avec *plan, niveau, domaine, secteur, matière* présentent une alternance de figement sur le modèle : (i) *Prép. (+ Dét.) + N + de* (suivi d'un SN); (ii) *Prép. + (Dét.) + N* (suivi d'un SAdj)

Tableau 1

Alternance de figement pour les locutions construites avec les N *plan, niveau, domaine, secteur, matière*

Nom figurant dans la Prép / Loc Prép	LOC 1 <i>Prép. (+ Dét.) + N + de</i>	LOC 2 <i>Prép. (+ Dét.) + N</i>
<i>propos</i>	<i>à propos de [la stratégie]</i>	<i>*à propos [stratégique]</i>

⁵ Moyennant quelques ajustements morphosyntaxiques liés aux contraintes d'emploi de l'expression substituée.

<i>matière</i>	<i>en matière de [stratégie]</i>	<i>en matière [stratégique]</i>
<i>sujet</i>	<i>au sujet de [la stratégie]</i>	<i>*au sujet [stratégique]</i>
<i>plan</i>	<i>sur le plan de [la stratégie]</i>	<i>sur le plan [stratégique]</i>
<i>niveau</i>	<i>au niveau de [la stratégie]</i>	<i>au niveau [stratégique]</i>
<i>domaine</i>	<i>dans le domaine de [la stratégie]</i>	<i>dans le domaine [stratégique]</i>
<i>secteur</i>	<i>dans le secteur de [la stratégie]</i>	<i>dans le secteur [stratégique]</i>

- Syntactiquement, ces expressions présentent des différences constructionnelles. Si l'on choisit pour critère la catégorie morphosyntaxique de la tête de leur régime, on aboutit à un regroupement en deux grandes classes, la seconde pouvant elle-même être raffinée en deux sous-classes. Dans les tableaux suivants, nous proposons d'ajouter les expressions *concernant*, *pour ce qui concerne*, *en ce qui concerne*, et *pour ce qui est de*, afin d'équilibrer la comparaison entre les deux classes de marqueurs. Elles ne seront pas prises en compte dans l'étude sur corpus, qui ne portera que sur les marqueurs figurant dans la liste (i) *supra*.

Tableau 2

Premier classement des marqueurs sur le critère syntaxique de leur complémentation

	Préposition / locution prépositionnelle		Régime			
			SN	Npr ⁶	ADJ	Vinf
1	<i>Sur le plan</i>	<i>de</i>	+	-	-	-
		\emptyset	-	-	+	-
	<i>Au niveau</i>	<i>de</i>	+	-	-	-
		\emptyset	-	-	+	-
	<i>Dans le domaine</i>	<i>de</i>	+	-	-	-
		\emptyset	-	-	+	-
	<i>Dans le secteur</i>	<i>de</i>	+	-	-	-
		\emptyset	-	-	+	-
<i>En matière</i>	<i>de</i>	+	-	-	-	
	\emptyset	-	-	+	-	
2a	<i>Au sujet de</i>		+	+	-	-
	<i>A propos de</i>		+	+	-	-
	<i>Concernant</i>		+	+	-	-
	<i>Pour ce qui concerne</i>		+	+	-	-
	<i>En ce qui concerne</i>		+	+	-	-
2b	<i>Quant à</i>		+	+	-	+
	<i>Pour ce qui est de</i>		+	+	-	+

- On peut raffiner ce classement en introduisant un critère de variabilité morphosyntaxique ayant une contrepartie sémantico-référentielle : l'expression présente-t-elle ou non une variabilité lui permettant d'inclure une reprise anaphorique ?

Tableau 3

Possibilité pour certains marqueurs d'inclure une reprise anaphorique

Prép / Loc Prép	Reprise anaphorique
<i>sur le plan (de)</i>	<i>sur ce plan</i>
<i>au niveau (de)</i>	<i>à ce niveau</i>

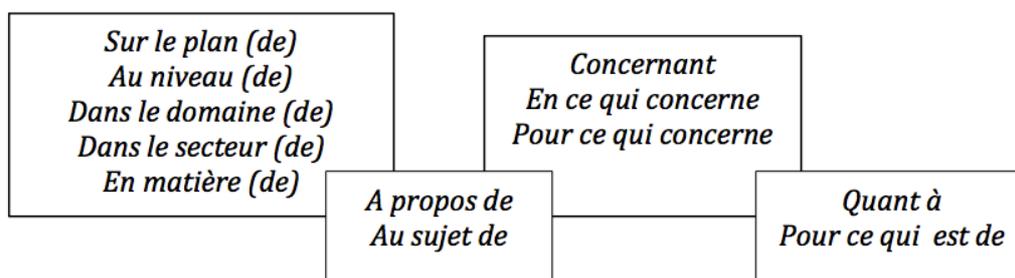
⁶ Nom propre.

<i>dans le domaine (de)</i>	<i>dans ce domaine</i>
<i>dans le secteur (de)</i>	<i>dans ce secteur</i>
<i>en matière (de)</i>	<i>en (cette +la) matière</i>
<i>au sujet de</i>	<i>à ce sujet</i>
<i>à propos de</i>	<i>à ce propos</i>
<i>concernant</i>	∅
<i>pour ce qui concerne</i>	∅
<i>quant à</i>	∅
<i>pour ce qui est de</i>	∅

On aboutit finalement au regroupement suivant :

Tableau 4

Regroupement en quatre sous-classes des expressions prises en compte dans les tableaux 2 et 3



Ce tableau donne une photographie de quelques-unes des formes à même de fonctionner en français contemporain comme des marqueurs de topique de discours. Ces formes ne sont visiblement pas toutes aussi spécialisées dans ce genre de fonction. Les expressions et locutions sans N comme *quant à* et *pour ce qui est de* que l'on a coutume – comme *concernant*, *pour/en ce qui concerne* – d'appeler *topicalisateurs* (Combettes & Prévost, 2001, 2003) semblent majoritairement, voire exclusivement, employées en français contemporain comme des marqueurs de topique de discours (Prévost, 2003, 2007 et 2010). Cela est assez évident avec *quant à* où l'on a du mal à imaginer aujourd'hui des usages comme **il m'a parlé quant à Paul*, *?? Marie estime quant à Paul qu'il est très serviable*. Il en va autrement avec *sur le plan (de)*, *au niveau (de)*, *dans le domaine (de)*, *dans le secteur (de)* et *en matière (de)* où les emplois non idiomatiques viennent facilement à l'esprit : *les scouts ont campé dans le domaine du château*; *il faut respecter le tracé sur le plan de l'architecte*; *la rivière est montée jusqu'au niveau du 1^{er} étage*; *ce mur est construit en matière végétale*; *il s'est implanté dans le secteur de X*.

Si l'on se tourne vers les dictionnaires, on observe que les expressions appartenant à la seconde classe du tableau 2 sont plus souvent recensées comme locutions figées (sauf *pour ce qui est de* et *en ce qui concerne*) que celles du premier groupe. De même pour leur datation : les données fournies par le *Trésor de la Langue Française informatisé (TLFi)* et *Le Dictionnaire Historique de la Langue Française (DHLF)* aussi bien que celles que nous avons recueillies sur *Frantext* font apparaître une opposition entre d'un côté *sur le plan (de)*, *au niveau (de)*, *dans le domaine (de)*, *dans le secteur (de)* qui sont de formation récente, et de l'autre *quant à* attesté dès l'ancien français ; avec, entre les deux, *au sujet de*, *à propos de* et *en matière (de)*.

Tableau 5

**Locutions figées recensées (ou non) dans le *Trésor de la Langue Française informatisé (TLFi)* et *Le Dictionnaire Historique de la Langue Française (DHLF)*
Date de première attestation dans Frantext**

	Expression notée comme locution		Date de 1 ^{ère} attestation	
	<i>TLFi</i>	<i>DHLF</i>	Dans <i>TLFi</i> / <i>DHLF</i>	dans <i>Frantext</i> (235 000 000 mots) ⁷
<i>Dans le domaine (de)</i>	-	-	(1778) ⁸	1874 (SN), 1923 (Adj)
<i>Sur le plan (de)</i>	+	-		1934 (SN), 1925 (Adj)
<i>Au niveau (de)</i>	-	+		1942 (SN), 1948 (Adj)
<i>Dans le secteur (de)</i>	-	-		1958 (SN), 1960 (Adj)
<i>En matière (de)</i>	-	+		1633 (SN), 1790 (Adj)
<i>Au sujet de</i>	+	+	1617	1734 (SN/Npr)
<i>A propos de</i>	+	+	1580	1552 (SN/Npr)
<i>Quant à</i>	+	+	1269 / 1278	1347 (SN/Npr)
<i>Pour ce qui est de</i>	-	-		1593 (SN/Npr)
<i>Concernant</i>	+	+		1789 (SN/Npr)
<i>Pour ce qui concerne</i>	+	+	1690 / 1690	1601 (SN/Npr)
<i>En ce qui concerne</i>	+	+	1690 / 1690	1833 (SN/Npr)

Le fait que *en matière (de)* soit employé comme marqueur de topique depuis le XVII^e et XVIII^e siècle, bien avant *sur le plan (de)*, *au niveau (de)*, *dans le domaine (de)*, *dans le secteur (de)*, incite à le disjoindre de ces marqueurs dans le tableau 4 et à le rapprocher de *à propos de* et de *au sujet de* dont les premières attestations datent de la même période. Ce rapprochement se justifie sémantiquement dans la mesure où le N *matière* est plus proche des N *sujet* et *propos* avec qui il peut partager l'idée de 'contenu verbal', alors qu'il n'a aucun rapport évident avec les N spatiaux *plan*, *niveau*, *domaine*, *secteur*. À ce propos, il n'est pas sans intérêt de relever que *quant à*, qui est le marqueur de topique de discours le plus spécialisé, le plus ancien et le plus étudié (Choï-Jonin, 1997 ; Fløttum, 1999, 2003 ; Debaisieux, 2001 ; Anscombe, 2006 ; Prévost, 2003a, 2008, 2010 ; Lagae, 2003), se différencie radicalement des autres marqueurs par son opacité sémantique, au point que l'on peut sans risque faire l'hypothèse qu'il est le plus grammaticalisé, au contraire de *sur le plan de/Adj*, *au niveau de/Adj*, *dans le domaine de/Adj*, *dans le secteur de/Adj*, qui semblent beaucoup moins avancés sur ce chemin.

Tableau 6
Hypothèse sur les positions relatives occupées par les expressions étudiées
sur une échelle de grammaticalisation

<i>Sur le plan (de)</i> <i>Au niveau (de)</i> <i>Dans le domaine (de)</i> <i>Dans le secteur (de)</i>	<i>En matière (de)</i>	<i>A propos de</i> <i>Au sujet de</i>	<i>Quant à</i>
- grammaticalisées (formation) récente	←————→		+ grammaticalisées (formation) ancienne

⁷ Interrogation sur chacune des expressions étudiées précédées d'un point, pour sélectionner leurs emplois en tête de phrase où elles ont plus de chances d'être des marqueurs de topique de discours.

⁸ Après avoir noté que le N *niveau* a pris au XVIII^{ème} siècle un sens abstrait pour parler de la position sociale ou économique, le *DHLF* relève (p. 2376) que « les locutions *au niveau de* (av. 1778) suivent ces emplois abstraits. Elles sont devenues au XX^{ème} siècle des emplois passe-partout diffusés par le langage des ingénieurs, pour « dans le domaine de », « pour » et sont critiquées ». La datation « av. 1778 » ne vaut pas pour ces emplois contemporains qui nous intéressent.

2. Emplois des marqueurs de topique du discours dans *Le Figaro* de la fin du XIX^e et du début du XXI^e siècle. Étude sur corpus.

Pour tester l'hypothèse ci-dessus, nous proposons de nous tourner vers le corpus complémentaire de Presto qu'est le *Corpus diachronique de la presse française*, vaste échantillon de périodiques nationaux et régionaux qui se divise en deux parties couvrant d'une part la fin du XIX^e siècle, d'autre part le début du XXI^e siècle. L'ensemble du corpus a été catégorisé et lemmatisé au moyen de l'analyseur *TreeTagger* (Stein & Schmid 1995).

2.1. Constitution du Corpus

Pour constituer notre corpus de travail, nous avons sélectionné les articles du quotidien *Le Figaro*, ce qui nous donne deux sous-corpus, dont le tableau suivant présente un aperçu :

Sous-corpus	Années	Total numéros	Total mots-occurrences
<i>Le Figaro</i> XIX ^{ème} siècle [désormais LF-19]	1885, 1890, 1895, 1896	964	37.446.988
<i>Le Figaro</i> XXI ^e siècle [désormais LF-21]	2002	313	30.497.189
TOTAUX		1277	67.944.177

La plateforme web BTLC⁹ donne accès à différentes applications permettant d'explorer ce corpus ; ces applications sont implémentées par la boîte à outils PrimeStat dont les fonctionnalités suivantes ont été utilisées pour cette étude : création de sous-corpus et de partitions, définition d'expressions de requête complexes en termes de propriétés lexicales et de nombre de segments ciblés, extraction et manipulation de concordances KWIC, calcul de spécificités fréquentielles.

2.2. Fréquence d'occurrence des marqueurs de topique du discours étudiés.

Notre objectif consiste à nous faire une idée la plus précise possible de l'évolution des usages dans notre corpus des marqueurs :

à propos de, en matière (de), au sujet de, sur le plan (de), au niveau (de), dans le domaine (de), dans le secteur (de), quant à

Pour ce faire, nous avons comparé les fréquences relatives de ces expressions dans les deux sous-corpus considérés comme deux tranches synchroniques.

Nous avons d'abord extrait au sein des deux sous-corpus LF-19 et LF-21 les expressions recherchées [étape E1]. Parmi les occurrences extraites figuraient des séquences non pertinentes pour notre recherche. Afin d'évaluer quelle part occupait le « bruit » dans cette première étape, nous avons constitué pour chaque expression étudiée un échantillon de cent occurrences¹⁰ tirées aléatoirement parmi celles obtenues à l'issue de E1. Soit huit cents occurrences triées et codées en tout [étape E2].

Notre codage, outre la catégorie « bruit », a distingué deux critères de classement :

⁹ Plateforme développée (ainsi que PrimeStat) par S. Diwersy, d'abord à l'Université de Cologne, et depuis 2015 à l'Université de Montpellier 3.

¹⁰ 50 occurrences pour LF-19 et 50 pour LF-21

- (i) un critère sémantique : emploi *topique* [TOP] - *non topique* [Non TOP]
- (ii) un critère syntaxique : position intraprédicative *versus* extraprédicative (Guimier 1996 :5)

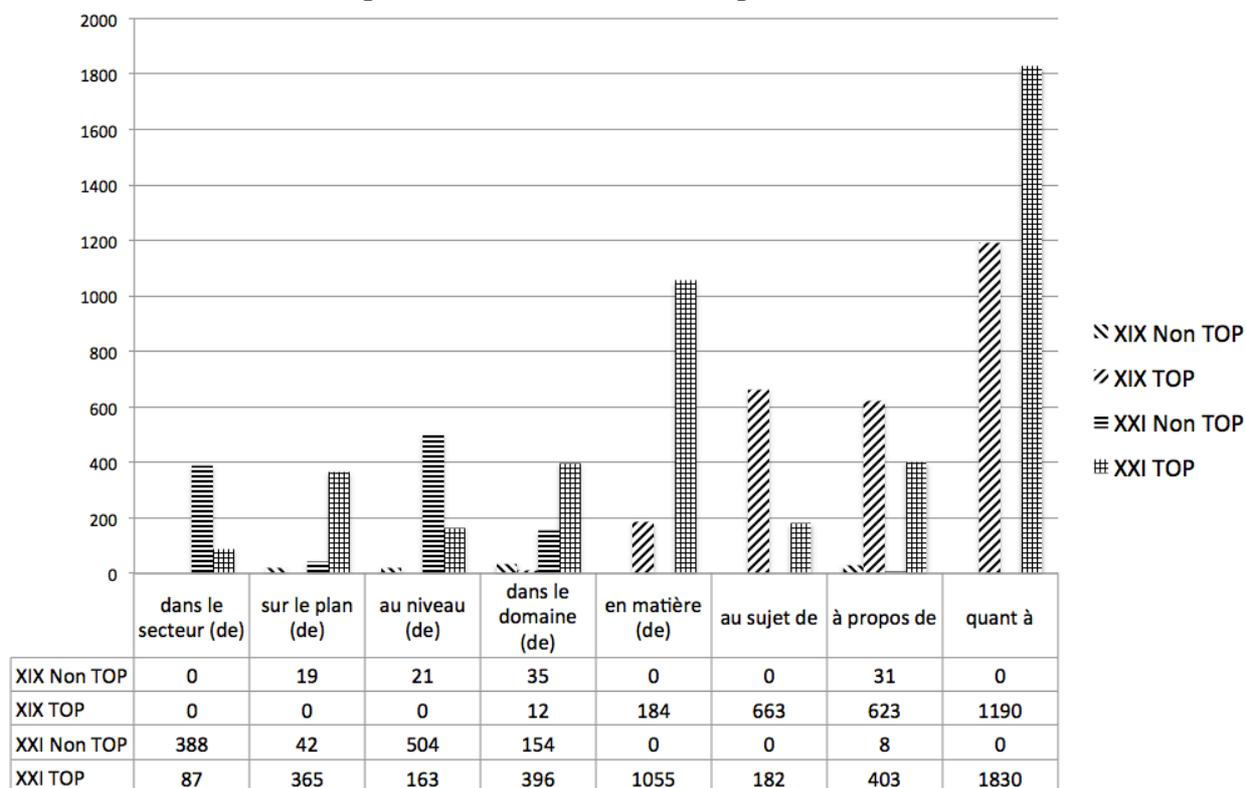
Le travail a été accompli par deux experts séparément et a été suivi d’une phase de discussion autour des occurrences dont le codage divergeait, jusqu’à accord des annotateurs.

Pour chaque marqueur ont été calculés les pourcentages recueillis par chacune des cinq sous-catégories distinguées et codées au sein de chaque échantillon. Ces pourcentages ont permis de quantifier par extrapolation la fréquence attendue des occurrences relevant de chacune des cinq sous-catégories dans LF-19 et dans LF-21. Ces fréquences théoriques ont ensuite été pondérées en fréquences relatives. Les valeurs ainsi obtenues ont été multipliées par 10^7 pour aboutir à des fréquences plus aisément interprétables et comparables [étape E3].

2.3. Résultats

Voici le diagramme des fréquences relatives calculées pour les catégories (sémantiques) TOP et Non-TOP après extrapolation des résultats obtenus à l’issue du codage des échantillons.

Diagramme 1
Diagramme des fréquences relatives (pour 10 000 000) des marqueurs pour les catégories sémantiques TOP et Non-TOP. Corpus LF-19 et LF-21



A la fin du XIX^e siècle, deux groupes de marqueurs se détachent nettement avec d’un côté, *dans le secteur (de)*, *sur le plan (de)*, *au niveau (de)*, *dans le domaine (de)* dont l’usage est rare (voire inexistant) et presque exclusivement cantonné à une valeur non-topique ; de l’autre côté, *en matière de*, *au sujet de*, *à propos de* et *quant à* qui s’utilisent dès cette période comme marqueurs de topiques de discours – que ce soit exclusivement pour *en matière (de)*, *au sujet de*, *quant à*, ou majoritairement pour *à propos de*.

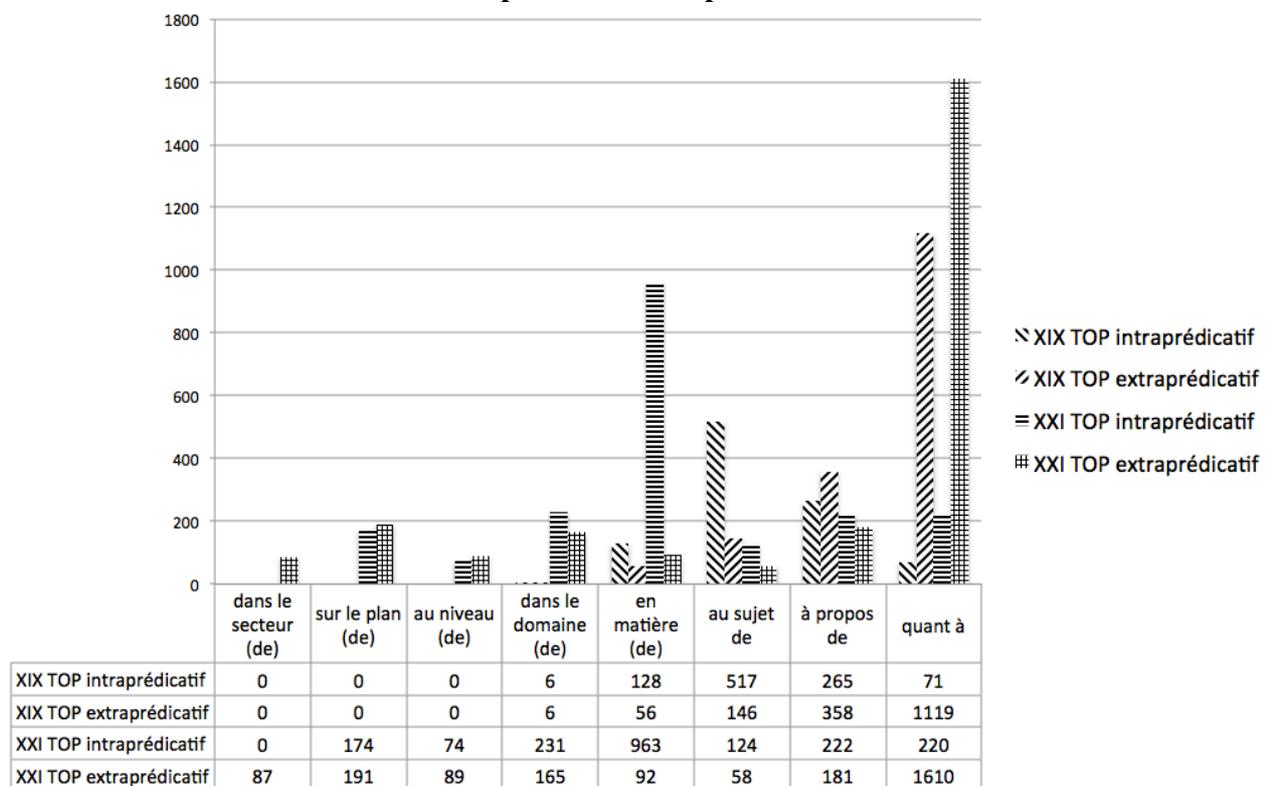
Au début du XXI^e siècle, on observe plusieurs changements. *Dans le secteur (de)*, *sur le plan (de)*, *au niveau (de)* présentent dans LF-21 des emplois comme marqueurs de topique de discours,

inexistants dans LF-19 ; quant à *dans le domaine (de)*, il est déjà employé avec cette valeur TOP dans LF-19 mais avec une fréquence relative très inférieure à ce qu'elle est au XXI^e siècle. Cette évolution est assez remarquable pour une si courte période et si l'on tient compte du fait qu'elle se traduit, pour les expressions en question, par un changement sémantique et fonctionnel important. *Quant à* reste de loin le plus topical des marqueurs : il présente plus d'emplois de ce type qu'au XIX^e siècle et devance nettement *au sujet de* et *à propos de* dont les emplois comme topique de discours déclinent du XIX^e au XXI^e siècle. *En matière (de)* qui, comme on l'a relevé, ne connaît dès le XIX^e siècle que des emplois comme marqueur de topique de discours, persiste dans cette exclusivité et son usage s'intensifie entre la fin du XIX^e siècle et le début du XXI^e siècle.

Voici un second diagramme qui présente les fréquences relatives des expressions étudiées, selon qu'elles occupent une position intraprédicative ou extraprédicative dans leurs emplois topicaux.

Diagramme 2

Fréquences relatives (pour 10 000 000) des marqueurs pour les catégories TOP intraprédicatif et TOP extraprédicatif. Corpus LF-19 et LF-21



On retrouve le contraste entre *dans le secteur (de)*, *sur le plan (de)*, *au niveau (de)*, *dans le domaine (de)* et les autres marqueurs, les premiers n'émergeant comme TOP que dans la seconde période. Les fréquences pour *dans le secteur (de)* et *au niveau (de)* sont particulièrement faibles, ce qui tend à les distinguer des deux autres qui, sous ce rapport, s'approchent des valeurs fréquentielles de *à propos de*.

Voici deux exemples d'emploi de *dans le domaine (de)* respectivement intraprédicatif (9) et extraprédicatif (10) :

- (9) Dans nombre de cas, les journaux n'avaient-ils pas signalé l'ingérence quotidienne du président du Conseil **dans le domaine des affaires extérieures**, ses entretiens fréquents avec tels ou tels membres du corps diplomatique étranger ? (LF-19)
- (10) On est bien entendu très loin de l'explosive affaire Enron. Mais l'on sent bien que la crainte cachée de tous est qu'un équivalent ne se produise en France. Or **dans le domaine de l'honnêteté**, il n'y a pas de sécurité possible. (LF-21)

Quant à présente beaucoup d’emplois extraprédicatifs (11) mais il garde des emplois intraprédicatifs à la fin du XIX^e et au début du XXI^e, comme (12), quoique dans une faible proportion :

- (11) Le texte est de notre collaborateur M. Jules Richard que ses connaissances des choses militaires désignaient tout naturellement pour la partie littéraire. **Quant à M. Détaille**, il a résolu le problème de faire revivre les types de l’armée sans tomber dans la monotonie. (LF-19)
- (12) La compétence ‘teutonne’ de notre historienne est à ce point considérable qu’on lui pardonne volontiers ces petites lacunes **quant aux documents de la francité**. Ajoutons que vichystologue, elle n’est pas nécessairement vichystophage, performance qui n’est pas si fréquente par le temps qui court. (LF-21)

Au sujet de et *à propos de* ont des profils assez différents : avec *au sujet de*, les emplois intraprédicatifs l’emportent sur les extraprédicatifs aux deux périodes. *A propos de* a nettement plus d’emplois extraprédicatifs que *au sujet de*, sur les deux périodes, mais ces emplois ne sont pas majoritaires au XXI^e siècle. Le profil de *en matière (de)* est à la fois différent de celui de *au sujet (de)* et de *à propos (de)* mais aussi de *quant à*. *En matière (de)* présente en effet beaucoup plus d’emplois intraprédicatifs (comme (13)) qu’extraprédicatifs (comme (14)) à chaque période et cet avantage tend à s’accroître au XXI^e siècle :

- (13) Après Denis Olivennes, le patron opérationnel de Canal + qui s’exprimait devant le Parti socialiste il y a deux mois, et Pierre Lescure qui, en décembre dernier, le martelait à qui voulait l’entendre, le voici qui réclame, à son tour, l’“équité” devant les obligations de production **en matière de cinéma**. (LF-21)
- (14) Espérons qu’on reviendra sur cette chinoiserie. Mais que voulez-vous, **en matière d’avancement**, on individualise, on ne généralise pas. (LF-21)

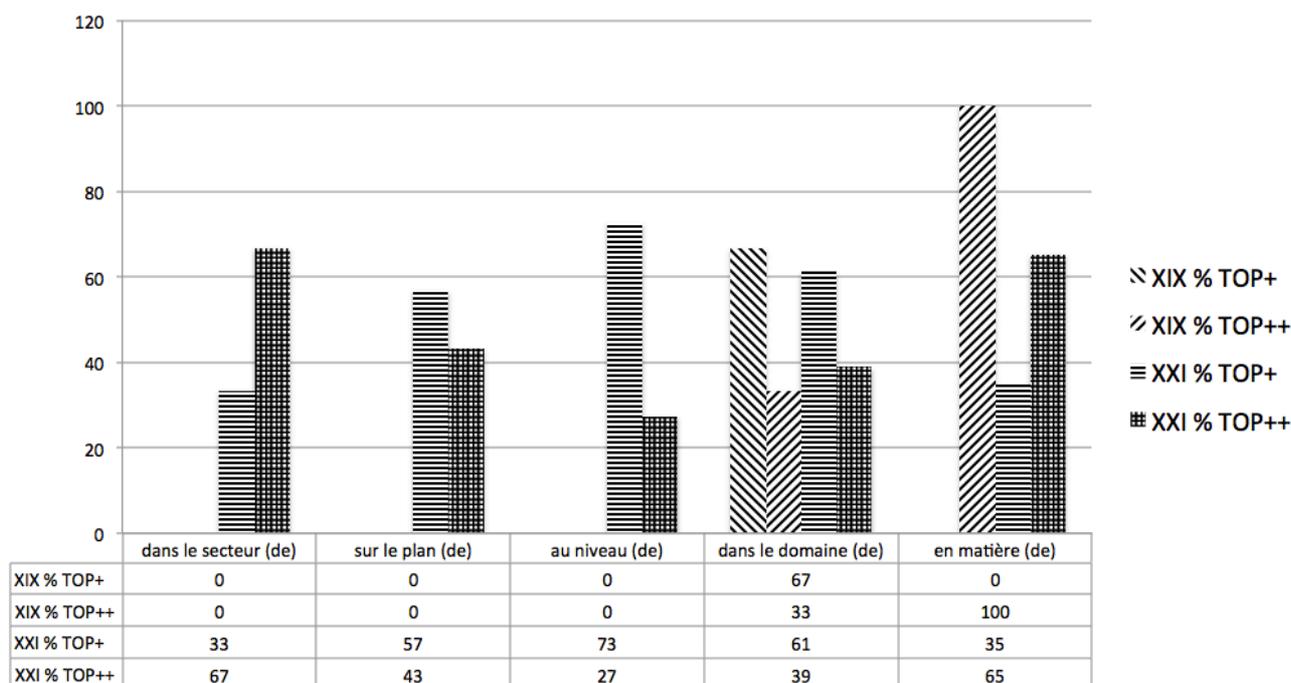
Si *en matière (de)* et *quant à* ont nettement plus d’emplois comme TOP au XXI^e siècle que tous les autres, cette suprématie n’a pas la même origine : elle tient, avec *quant à*, à la prééminence des emplois extraprédicatifs et, avec *en matière (de)*, à celle des emplois intraprédicatifs.

Le diagramme ci-dessous présente les résultats d’un test que nous avons appliqué aux extraits comportant *sur le plan (de)*, *au niveau (de)*, *dans le domaine (de)*, *dans le secteur (de)* et *en matière (de)* TOP intra- et extraprédicatifs. Ce test vise à départager, parmi les emplois de ces marqueurs comme TOP, ceux qui semblent plus topicaux que d’autres. Le test utilisé ne s’applique pas aux occurrences de *au sujet de* et *à propos de* parce que tout simplement il repose sur la possibilité de commuter tel emploi de ces marqueurs soit avec *à propos de / au sujet de* (TOP++) soit avec *concernant / en/pour ce qui concerne* (TOP +)¹¹.

Diagramme 3

Pourcentages de marqueurs codés « TOP+ » et « TOP++ » dans chacun des deux sous-corpus : LF-19 et LF-21

¹¹ Le fait de considérer que les emplois dans lesquels *sur le plan de/Adj*, *au niveau de/Adj*, *dans le domaine de/Adj*, *dans le secteur de/Adj* et *en matière de/Adj* commutent avec *à propos (de)* et *au sujet (de)* sont les plus topicaux remonte à Reinhart (1980) et (1981) qui en fait un test pour identifier les topiques de discours, et elle est reprise par Velghe & Lahousse (2015) et Velghe (2015). Sur *à propos de* voir Prévost (2007, 2011), Lagae (2011) et Porhiel (2001) qui le compare à *au sujet de*.



Là encore, *en matière (de)* se distingue des autres marqueurs puisque ses emplois les plus topicaux dominent dès le XIX^e siècle. À l'inverse, *dans le secteur (de)*, *sur le plan (de)*, *au niveau (de)* voient leur emploi TOP++ n'apparaître qu'au XXI^e siècle (*au niveau (de)* se classant en dernier). Enfin, *dans le domaine (de)* occupe une position intermédiaire : quoiqu'on le rencontre dans des emplois très topicaux dès le XIX^e siècle, ces derniers s'avèrent minoritaires aux XIX^e et XXI^e siècles.

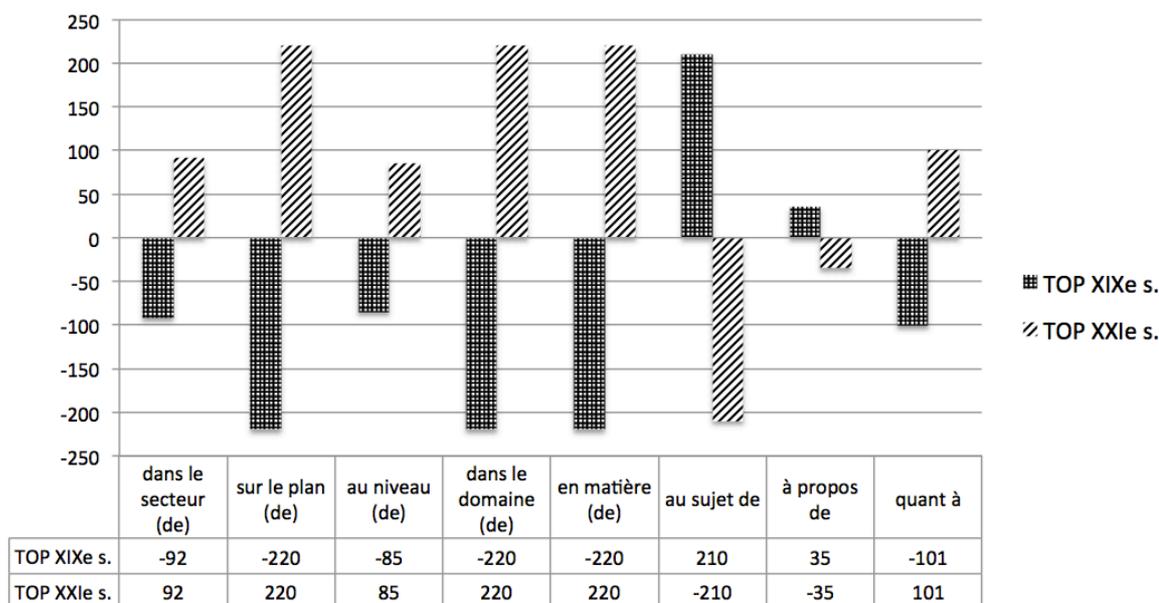
3. Mise en œuvre de calculs de spécificités

Pour compléter les analyses présentées dans la section précédente et qui s'appuient sur l'examen des fréquences relatives calculées pour les deux sous-corpus LF-19 et LF-21, nous recourons ci-dessous au calcul des spécificités¹² (cf. Lafon 1980). Dans un premier temps, nous avons fait un calcul de manière à déterminer si les marqueurs étudiés s'avéraient sur- ou sous-représentés dans chacune des deux parties (LF-19 et LF-21) de notre corpus total. Voici les indices auxquels nous parvenons :

Diagramme 4

Indices de spécificité calculés pour les marqueurs employés comme topiques quelle que soit leur position syntaxique (intraprédicatifs et extrapredicatifs). Corpus total partitionné en deux parties (LF-19 et LF-21)

¹² Pour une présentation de ce calcul et des paramètres qu'il met en jeu, voir D. Vigier dans ce numéro.



Ce qui frappe d'abord, c'est qu'aucun des marqueurs étudiés ne possède une distribution « banale » (au sens statistique) dans les deux parties du corpus¹³ : tous sont soit *sur-représentés* soit *sous-représentés*. Notre corpus nous invite donc à considérer que quel que soit le marqueur, *il se passe quelque chose de statistiquement significatif* quant à son usage comme marqueur TOP entre le XIX^e et le XXI^e siècle dans le journal *Le Figaro*.

En second lieu, il est frappant également que la répartition des sur- et des sous-emplois de *au sujet de*, *à propos de* dans les deux parties du corpus soit symétrique de celle observée pour les autres marqueurs. Tandis que les deux premiers sont sous-représentés au XXI^e siècle, les autres s'avèrent sur-représentés pour la même période. On retrouve ici la désaffection relative qui frappe l'usage de ces deux marqueurs dans notre corpus et que la comparaison des fréquences relatives révélait déjà.

Autre observation : les « profils » de sur- et de sous-emploi des trois marqueurs *sur le plan (de)*, *dans le domaine (de)* et *en matière (de)* apparaissent très proches dans le diagramme. Le calcul montre en effet que, pour ce qui regarde leur usage comme marqueurs de topique, ce sont eux qui connaissent les plus grands changements entre le XIX^e et le XXI^e siècle. Le sort de *dans le secteur (de)* et de *au niveau (de)* – aux profils proches aussi – est à cet égard moins contrasté. Et c'est *à propos de* qui présente finalement le moins de contraste en termes d'usage comme TOP entre les deux périodes considérées.

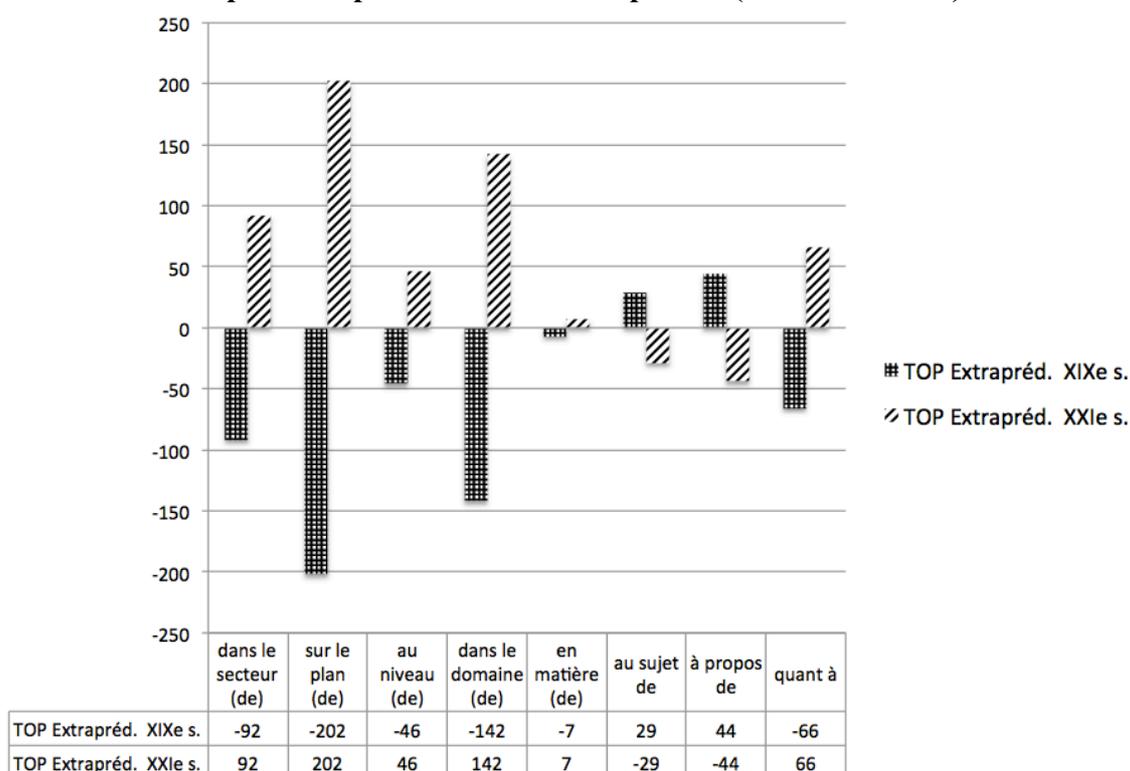
Faut-il enfin s'étonner que les indices de sur- et de sous-spécificité affectés à *quant à* soient très inférieurs à ceux calculés pour *sur le plan (de)*, *dans le domaine (de)* et *en matière (de)* notamment, alors que nous avons plusieurs fois souligné sa très haute fréquence d'emploi comme topique aux XIX^e et XXI^e siècles ? En fait, il n'y a là rien d'étonnant : les deux indices calculés montrent en effet que si *quant à* est (encore) plus souvent utilisé comme TOP dans *Le Figaro* XXI^e siècle qu'il ne l'était dans *Le Figaro* XIX^e siècle, le contraste entre sa fréquence d'emploi comme TOP entre ces deux périodes est statistiquement moins significatif qu'il ne l'est pour les trois autres marqueurs.

Alors que le diagramme 4 présentait les indices de spécificité calculés pour les emplois TOP des marqueurs étudiés dans les deux parties de notre corpus quelle que soit leur position syntaxique, le diagramme 5 restreint ce calcul à leurs seuls emplois TOP extrapredicatifs.

¹³ ce qui serait le cas si la valeur absolue de l'indice de spécificité était inférieure (au sens strict) à 3.

Diagramme 5

Indices de spécificité calculés pour les marqueurs employés comme topiques extrapredicatifs. Corpus total partitionné en deux parties (LF-19 et LF-21)



On observera d'abord la relative faiblesse des deux indices affectés aux fréquences d'emplois de *en matière (de)*. Si on les compare à celui affecté au même marqueur dans le diagramme 4, on est frappé par leur grande différence puisqu'on passe (en valeur absolue) de 220 à 7. Cette observation révèle une caractéristique importante de l'évolution diachronique des emplois de *en matière (de)* : quoique ce marqueur voie son usage comme TOP s'accroître de manière spectaculaire entre le XIX^e et le XXI^e siècle (dans notre corpus : cf. diagramme 1 *supra*), cet essor concerne avant tout ses emplois intrapredicatifs, et non extrapredicatifs. A cet égard, la comparaison avec *sur le plan (de)* et *dans le domaine (de)* ne manque pas d'intérêt : nous avons souligné *supra* (commentaire du diagramme 4) que ces trois marqueurs se différencient de tous les autres en ceci que le contraste entre leur usage comme TOP au XIX^e et au XXI^e siècles apparaît comme très marqué (valeur absolue des indices de sur- et de sous spécificité > 200). Or si ce contraste tend à s'évanouir (même s'il reste significatif) pour *en matière de* dans le diagramme 5 (du fait, répétons-le, de ses emplois de TOP essentiellement intrapredicatifs), tel n'est pas le cas des deux autres marqueurs : preuve que leur usage comme TOP extrapredicatif au XXI^e siècle diffère aussi nettement de celui qu'il était au XIX^e siècle.

On gagne enfin à étendre au marqueur *au sujet (de)* la comparaison des indices présentés dans les diagrammes 4 et 5. Leur évolution se révèle en effet à la fois très proche et aux antipodes de celle de *en matière (de)*. Proche, car l'on observe une chute tout aussi spectaculaire de la valeur absolue de l'indice des spécificités (= 29) affecté aux emplois TOP extrapredicatifs de *au sujet (de)* (diagramme 5) alors qu'elle était très élevée (= 210) pour ses emplois TOP (intra- et extra-predicatifs : cf. diagramme 4). Si l'emploi comme TOP de ce marqueur a donc significativement changé en un siècle, ce changement – comme pour *en matière de* – concerne moins ses emplois extrapredicatifs que ses emplois intrapredicatifs. Aux antipodes enfin, parce que l'évolution globale des emplois TOP de *au sujet de* est à rebours de celle de *en matière de* : alors que la première connaît un déclin général (plus ou moins creusé selon qu'on se focalise sur les seuls emplois extrapredicatifs ou bien sur tous), la seconde connaît au contraire un essor (limité aux TOP intrapredicatifs).

Conclusion

Les données quantitatives présentées dans cette étude montrent que les emplois de *sur le plan (de)*, *au niveau (de)*, *dans le domaine (de)*, *dans le secteur (de)* ont connu un essor remarquable dans *Le Figaro* entre la fin du XIX^e et le début du XXI^e siècle. Cet essor que l'on peut imputer au fait qu'elles ont commencé à se grammaticaliser comme des marqueurs de topique de discours contraste avec les données rassemblées sur *quant à* qui n'a fait, pendant le même temps, que se maintenir dans cette fonction où il est de très loin le plus usité. *Au sujet de* et *à propos de* qui avaient, comme *quant à*, plusieurs siècles d'avance sur les expressions comme *sur le plan (de)*, *au niveau (de)*, *dans le domaine (de)*, *dans le secteur (de)*, paraissent par contre en perte de vitesse par rapport à *quant à*, surtout pour les emplois extraprédicatifs. Pour ce qui concerne enfin *en matière (de)*, ses emplois comme Marqueur de Topique de Discours restent au début du XXI^e siècle à un bon niveau, mais essentiellement comme intraprédicatifs.

Par ailleurs, si *sur le plan (de)*, *au niveau (de)*, *dans le domaine (de)*, *dans le secteur (de)* ont connu un même essor de la fin du XIX^e au début du XXI^e siècle, il demeure des différences importantes entre d'un côté *sur le plan (de)* et *dans le domaine (de)* et, de l'autre, *dans le secteur (de)* et *au niveau (de)*. Ces différences demanderaient à être précisées et testées sur plus d'extraits que ceux que nous avons pu annoter, et sur d'autres expressions comme *en regard de* ou *du point de vue (de)* qui pourraient être ajoutées à celles que nous avons prises en compte dans cette étude. Il y a en effet tout lieu de penser que, dans le domaine qui nous préoccupe comme dans beaucoup d'autres, les marqueurs linguistiques assument des fonctions très spécifiques qui doivent être étudiées pour elles-mêmes, en synchronie, comme le font Delahaie & Flament-Boistrancourt (2013) à l'oral pour *niveau*, *par rapport à* et *en termes (de)*, et Flament-Boistrancourt (2002) pour *au niveau (de)*. On renvoie également à Velghe et Lahousse (2015) et Velghe (2015) qui apportent des données synchroniques sur, notamment, *au niveau (de)*, *au sujet de*, *en matière (de)* en français parlé et écrit (corpus *Le Monde*) et le « Yahoo-based Contrastive Corpus of Questions and Answers » (entre l'écrit et l'oral).

Le travail ne manque donc pas sur le sujet, sans compter un aspect que nous n'avons pas discuté ici mais qui ressort nettement de la fréquentation de notre corpus : très souvent, les emplois comme marqueurs de topique de discours étudiés dans cet article apparaissent en série, sur des fragments de texte relativement étoffés.

Références bibliographiques

- ANSCOMBRE J.-C. (2006), « Les locutions *quant à*, *pour ce qui est de*, *en ce qui concerne*. Chronique d'un discours annoncé », *Modèles Linguistiques*, 25.2, 155-169.
- CHAROLLES M. (1997), « L'encadrement du discours - Univers, Champs, Domaines et Espaces », *Cahier de Recherche Linguistique*, 6, 1-73.
- CHOI-JONIN I. (2003), « Ordre syntaxique et ordre référentiel : Emplois de la locution prépositive *quant à* », in B. Combettes, C. Schnedecker & A. Theissen (éds), *Ordre et distinction dans la langue et le discours*, Paris : Champion, 133-147.
- COMBETTES B. et PREVOST S. (2001), « Évolution des marqueurs de topicalisation », *Cahiers de Praxématique* 37, 103-124.
- COMBETTES B. & PREVOST S. (2003), « Texte argumentatif et topicalisation d'une proposition : une approche diachronique », *Scolia* 16, p. 63-75
- DEBAISIEUX J.-M. (2001), « Contraintes syntaxiques et discursives des emplois de *quant à* et *en ce qui concerne* dans les corpus oraux », *Cahiers de praxématique* 37, 125-164.
- DELAHAIE J. & FLAMENT-BOISTRANCOURT D. (2013), « Questions de variation : autour de quelques locutions méconnues de l'oral, *niveau*, *par rapport à*, *en termes de*. », in *Actes du colloque LPTS*,

- « *Across the line of speech and writing variation* » (octobre 2011, Louvain-la-Neuve), Presses Universitaires de Louvain, collection « *Corpora and Language in Use* ».
- FLAMENT-BOISTRANCOURT D. (2002), « *Et au niveau tarifs, ça donne quoi ? L'éclairage du corpus Lancom sur le mot niveau* », in J. Binon, P. Desmet, J. Elen, P. Mertens & L. Sercu. (éds), *Tableaux vivants. Opstellen over taal-en-onderwijs aangeboden aan Mark Debrock*, Leuven : Universitaire Pers, 117-132.
- FLØTTUM K. (1999), « *Quant à – thématiseur et focalisateur* », in C. Guimier (éd.), *La thématization dans les langues*. Berne : Peter Lang, 135-149.
- FRASER B. (2009a), « *Towards a theory of discourse markers* », in K. Fischer (ed.) *Approaches to Discourse Particles*, Amsterdam: Elsevier Ltd, 189-204
- FRASER B. (2009b), « *Topic Orientation Markers* », *Journal of Pragmatics* 41, 892-898
- GOUTSOS D. (1996), « *A model of sequential relations in expository text* », *Text* 16 (4), 501-533.
- GUIMIER C. (1996), *Les adverbes du français. Le cas des adverbes en -ment*, Paris : Ophrys.
- KUPPEVELT J. Van (1995), « *Discourse Structure, Topicality and Questioning* », *Journal of Linguistics* 31, 1, 109-147.
- LAFON P. (1980), « *Sur la variabilité de la fréquence des formes dans un corpus* », *Mots* 1, 127-165.
- LAGAE V. (2003), « *Quant aux livres/De livre, il n'en a lu aucun. Étude syntaxique de deux constructions détachées* », *Linguisticae Investigationes* 26.2, 237-260.
- LAGAE V. (2011), « *A propos de : un marqueur thématique très particulier* », in D. Amiot, W. De Mulder, E. Moline & D. Stosic (éds), *Ars Grammatica. Hommage à Nelly Flaux*, Berne : Peter Lang, 273-288.
- LAMIROY B. & CHAROLLES M. (2005), « *Constitution et exploitation de corpus dédiés à l'analyse linguistique des adverbiaux* », in A. Condamines ed., *Sémantique et corpus*, Paris Hermès, 109-146
- LEBART L. & SALEM A. (1994), *Statistique Textuelle*, Dunod : Paris.
- MARANDIN J.-M. (1988), « *À propos de la notion de thème de discours. Éléments d'analyse dans le récit* », *Langue française* 78, 67-87.
- PORHIEL S. (1998), *Les indicateurs d'intérêt*, Thèse, Université de Paris 13/Presses du Septentrion.
- PORHIEL S. (2001), « *Au sujet de et à propos de - une analyse lexicographique, discursive et linguistique* », *Travaux de linguistique* 2001/1, 42-43, 171-181.
- PORHIEL S. (2003), « *Les indicateurs d'intérêt dans l'organisation textuelle* », in B. Combettes, C. Schnedecker & A. Theissen (éds), *Ordre et distinction dans la langue et le discours*, Paris : Champion, 425-442.
- PORHIEL S. (2004), « *Les introducteurs de cadre thématique* », *Cahiers de lexicologie* 85, 2004-2, 9-45.
- PREVOST S. (2003), « *Quant à : analyse pragmatique de l'évolution diachronique (14^{ème} - 16^{ème} siècles)* », in B. Combettes, C. Schnedecker & A. Theissen (éds), *Ordre et distinction dans la langue et le discours*, Paris : Champion, 443-459.
- PREVOST S. (2007), « *À propos de X / à ce propos / à propos : évolution du XIV^e au XVI^e siècle* », *Langue française* 156, 108-126.
- PREVOST S. (2008), « *Quant à X et à propos de X du 14^{ème} au 16^{ème} siècle : émergence de deux marqueurs de topicalisation* », *L'information grammaticale* 118, 38-43.
- PREVOST S. (2010), « *Quant à X : du complément à l'introducteur de topique en passant par l'introducteur de cadre* », in B. Combettes, C. Guillot, E. Oppermann-Marsaux, S. Prévost & A. Rodríguez Somolinos (éds), *Le changement en français. Etudes de linguistique diachronique*, Bern : Peter Lang, 325-343.
- PRÉVOST S. (2011), « *A propos : from verbal complement to 'utterance marker' of discourse shift* », *Linguistics* 49 : 2, 391-413.
- REINHART T. (1980), « *Condition for text coherence* », *Poetics to day* 1/4, 161-180.
- REINHART T. (1981), « *Pragmatics and linguistics: An analysis of sentence topics* », *Philosophica* 27, 53-94.
- ROBERTS C. (2004), « *Discourse context in dynamic interpretation* », in L. Horn & G. Ward (eds), *Handbook of Contemporary Pragmatic Theory*, Oxford: Blackwell, 197-220.

- STEIN A. & SCHMID H. (1995), « Etiquetage morphologique de textes français avec un arbre de décisions », *Revue T.A.L* Vol. 36.
- TRAUGOTT, E. & E. KÖNIG (1991), « The semantics-pragmatics of grammaticalization revisited », in Traugott, E. & Heine B. (eds.), *Approaches to grammaticalization*, Amsterdam: Benjamins, 189-218.
- TRAUGOTT, E. & DASHER B. (2002), *Regularity in Semantic Change*, Cambridge : Cambridge University Press.
- VAN DIJK T. A. (1977), « Sentence Topic and Discourse Topic », *Papers in Slavic Philology* 1, 1977, 49-61
- VELGHE T. & LAHOUSSE K. (2015), « Thematic markers in informal written French : *pour ce qui est de, au niveau (de), and en matière de* », *Journal of French Language Studies* vol. 25 (3), 423-444.
- VELGHE T. (2015), *Les marqueurs de thématisation en français : Syntaxe, Discours, Prosodie*, Thèse de Doctorat, KU Leuven.