



HAL
open science

Exploration textométrique d'un corpus de motifs juridiques dans le droit international des transports

Fadila Taleb, Maryvonne Holzem

► **To cite this version:**

Fadila Taleb, Maryvonne Holzem. Exploration textométrique d'un corpus de motifs juridiques dans le droit international des transports. JADT 2018 : 14es Journées internationales d'Analyse statistique des Données Textuelles, Jun 2018, Rome, Italie. hal-02378331

HAL Id: hal-02378331

<https://normandie-univ.hal.science/hal-02378331>

Submitted on 25 Nov 2019

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.

Exploration textométrique d'un corpus de motifs juridiques dans le droit international des transports

Fadila Taleb¹, Maryvonne Holzem²

¹Université Rouen Normandie – fadila.taleb1@univ-rouen.fr

²Université Rouen Normandie – maryvonne.holzem@univ-rouen.fr

Abstract

Within the framework of a research whose objective consists of responding to a need formulated by the IDIT, which helps to interpret the jurisprudential texts contained in its database, we are looking to highlight the interpretive paths considered as modal scenarios. We propose here a preliminary textometric analysis in order to define the linguistic profile of the corpus and to detect certain repeated segments that may represent a relevant constraint to complete and enrich the interpretive paths identified in the case law.

Résumé

Dans le cadre d'une recherche dont l'objectif consiste à répondre à un besoin formulé par l'IDIT¹, celui d'aider à l'interprétation des textes jurisprudentiels contenus dans sa base de données, nous cherchons à mettre au jour des parcours interprétatifs envisagés comme des scénarios modaux. Nous proposons ici une analyse textométrique préalable afin de cerner le profil linguistique du corpus et de détecter certains *segments répétés* pouvant représenter une contrainte pertinente pour compléter et enrichir les parcours interprétatifs identifiés dans les textes jurisprudentiels.

Keywords: textométrie, parcours interprétatif, scénario modal, segments répétés, motifs juridiques, droit des transports.

1. Introduction

1.1. Contexte

Dans le cadre d'un projet pluridisciplinaire « PlaIR »², des chercheurs informaticiens, linguistes, juristes posent la question de l'aide à l'interprétation³ du fond jurisprudentiel de la base de données de l'IDIT. Du point de vue linguistique, notre tâche préalable à une implémentation consiste en l'étude de décisions de justice dans le but de comprendre leur structure, le mécanisme argumentatif mis en œuvre et les mouvements de transformations textuelles susceptibles de déclencher des parcours interprétatifs pouvant aider à la lecture de ces décisions. Notre recherche s'écarte des modèles prédictifs, *justice prédictive ou legaltech*, qui, sous l'influence des *big data et du Machine Learning*, produisent des résultats de contentieux sur des bases algorithmiques. De ce point de vue, nous partageons les craintes de

¹ Institut du Droit International des Transports.

² *Plateforme d'Indexation Régionale*

³ Notre objectif est celui d'une aide instrumentée centrée sur l'agir de l'utilisateur *cf.* travaux du groupe *v* (Holzem et Labiche, 2017).

bon nombre de juristes de voir ces *legaltech* « devenir eux mêmes une nouvelle forme de justice » (Garapon, 2017). Il s'agit d'une pratique textuelle (et intertextuelle) comprise comme régime de transformation et d'interprétation. Dans cette perspective, notre recherche se place donc du côté de la *jurilinguistique* et son objectif est d'essayer de comprendre dans une approche linguistique et à travers l'étude du matériel textuel les décisions de justice et les stratégies argumentatives mises en œuvre pour ainsi aider à leur interprétation.

1.2. Questionnement et hypothèse

Pour aider à l'interprétation nous cherchons à cerner les stratégies argumentatives mises en œuvre par le juge, notamment dans sa manière d'intégrer et de prendre en charge les discours des autres (celui des parties du procès, celui des experts, celui du législateur etc.). Notre hypothèse est fondée sur des recherches antérieures (Holzem 2014 et Taleb 2014⁴) qui ont montré l'intérêt de la prise en compte des modalités linguistiques suivant le modèle développé dans (Gosselin 2010) pour la constitution d'un parcours interprétatif (Rastier 2001) envisagé ici comme scénario modal susceptible d'aider à l'interprétation. Mais avant de procéder à une telle analyse textuelle menée directement sur des textes pleins, nous avons eu besoin de cerner dans sa globalité et ses spécificités le profil linguistique de notre corpus d'étude. Pour cela, nous avons eu recours à une analyse textométrique approfondie, menée avec le logiciel TXM. Au fil de nos investigations textométriques, nous nous sommes rendu compte de l'importance de certaines fonctions offertes par ces outils pour la détection, par exemple, de *segments répétés*⁵, qui peuvent représenter une contrainte pertinente pour compléter les parcours interprétatifs identifiés grâce à l'étude des modalités. L'objectif de cet article est de présenter, dans ses grandes lignes, en raison de la place, l'analyse textométrique menée sur notre corpus.

2. Corpus et méthodologie

2.1. Description globale

Nous avons, à la suite de Rastier (2011), retenu le critère du genre comme critère définitoire du corpus de référence. Il regroupe des textes (décisions de justice) relevant du discours judiciaire⁶ et appartenant au genre jurisprudentiel⁷.

En reprenant la typologie du corpus proposée par B. Pincemin (1999) et reprise par (Rastier 2011), nous avons distingué quatre niveaux de corpus : (i) un corpus *existant/latent* (*archives* pour Rastier) qui correspond dans notre recherche à la base de données de l'IDIT ; (ii) un corpus de *référence* qui renvoie à l'ensemble des documents numérisés dans le fond jurisprudentiel de l'IDIT ; (iii) un corpus *d'étude* qui contient un nombre délimité de ces décisions sélectionnées pour les besoins de notre recherche et enfin (iv) un corpus *distingué* (*corpus d'élection ou sous-corpus* pour Rastier) correspondant à des passages précis des textes étudiés nommés « *les motifs* ». Ces derniers constituent le cœur du jugement, le juge exposant « (...) *les raisons de faits et de droit qui justifient la décision* (...)» (Cohen et

⁴ Un mémoire de master 2 recherche en science du langage soutenu en juin 2014 : « *Étude du scénario modal et du syllogisme juridique pour la compréhension du processus de production du texte. Cas des textes du droit.* »

⁵ *Suite de formes graphiques identiques attestées plusieurs fois dans le texte.*

⁶ Il renvoie aux discours produits par (ou au sein) des juridictions. Il est à distinguer du discours *juridique* qui désigne, entre autre, les domaines du droit ou ses sources (lois, réglementation etc.). L'un concerne la création du droit, l'autre rend compte de son aspect applicatif.

⁷ Le terme de jurisprudence renvoie ici à l'« *ensemble des décisions rendues par les tribunaux d'un pays, pendant une certaine période dans une certaine manière.* » (Dictionnaire du vocabulaire juridique 2017, éd. LexisNexis) P.322)

Pasquino, 2013). Notre intérêt pour cette zone textuelle est doublement motivé. Premièrement notre objectif consiste à repérer les moments clés de transformations du jugement pour cerner les stratégies argumentatives mises en œuvre et partant aider à leur interprétation. Deuxièmement la motivation est une composante commune⁸ à toutes les décisions de toutes les juridictions. Elle doit faire face à une double exigence : logique et persuasive. L'une est due à la forme syllogistique du raisonnement juridique imposée et l'autre à la nécessité de persuader l'auditoire de la décision⁹ de sorte à éviter les recours et faire accepter la solution juridique apportée comme étant la seule possible.

2.2. Caractéristiques quantitatives

Le volume textuel du corpus d'étude est de 878848 occurrences dont 22456 formes. Le sous-corpus des motifs représente à lui seul près de la moitié des occurrences du corpus d'étude. Il contient 393092 occurrences pour 14742 formes. La dysmétrie de la distribution des formes dans les différentes zones délimitées montre l'importance et le rôle des motifs dans les décisions de justice, ils sont leur raison d'être, et tout juge est dans l'obligation de motiver son jugement.

2.3. Encodage et prétraitement

Notre corpus présente l'avantage d'être accessible en ligne. Cependant, l'ensemble des textes au format PDF n'est pas homogène : certains documents proviennent d'un format image (non *océrisé*¹⁰). Le format PDF n'étant pas pris en charge par TXM, nous avons tout d'abord procédé à une conversion (avec la technique d'océrisation pour les fichiers annotés et numérisés) au format TXT, puis dans un second temps à un codage XML en s'inspirant des recommandations de la TEI¹¹ pour l'encodage des données textuelles. Ce dernier nous permet une navigation plus fine dans le corpus grâce à des métadonnées péritextuelles, comme celles relatives au type de la juridiction : *tribunal de commerce (TC)*, *cour d'appel (CA)*, *cour de cassation (CC)*, à la date et au lieu, et des métadonnées intratextuelles, telles que celles relatives à des parties spécifiques dans les textes. Nous avons relevé quatre parties principales : *faits*, *moyens*, *motifs*, *conclusions*. *Les motifs et les conclusions* sont présents dans toutes décisions étudiées. *Les faits* sont absents des CC, et *les moyens* ne sont pas toujours indiqués comme tels dans les arrêts CA, ils sont souvent rappelés dans la zone des faits sous forme de discours indirect.

La figure suivante représente les différentes phases de préparation du corpus avant son traitement textométrique :

⁸ Ce qui n'est pas le cas pour les autres composantes. Ainsi, *l'exposé du litige* ne figure pas dans les arrêts de la cour de cassation, car celle-ci étant une juridiction d'ordre suprême, son rôle est de veiller à la bonne application des normes juridiques, elle considère l'appréciation des faits par les juges de fond comme étant souveraine.

⁹ Composée certes des parties du litige directement concernées par la décision, mais aussi les autres juges des autres juridictions et un public encore plus large, le destinataire universel.

¹⁰ OCR (*Optical Character Recognition*) Reconnaissance Optique de Caractères, étape nécessaire pour déchiffrer les formes et les traduire ici en lettres.

¹¹ *Text Encoding Initiative* : recommandations standard pour l'encodage des documents numériques.

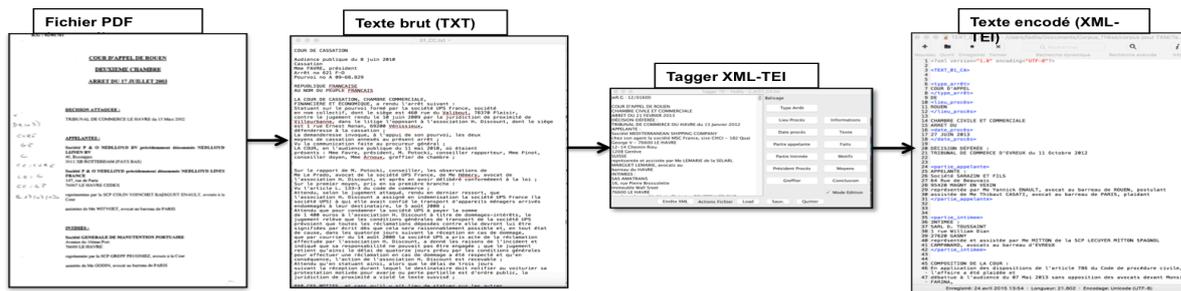


Figure 1 Les étapes de préparation du corpus

Pour le passage du format TXT au format XML-TEI nous avons créé les balises spécifiques au genre du corpus étudié : <TypeJurisdiction>, <DateProcès>, <LieuProcès> etc. Nous avons eu recours à un encodage semi-automatique au moyen d'un *tagger* conçu spécialement pour notre étude par Eric Trupin, MCF en informatique au laboratoire LITIS¹². Cette étape indispensable de préparation du corpus pour le traitement textométrique a été à la fois chronophage et délicate : traitement des annotations manuscrites et nettoyage de documents plus anciens.

3. Exploration textométrique du corpus distingué : la zone des motifs

3.1. Etude occurrentielle : les spécificités lexicales

Une première étude contrastive au moyen d'un traitement textométrique phare, *le calcul de spécificités*¹³, permet d'avoir une vue globale sur les caractéristiques lexicales du corpus distingué « *les motifs* ». Le tableau ci-dessus dresse la liste des 20 premières formes les plus spécifiques à cette zone. Il est trié par ordre décroissant sur l'indice de spécificité de celle-ci :

Unités	Fréquence T=828700	MOTIFS t=391346	score √	CorpusR t=437354	score
Attendu	801	797	250,8	4	-250,8
que	16573	9373	129,9	7200	-129,9
Considérant	236	226	59,3	10	-59,3
Mais	163	163	53,1	0	-53,1
ne	7197	4031	50,3	3166	-50,3
est	4843	2778	45,1	2065	-45,1
attendu	181	170	41,3	11	-41,3
pas	4136	2372	38,5	1764	-38,5
moyen	396	308	34,9	88	-34,9
donc	802	530	26,4	272	-26,4
sera	516	356	22,9	160	-22,9
résulte	378	267	19,6	111	-19,6
Que	532	353	18,3	179	-18,3
succombe	59	58	17,4	1	-17,4
équité	51	51	16,6	0	-16,6
convient	240	177	16,2	63	-16,2
marchandises	1549	892	15,7	657	-15,7
suivant	156	123	15,3	33	-15,3
condamnera	55	53	14,6	2	-14,6
inéquitable	62	58	14,2	4	-14,2

Figure 2 : spécificité lexicales de la zone des motifs

¹² Laboratoire d'Informatique, du Traitement de l'Information et des Systèmes, Université Rouen Normandie

¹³ Le calcul de spécificités implémenté dans TXM repose sur la *loi hypergéométrique* développée par Lafon (1984). Le seuil de pertinence d'une distribution est fixé à 2 : +2 l'indice de spécificité est positivement significatif, -2 il est négativement significatif. L'indice se situant entre les deux est banal.

Nous portons ici attention à un usage excessif d'occurrences caractéristiques du discours judiciaire et constitutif de la zone des motifs : *Attendu, que, Considérant*¹⁴, *attendu*, de même pour les connecteurs : *Mais* et *donc*. L'ensemble de ces marqueurs jouent un rôle spécifique ici, celui de ponctuer l'argumentation du juge en assurant sa progression syllogistique. L'usage excessif du futur, représenté avec les verbes être (*sera* : 22,9) et condamner (*condamnera* : 14,6) n'est pas surprenant, car avant de prononcer le verdict final dans un acte exclusivement directif (énoncé réservé à la zone des dispositifs), les juges avancent au préalable dans la zone des motifs les résultats (comme le montre d'ailleurs le suremploi du verbe *résulte* (19,6)) de leurs argumentations : « *Le jugement entrepris sera confirmé en ses autres dispositions qui ne sont pas critiquées* ».¹⁵; « *Le tribunal condamnera Monsieur le capitaine du [...] ;* ».¹⁶ L'emploi significatif d'autres mots, comme *équité, marchandises, inéquitable* renvoie à la thématique des textes étudiés : *le droit des transports*. L'emploi significatif des adverbes de négation : *ne* (+50,3), *pas* (+38,5) révèle une caractéristique particulière de l'argumentation juridique car, fidèle au principe spinoziste *Determinatio negatio es*, la négation manifeste une valeur réplivative et résultative (i.e. portée référentielle en réponse à ce qui a été énoncé précédemment et qui n'a plus lieu d'être) préparatoire à la transformation juridique de l'énoncé.

3.2. Etude contextuelle

Au-delà des investigations menées sur des unités lexicales minimales, les outils que propose la communauté ADT problématiquement la notion de contexte selon des paliers différents pour privilégier un retour au texte. Nous allons ici donner l'exemple de la contextualisation des « *attendu* » dans la zone des motifs dont le suremploi a été relevé dans le tableau ci-dessus. Suite à une étude cooccurrence autour du mot-pôle « *attendu* », nous avons repéré une très forte attractivité avec le connecteur « *Mais* » (l'indice de spécificité¹⁷ = +95).

text_id, div, type	Contexte gauche	Pivot	Ca Contexte droit
0001 TC, motifs TC	à 25, 912, 64 € ;	Mais attendu	, comme le fait remarquer l'expert, que CFT n'a
0001 TC, motifs TC	très partiellement, responsable des événements litigieux ;	Mais attendu	que la lettre de garantie adressée à CPA dès le 2 mars
0004 TC, motifs TC	fait est constitutif d'une faute lourde ;	Mais attendu	que le même procès-verbal précise que la chaudière était sèche, la
0010 TC, motifs TC	devant les juridictions de cet Etat » ;	Mais attendu	que l'article 6, 2 de ladite Convention ordonne que le
0011 TC, motifs TC	les causes de l'incident demeurent inconnues ;	Mais attendu	que la société TERMINAUX DE NORMANDIE a reconnu sa faute ou s'est
0012 TC, motifs TC	à « l'économie du contrat » ;	Mais attendu	ou'une clause attributive de compétence de juridiction n'accortient pas à
0012 TC, motifs TC	dont le transporteur ne pouvait avoir connaissance ;	Mais attendu	ou'aux termes du rapport d'expertise versé aux débats, il
0013 TC, motifs TC	bénéficiaire de la présomption de livraison conforme ;	Mais attendu	que selon les mentions portées au connaissance, le transporteur maritime a
0013 TC, motifs TC	et l'inspection sanitaire du lendemain matin ;	Mais attendu	ou'il s'agit de relevés manuels, établis par ailleurs unilatéralement
0015 TC, motifs TC	l'intervention du tiers sur le conteneur ;	Mais attendu	que le connaissance fourni aux débats par les parties mentionne clairement
0021 TC, motifs TC	que MSC est déclarée responsable du sinistre ;	Mais attendu	ou'en application de l'article 27 § 1 de la loi du
0021 TC, motifs TC	498, 85 € décait usine ;	Mais attendu	ou'en application de la loi française du 18 juin 1966,
0030 TC, motifs TC	décait des faits était de cinq ans ;	Mais attendu	ou'aucun contrat n'est produit, que les assureurs ne précisent
0030 TC, motifs TC	le connaissance pour le compte de DELMAS ;	Mais attendu	que le rédacteur d'un document, surtout lorsqu'il s'agit
0030 TC, motifs TC	ne disposant pas d'entrepôt frigorifique ;	Mais attendu	que « emporté sous la responsabilité du transporteur » ne veut pas
0030 TC, motifs TC	Convention de Bruxelles du 25 août 1924 ;	Mais attendu	que DELMAS a délivré un connaissance sans réserves, alors qu'elle
0030 TC, motifs TC	DELMAS, in solidum avec celle-ci ;	Mais attendu	ou'elles ne produisent aucune pièce permettant de statuer sur l'étendue
0030 TC, motifs TC	mais que leur action est prescrite ;	Mais attendu	que le conteneur était sous la responsabilité de DELMAS lors du séjour
0030 TC, motifs TC	autres demandent une indemnité pour résistance abusive ;	Mais attendu	que le litige était réel ; que ni DELMAS, ni ses
0031 TC, motifs TC	50 -1098 COPENHAGEN (Danemark) » ;	Mais attendu	ou'agent a le pouvoir légal de représenter le transporteur en
0031 TC, motifs TC	Il y a prescription de livraison conforme ;	Mais attendu	que le connaissance est intitulé « Combined Transport Bill of Lading »
0031 TC, motifs TC	un montant de 28 500 US \$;	Mais attendu	ou'aucun document n'apporte l'affirmation de MAERSK, alors que
0036 TC, motifs TC	est elle-même adossée au contrat de transport ;	Mais attendu	ou'au regard de la CMR et de la convention de Genève
0036 TC, motifs TC	la décision définitive concernant cette affaire ;	Mais attendu	que selon la jurisprudence constante, et notamment arrêt de la
0037 TC, motifs TC	total 13, 936, 44 € ;	Mais attendu	que les constatations de l'expert du cabinet SETEX en date du
0040 TC, motifs TC	v avoir lieu à dommages et intérêts ;	Mais attendu	que la proposition commerciale de la société CENTRAL ETHERNET RELATION
0053 CA, motifs CA	3222 - 6 du code des transports ;	Mais attendu	que les circonstances dans lesquelles s'est produit l'accident caractérisent une
0053 CA, motifs CA	elle, la somme de 9955 € ;	Mais attendu	ou'elle invoque, en réalité le préjudice personnellement subi par ses
0063 CA, motifs CA	les mêmes capacités que la pelle endommagée ;	Mais attendu	que l'expert judiciaire note que les stères de bois ont conservé
0063 CA, motifs CA	être imputé à une indisponibilité de matériel ;	Mais attendu	que l'accident s'est produit le 27 mars 2007 ; que
0066 CA, motifs CA	consolidations a conclu le 12 juillet 2012 ;	Mais attendu	ou'ce notification est antérieure à l'ordonnance de clôture prise le
0068 CA, motifs CA	constat contradictoire portant sur les marchandises concernées ;	Mais attendu	que le rapport d'emportade établit que l'ensemble de la marchandise
0068 CA, motifs CA	allocation d'une indemnité pour préjudice moral ;	Mais attendu	ou'elle invoque, en réalité le préjudice personnellement subi par ses
0069 CA, motifs CA	à l'égard de la société Sramco ;	Mais attendu	que dès le mois de novembre 2010, notamment par mail du
0068 CA, motifs CA	autre état membre de l'Union Européenne ;	Mais attendu	ou'à la suite de l'enquête douanière menée par les agents
0069 CA, motifs CA	société MERLET, n'est pas rapportée ;	Mais attendu	ou'en sa qualité de professionnel, le commissionnaire en douane est
0069 CA, motifs CA	exigibles auraient diminué de 13 042 € ;	Mais attendu	sur le premier point, qu'il résulte du courriel du 21
0069 CA, motifs CA	des clients concernés par les messages susvisés ;	Mais attendu	que par les courriels susvisés, M Merlet responsable de la société
0072 CA, motifs CA	après règlement de la facture de garde-meubles ;	Mais attendu	que le procès-verbal du 11 février 2010 a été établi unilatéralement sans

Figure 3 : Concordancier "Mais attendu" dans la zone des motifs

¹⁴ Dans notre corpus la forme *Considérant* n'apparaît que dans les CA. Son absence dans les CC serait donc significative.

¹⁵ CA Rouen, 03/10/2013

¹⁶ TC de Rouen, 15/12/2003

¹⁷ Le calcul des cooccurrences qui repère les affinités et répulsions lexicales selon un indicateur de probabilité de rencontre repose sur le même modèle que celui du calcul des spécificités (Lafon, 1984).

Nous avons remarqué une systématique dans l'usage des « *Mais attendu* » qui vient clore un enchaînement de propositions subordonnées introduites par des « *Attendu que* », repris parfois par la conjonction *que*. L'étude approfondie des contextes de ce « *Mais attendu* » révèle une incidence particulière de celui-ci sur ses contextes droits :

« Attendu que les marchandises ont été totalement perdues du fait de leur décongélation. Attendu que la première évaluation des marchandises a été établie à 18. 498, 85 € départ usine, Mais attendu qu'en application de la loi française du 18 juin 1966, le montant de la marchandise s'évalue en valeur CIF (coût + assurance + fret). Attendu qu'en l'espèce la valeur CIF des marchandises se monte à 21. 163, 96 €, c'est bien ce montant que le tribunal retiendra en préjudice principal. ».

Dans l'extrait ci-dessus *Mais attendu* introduit non seulement un mécanisme de renforcement argumentatif¹⁸, mais il joue également le rôle de déclencheur de transformation modale entre deux modalités de type axiologiques¹⁹. Dans l'exemple cité ici, le *Mais attendu* accompagné d'une référence juridique « *application de la loi française [...]* assure cette transformation entre une norme liée au domaine du transport (*marchandises totalement perdues du fait de leur décongélation* : modalité axiologique négative) et les modes d'édition d'une norme juridique cette fois. La marchandise dépréciée se trouve alors revalorisée (axiologique positif du point de vue juridique) par le changement des co-occurents à droite (*valeur CIF (coût + assurance + fret)*).

4. Conclusion

À travers cette contribution nous avons voulu montrer l'intérêt que représente une étude textométrique pour l'appréhension de son corpus d'étude. Si notre objectif principal, celui de mettre au jour des parcours interprétatifs nommés scénarios modaux (Taleb 2015), est difficilement envisageable en se limitant à une stricte étude textométrique (car elle repose sur l'étude modale propre à chaque texte). L'approche textométrique s'est avérée néanmoins pertinente pour décrire et cerner le profil linguistique du corpus. Son principe différentiel essentiel du point de vue sémantique, nous a incitées à adopter cette démarche d'analyse contrastive indispensable. L'analyse contextuelle à plusieurs paliers nous a permis le repérage de *constructions lexicales répétitives*, comme l'exemple des « *Mais attendu* » exposé ici, qui se révèlent être des moments clés du jugement et donc parcours interprétatifs corrélatifs à une transformation modale.

Références

- Cohen M. et Pasquino P. (2013). *La motivation des décisions de justice, entre épistémologie sociale et théorie du droit. Le cas des Cours souveraines et des cours constitutionnelles*. CNRS, New York University, University of Connecticut.
- Ducrot A. (1982). *Le dire et le dit*. Les Éditions de minuit, Paris.
- Garapon A. (2017). Les enjeux de la justice prédictive. *La semaine juridique LexisNexis*, N°12: 47-52.
- Gosselin L. (2010). *Les modalités en français. La validation des représentations*. Amsterdam-New-York : Rodopi B.V.

¹⁸ Voir les travaux pionniers de A. Ducrot (1984) sur les valeurs argumentatives de *Mais*.

¹⁹ *Les modalités axiologiques sont propres aux jugements de valeur de nature morale, idéologique et/ou légale*. (Gosselin, 2010).

- Holzem M. (2014). Le Parcours interprétatif sous l'angle d'une transformation d'états modaux, dans Nunes Correia C. et Coutinho M. A. (eds), *Estudos Linguisticos : Linguistic studies*, n° 10, p. 283-295.
- Holzem M. Labiche J (2017) *Dessillement numérique : énonciation, interprétation, connaissances*. Bruxelles, Bern, Berlin : PIE Peter Lang.
- Lafon P. (1984). *Dépouillements et Statistiques en Lexicométrie*. Slatkine-Champion.
- Pincemin B. (1999). *Diffusion ciblée automatique d'informations : conception et mise en œuvre d'une linguistique textuelle pour la caractérisation des destinataires et des documents*, Thèse de Doctorat en Linguistique, Université Paris IV Sorbonne, chapitre VII.
- Rastier F. (2001). *Art et science du texte*. Puf. Rastier 2011
- Rastier F. (2011). *La mesure et le grain*. Paris, Éditions Honoré Champion.
- Taleb F. (2015). Les modalités linguistiques pour aider à l'interprétation de textes juridiques. *Actes Interface TAL IHM (ITI'2015), 22ème Congrès TALn*, Caen.