

Les modalités linguistiques pour aider à l'interprétation de textes juridiques

Fadila Taleb^{1,2}

(1) Laboratoire Dysola EA 4701 (Dynamiques sociales et langagières), Normandie Université, Université de Rouen

(2) Litis EA 4108 (laboratoire d'informatique, de Traitement de l'Information et des Systèmes), Normandie Université, Université de Rouen et INSA de Rouen
fadila.taleb1@univ-rouen.fr

Résumé. L'un des objectifs majeur que se fixent les concepteurs des bases de données est celui de permettre un accès facile et rapide à de multiples documents et informations soit dans un domaine ciblé ou dans des domaines variés. Ces ressources numériques doivent aussi prendre en compte et répondre efficacement aux besoins pratiques de leurs futurs utilisateurs. Cependant, l'accès aux documents d'un corpus numérique suscite de nombreuses interrogations, notamment sur la recherche d'informations. Cette dernière se voit souvent réduite à un réseau restreint de mots-clés, ce qui ne saurait satisfaire à tous les besoins des utilisateurs. Nous proposons dans cet article une approche linguistique dans la recherche d'informations en intégrant un module d'aide à l'interprétation qui, permettra à l'utilisateur un accès au plein texte, en lui proposant des scénarii textuels vu sous l'angle des scénarii modaux susceptibles de le guider vers des passages textuels répondant à ses requêtes en fonction de sa pratique en cours.

Abstract.

Study of linguistics Modality to assistant interpretation of legal texts.

One of the main objectives set by the designers of databases is to allow an easy and fast access to multiple documents and information either in a targeted area or in different areas. These digital resources must also consider and effectively address the practical needs of their future users. However, access to documents of a digital corpus raises many questions, including research information. The latter sees often reduced to a restricted network of keywords, which does not respond in any case the requirements of users. We propose here a linguistic approach in the search for information by integrating a help module for interpretation that will allow the user access to the full text by guiding and proposing textual scenarios that meet their requests and taking into account its social profile.

Mots-clés : Scénario modal, sémantique interprétative, traitement de texte, base de données..

Keywords: Modal script, Interpretive Semantic, word processor, database.

1 Introduction

A l'heure où les technologies numériques permettent un accès aux textes et aux corpus, nous questionnerons les interactions entre linguistique et informatique que résume imparfaitement l'appellation : « dialogue Homme-machine » puisqu'il serait plus juste de parler de « modèle de dialogue informatique », laissant perceptible la programmation sous-tendue.

Il devient en effet possible à la fois de ne plus encapsuler l'interprétation textuelle dans des réseaux sémantiques préexistants (réseaux de mots-clés), mais également de prendre pleinement en compte les contextes d'interprétations liés aux vécus des utilisateurs, à leur connaissance et à leur pratique sociale guidant leur parcours interprétatif. Dans ce contexte, nous cherchons à contribuer à une aide à l'interprétation des textes juridiques numérisés dans la base de données de l'Institut du Droit International des Transports (IDIT) basé à Rouen. Cette aide à l'interprétation est appréhendée par des approches linguistiques s'inscrivant dans le domaine de la sémantique temporelle et modale et de la sémantique textuelle. A partir d'un mémoire d'étude de Master 2 (2014), dans lequel nous avons étudié des fiches numériques résumant les arrêts de la cour d'appel dans le domaine du droit du transport, nous avons mis au jour un scénario qui de par ses contraintes modales est susceptible de guider l'interprétation des utilisateurs de la base de l'IDIT. Ce scénario modal inspiré de la théorie modulaire des modalités de (Gosselin, 2010) et de la notion de parcours interprétatif de (Rastier, 2001) : « (...) se manifeste aussi, bien comme arrière-plan de la signification de certains lexèmes, comme comportement stéréotypique des individus dans une société, et comme structure sous-jacente à certains textes (...) mais aussi [comme] trame textuelle et

argumentative des arrêts de cours d'appel» (Gosselin, 0000).

Ce travail se poursuit depuis la rentrée 2014 par une recherche doctorale¹ intitulée : « *Aide à l'interprétation : apport des maqueurs modaux à l'appréhension d'un genre textuel : application au discours juridique* » Nous nous focaliserons ici sur les contraintes modales et temporelles représentant des scénarii réguliers de la pratique juridique à l'œuvre dans ces textes. L'objectif de cet article consistera à montrer l'importance d'une analyse linguistique préalable à la conception d'une plate-forme d'aide à l'interprétation dédiée à cette pratique textuelle. Le but est d'adapter un système d'aide à l'interprétation aux besoins pratiques des utilisateurs lambda sur cette base de données. Nous présenterons tout d'abord le cadre général de notre étude et la problématique centrale de notre recherche, puis la notion du scénario modal telle qu'elle a été introduite par (Gosselin, 0000). Nous illustrerons alors notre démarche à partir d'un exemple. Nous pourrions alors conclure à une appréhension de la modalité comme faisceau de contraintes pouvant peser sur le parcours interprétatif d'un utilisateur et donc sur ses traces de navigation.

2 Contexte de la recherche

Notre étude s'inscrit dans le cadre général du projet PlaIr 2.0 (Plate-forme d'Indexation Régionale) qui se développe en partenariat avec l'IDIT², l'Université et l'INSA de Rouen. Cela fait plusieurs années que des chercheurs informaticiens et linguistes (Ferrari *et al.*, 2012) des deux laboratoires du LITIS et du DYSOLA tentent d'améliorer les interactions entre les abonnés de cette base jurisprudentielle et le système d'information en vue de répondre aux besoins réels des différentes catégories professionnelles (juristes, assureurs, transporteurs, étudiants en droit ...).

Les recherches au sein du projet PlaIr 2.0 ont été jusqu'à aujourd'hui essentiellement menées du point de vue ingénierie des systèmes, mais sans qu'aucun personnel (stagiaire de M2 ou doctorant) ne soit spécifiquement chargé d'une recherche linguistique (à l'inverse des stagiaires, ingénieurs, doctorants qui ont eu pour mission les développements informatiques de cette plate-forme). Notre travail est de ce point de vue sans précédent.

2.1 Le corpus de travail

Notre corpus de travail nous a été fourni par l'institut du droit international des transports qui donne accès à une base de données contenant environ 40.000 références organisées en 4 groupes :

- jurisprudences
- articles
- réglementation
- fonds documentaire

Notre étude porte sur le fond numérisé des arrêts de jurisprudence (tribunal de commerce, cour d'appel, cour de cassation). Ces arrêts sont accompagnés de fiches présentées comme des réécritures plus abrégées de ces arrêts. Elles sont classées par numéro et répondent toutes à la même structuration. Dans la perspective de mener une étude approfondie sur les textes pleins, nous avons choisis d'explorer tout d'abord les fiches, supposées être représentatives du contenu des arrêts. Cette pré-analyse sur les fiches nous a permis d'expérimenter la notion de scénario modal.

3 La notion de scénario modal

Notre hypothèse est que l'enchaînement des actions répétées effectuées par un sujet dans le cadre d'une pratique constituerait une sorte de mémoire actancielle à l'image d'un apprentissage. Afin de rendre compte de l'organisation schématique de cette expérience mémorielle, comme suite organisée de tâches sur le modèle d'un script, une notion s'avère être centrale : le scénario.

Le concept de scénario (script, schéma, ou cadre de connaissance) est apparu dans des travaux de l'intelligence artificielle (Schank & Abelson, 1977) dans le but de modéliser les comportements humains. En informatique, il peut être défini

1. Thèse régionale financée dans le cadre du GRR CSN axe 6 interdisciplinarité et technologies numériques en SHS : apport linguistique au projet PlaIr 2.

2. La base de données de l'IDIT se propose de répondre à toutes les questions d'ordre juridique, économique ou technique, intéressant les transports de quelque nature qu'ils soient, tant sur le plan national qu'international.

comme « *une succession ordonnée d'un ensemble d'opérations (. . .) [ou] un processus qui se déroule suivant un plan défini* » (Saidali, 2002). En ingénierie informatique, cela suppose d'emmagasiner les connaissances des experts liées à la façon d'agencer et d'utiliser des outils pour atteindre un but, notamment dans la construction de plate-forme de traitement d'images comme ce fut le cas pour la thèse de (Saidali, 2002). La notion de scénario est alors nécessairement à visée pragmati

- la définition et la catégorisation du contexte
- la décomposition de l'intention en sous-intentions (à l'instar d'un programme informatique)
- l'élabque. Un scénario contiendra dans ce cas : oration d'une méthode décrite par étapes
- l'enchaînement des étapes
- le choix et le réglage des paramètres pour chaque méthode
- la perception et l'évaluation des résultats (ibid)

En linguistique, l'idée de scénario comme « représentation d'une situation » trouve son origine chez Ch. Fillmore dans sa sémantique des frames (ou sémantique des cadres) (Fillmore, 1982) qu'il introduit pour décrire des enchaînements stéréotypiques d'actions, de situations et d'événements de la vie quotidienne. Pour (Fillmore, 1982), il s'agit d'abord de décrire les propriétés des cadres auxquels renvoient les verbes avant de déterminer leurs contenus lexicaux. Dans cette perspective, ce sont les événements et les situations dans lesquels s'inscrivent les mots qui déterminent leurs significations. Tel est l'exemple d'une scène de transaction commerciale qui évoque un ensemble de lexèmes : vendre, vente, achat, acheter, payer, facture, vendeur, acheteur etc. Chacun d'eux renvoie à un aspect différent du cadre mais s'y rapporte nécessairement. Vendre par exemple met en relief les éléments : vente/vendeur, facture met en avant les lexèmes achat/acheter etc. Cette mise en relief renvoie à ce que R. Langacker ((1987) cité par (Gosselin, 0000) désigne par « profilage » et qui représente un arrière-plan conceptuel où le lexème profile des éléments. Pour comprendre par exemple la signification du concept de l'hypoténuse, il est impératif de connaître le concept de triangle rectangle et qui exige à son tour d'avoir la connaissance de triangle et d'angle droit.

En s'inspirant des travaux de Ch. Fillmore sur les Frames, mais également de la théorie des modèles comportementaux de R. Jackendoff (2007) et de la sémiotique narrative de A.J. Greimas (1976, 1983) cité dans (Gosselin, 0000), L. Gosselin propose d'introduire la notion de scénario dans le cadre de sa Théorie Modulaire des Modalités (2010) sous le nom de « scénario modal ». Ce dernier « (. . .) se manifeste aussi, bien comme arrière-plan de la signification de certains lexèmes, comme comportement stéréotypique des individus dans une société, comme structure sous-jacente à certains textes (. . .) » (Gosselin, 0000).

Sans retenir l'idée d'un profilage ni celle d'un *a priori* conceptuel notre étude s'en tiendra à la trame textuelle et l'impact d'un réseau de contraintes modales sur l'interprétation des utilisateurs, En effet, l'intérêt de l'approche de L. Gosselin est qu'elle ne cherche pas à définir une liste de marqueurs spécifiquement dévolue à l'expression de tel ou tel calcul de sens (approche de nature réductionniste) mais contribue à une prise en compte globale (basée sur une conception large des modalités) des phénomènes lexicaux et grammaticaux constitutifs de la cohérence textuelle et de la construction du sens en contexte. De ce point de vue, l'étude des modalités et la formulation de scenarii modaux présentent un intérêt considérable pour la constitution d'une aide à l'interprétation dans un corpus numérique.

4 Démarche de l'analyse du corpus

Dans cette section nous présentons en grands traits la démarche adoptée pour l'analyse du corpus illustré d'un exemple d'une fiche résumant un arrêt de la cour d'appel de Rouen. Nous avons tout d'abord commencé par relever et classer dans un tableau les types de modalités qui caractérisent le texte contenu dans la zone « objet » de la fiche. À l'aide de ses modalités, nous avons retracé le scénario modal qui nous a permis par la suite de mettre en relief la structure interprétative des textes.

4.1 Extrait d'une fiche de la base de données de l'Idit

Référence :

Cour d'appel de Rouen, 2e ch.

17 juillet 2003

Sté P & O Nedlloyd BV et autre c/ Sté Générale de Manutention Portuaire et autres Navire "Bunga Pelangi"

Objet :

Transport d'un lot de cartons de vin empotés dans un conteneur réfrigéré - Température pré-réglée à -18°C au lieu de $+12^{\circ}\text{C}/+14^{\circ}\text{C}$ - Marchandise livrée congelée - Recours du transporteur-affréteur contre le manutentionnaire et le fréteur - Responsabilité du manutentionnaire (oui) - Preuve de ses manquements aux obligations fixées par le contrat de manutention - Défaut de branchement du conteneur sur le terminal et de vérification de l'indexation de la température - Responsabilité du fréteur (non) - Contrat d'affrètement d'espace - Clauses - Fréteur responsable de l'indexation de la température des conteneurs (non) - Température pré-réglée par l'affréteur - Connaissance de la température requise par le fréteur (non) - Obligations du fréteur limitées à la fourniture de l'alimentation électrique et à la vérification du bon fonctionnement des conteneurs.

Relevons que les périphrases de la zone objet sont construites sur le principe de l'« Economy Grammar » (Halliday, 1967) principe qui participe d'un mouvement de décontextualisation bien connu en documentation. La qualification juridique des faits, fondée sur un principe dichotomique (oui/non) intervient juste après leur rappel : Responsabilité du manutentionnaire (Oui) » alors qu'elle n'apparaît bien entendu que dans les conclusions de l'arrêt de la cour.

4.2 Relevé des modalités

Afin de pouvoir retracer l'enchaînement des modalités et d'établir un scénario modal pour la situation évoquée, nous devons tout d'abord relever les principaux marqueurs modaux qui jouent un rôle dans l'argumentation. Rappelons que les textes étudiés sont une première réécriture des arrêts de cours d'appel présentés comme une suite de mots-clés. Les textes se caractérisent par des modalités intrinsèques (associées ou dénotées aux lexèmes) et/ou des modalités inférées sur la base de connaissances encyclopédiques (Gosselin, 2010). Le tableau ci-dessous présente les marqueurs modaux mis en gras et les types de catégories modales exprimées, ainsi que les valeurs.

Modalités				
Les marqueurs modaux	catégories	Valeurs	Obtenues par inférence	valeurs
Température pré-réglée à 18°C au lieu de $+12^{\circ}\text{C}/+14^{\circ}\text{C}$.	Appréciative	Négative		
Marchandise livrée congelée .			Appréciative	négative
Recours du transporteur-affréteur contre le manutentionnaire et le fréteur	Aléthique	positif	Boulisque	positive
Responsabilité du manutentionnaire (oui)	Déontique	Positif		
Preuve de ses [manquements aux obligations] fixées par le contrat de manutention.	Aléthique (preuve) Axiologique (manquements) Déontique (obligations)	Positive négative positive		
Défaut de branchement du conteneur sur le terminal et de vérification de l'indexation de la température	Axiologique	négatif		
Responsabilité du fréteur (non)	Déontique	négative		
Fréteur responsable de l'indexation de la température des conteneurs (non)	Déontique	négative		
Connaissance de la température requise par le fréteur (non)	Epistémique	négative	déontique	négative
Obligations du fréteur limitées à la fourniture de l'alimentation électrique et à la vérification du bon fonctionnement des conteneurs	Déontique. Appréciatif	Positive positive		

4.3 Le scénario modal

Légende :

X = La société appelante / Y1 et Y2 = les intimées / Z = représentant de la loi

E1 = Transport d'un lot de cartons de vins empotés d'un conteneur réfrigéré.

Sit1 = température pré réglée -18°C.

Sit2 = marchandise congelée.

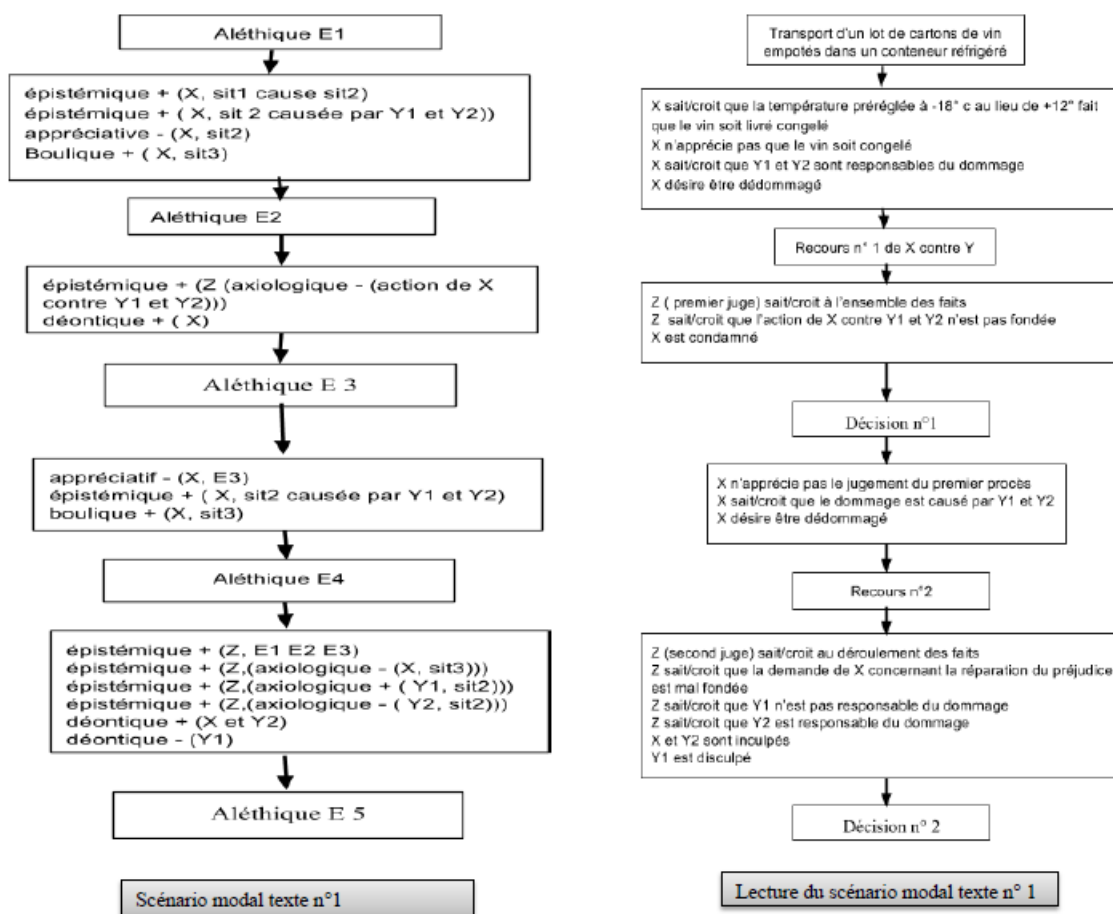
E2 = Premier recours de X contre Y1 et Y2

E3 = Décision du premier procès

E4 = Deuxième recours de X contre Y1 et Y2

E5 = Décision du second procès

La présentation est faite en deux plans. Nous avons dans le côté gauche de la figure le déroulé modal organisé grâce aux marqueurs modaux relevés dans le texte. A droite nous avons mis en correspondance les passages textuels correspondants. Cette démarche nous a permis de relever la structure chronologique et argumentative de la zone objet de la fiche supposée être celle du texte plein qu'elle représente.



4.4 Commentaires

Dans le côté gauche de la figure ci-dessus nous apercevons une présentation chronologique et narrative du contenu des arrêts étudiés. Cette chronologie a été mise au jours grâce à des marqueurs modaux qui témoignent de l'ordre dans lequel les événements sont successivement présentés. A droite nous avons repris les passages textuels figurants sur la fiche et nous les avons organisés selon le modèle du scénario. Ce qui ressort de cet enchaînement modal est le protocole discursif mis en place dans l'argumentation juridique. Dans cet arrêt de cour d'appel, le juge reprend d'abord la description des

faits du litige (illustré dans le scénario par les modalités aléthiques), suivi de leur appréciation (négative ou positive, en fonction de l'orientation du juge (condamnation ou acquittement). Les appréciations sont désignées dans le scénario par les modalités appréciatives. Ces dernières sont suivies de jugements d'ordre légal sur les faits (exprimer ce qui est blâmable ou louable) on parle alors de modalités axiologiques. Ce n'est qu'à ce stade du parcours que le juge prononce sa sentence (dire l'interdit et ou le permis) ce qui renvoie au dernier bloc du scénario, les modalités déontiques.

5 Conclusion et perspective

Nous avons présenté dans cet article la notion de scénario modal avec laquelle nous avons reconstitué le possible parcours interprétatif d'un arrêt de cour d'appel. Les premiers résultats obtenus à partir de ce corpus de fiches laissent entrevoir une possible implémentation dans un système d'aide à l'interprétation. À partir du scénario modal nous formulons l'hypothèse de l'existence au sein des textes pleins d'une structuration en « zones d'isotopies³ modales ». Ces zones, une fois délimitées, sont susceptibles de contenir des informations différentes pouvant répondre à des requêtes précises formulées par les utilisateurs. Afin de conduire un utilisateur vers une zone pertinente, nous envisageons un système d'annotation des marqueurs modaux spécifiques à certaines zones de textes. Nous pensons également possible, avec une étude approfondie sur les textes pleins, de proposer différents chemins de lecture et donc d'interprétation correspondants à la pratique en cours de l'utilisateur. L'un de nos objectifs est de valider cette étude de scénario modal sur un corpus de textes pleins. Pour faire cela, nous procédons actuellement à l'aide d'un logiciel textométrique (TXM)⁴ à une étude statistique afin de relever toutes les caractéristiques linguistiques comparatives entre trois corpus de textes (cour d'appel, cour de cassation et tribunaux de commerce). Cette démarche nous permettra de cerner leur structuration spécifique et d'y adapter notre aide à l'interprétation supposant avoir affaire à des pratiques sociales différentes, malgré la proximité des genres. Une fois ce travail accompli, nous mènerons une enquête qualitative auprès d'un panel d'utilisateurs de la base de l'IDIT dans le but de cerner leur parcours de navigation et de le confronter à notre scénario modal. Nous pourrions alors mieux comprendre leurs besoins réels et apprécier la portée de nos résultats encore théoriques au regard des traces de leur navigation. Le lecteur humain est évidemment le meilleur interprète d'un document, par comparaison à tout traitement automatique et machinal, qui n'est d'ailleurs jamais une véritable lecture.

Remerciements

Les travaux menés dans cet article bénéficient d'un financement du Grand Réseau de Recherche Culture & Société en Normandie (CSN) de la Région Haute-Normandie. Et nous avons le soutien du projet PlaIR 2,0 (2013-2016) pris en charge par le GRR Transport Logistique Technologies de l'Information de la Région Haute-Normandie.

Références

- FERRARI S., MAUREL F., BEUST P., MAUGER S., HOLZEM M., BAUDOUIN N., TRUPIN E., SAIDALI Y., LABICHE J. & DIONISI D. (2012). Pour une recherche d'information et une veille juridique interactives et socio centrées. ENT éactif & veille en droit du transport. *Ingénierie des Systèmes d'Information*, **2**, pages 17–40.
- FILLMORE C. (1982). Frame semantics. in linguistics in the morning calm. *Revue Ingénierie des systèmes d'information - RSTI série ISI-*, **17**, 111–137.
- GOSELIN L. (2000). à paraître : De l'opposition modus/dictum à la distinction entre modalités extrinsèques et modalités intrinsèques. *Bulletin de la Société de Linguistiques de Paris*, p.35.
- GOSELIN L. (2010). *Les modalités en français : la validation des représentations*. Rodopi.
- HALLIDAY M. A. K. (1967). Notes on transitivity and theme in English : Part 2. *Journal of Linguistics*, **3**(02), 199–244.
- RASTIER F. (2001). *Arts et sciences du texte*.
- SAIDALI Y. (2002). *Modélisation et acquisition de connaissances : application à une plate-forme de traitement d'images. Thèse d'informatique automatique et systèmes*. PhD thesis, Université de Rouen.

3. La notion d'isotopie est définie par F. Rastier (1972) comme l'effet de la récurrence d'un même sème, dit sème isotopant.

4. Un logiciel gratuit de traitement statistique des textes qui se développe dans l'ENS de Lyon.

SCHANK R. C. & ABELSON R. P. (1977). *Scripts, Plans, Goals, and Understanding : An Inquiry Into Human Knowledge Structures*. Psychology Press.