

---

# Une expérience d'un déploiement industriel de LELIE: une relecture intelligente des exigences

Juyeon Kang<sup>1,2</sup>, Patrick Saint-Dizier<sup>2</sup>

1. Prometil

42, avenue du Général de Croutte 31100 Toulouse, j.kang@prometil.com

2. IRIT-CNRS

118, route de Narbonne 31062 Toulouse cedex, stdizier@irit.fr

---

*RESUME.* Cet article montre à travers de nombreux exemples et de nombreuses situations industrielles comment le prototype universitaire LELIE, un outil intelligent d'aide à la rédaction de documents techniques, a été transféré et valorisé dans un cadre industriel. LELIE est tout d'abord adapté de façon générale au traitement des exigences, suivant les normes et pratiques en vigueur. Nous montrons ensuite comment LELIE pour les exigences est adapté au contexte de chaque industriel, en particulier au niveau des ressources linguistiques et du suivi des pratiques des rédacteurs d'exigences.

*ABSTRACT.* In this article, we show how the academic prototype LELIE, an intelligent authoring tool for technical texts, developed at IRIT, has been transferred to an industrial context. We first show how it has been customized to the norms and general practices of requirement authoring. We then show how LELIE for Requirements is adapted to the local context of specific companies, in terms of linguistic resources and companies authoring practices of technical writers.

*MOTS-CLES :* Informatique linguistique, exigences, aide à la rédaction de textes techniques

*KEYWORDS:* Computational linguistics, requirements, technical text authoring tool

---

## 1. Problématique et motivations

Les documents techniques (procédures, exigences, spécifications) forment un genre conceptuel particulier qui suit des contraintes linguistiques fortes en termes de choix lexical (par exemple, quel terme utiliser dans une terminologie de termes métier), de syntaxe (structure des phrases et énoncés), typographique, de style et de contraintes métier variées. Les documents techniques sont conçus de telle façon qu'ils soient faciles à lire, efficaces (rien de superflu et d'un niveau de granularité adéquat), adaptés au lecteur ou à l'opérateur et le moins ambigus possibles. Ces documents doivent être explicites de façon à laisser le moins possible d'espace aux interprétations personnelles des utilisateurs, génératrices de risques. Dans ce but, plusieurs principes et recommandations de rédaction se sont développés : des principes généraux : ceux de la langue contrôlée générale (CNL) et ceux de nombreuses entreprises sous forme de guides ou de recommandations. Le résultat est un ensemble un peu hétérogène de principes qu'il faut adapter à chaque situation.

Les principes de la CNL incluent par exemple : (1) l'interdiction de l'emploi de la négation, du futur ou du passif, de termes passe-partout ou vagues, de termes flous, (2) la limitation de la complexité des énoncés : coordinations et relatives simples et limitées, phrases de longueur modérée, (3) le contrôle des titres, des renvois, du style : régularité des énumérations, position de certaines structures (conditions, expression du but, etc.). Ces principes, dans la réalité de la rédaction technique, sont souvent perçus comme trop contraignants et très complexes à mettre en œuvre. Lorsque l'on ajoute à cette difficulté le fait que les documents techniques sont souvent produits à partir de documents plus anciens via des 'copier-coller', qu'un texte est souvent rédigé par plusieurs rédacteurs et que la relecture d'un document est souvent fastidieuse et couteuse en temps, on comprend qu'il reste dans les textes techniques, même dûment validés, de nombreuses erreurs qui nuisent à leur bonne intelligibilité. Il en résulte des coûts parfois très élevés pour régler a posteriori des problèmes qui auraient pu être évités par une relecture plus soignée.

### 1.1 Le noyau de LELIE

Ces considérations ont motivé le développement du projet LELIE (projet ANR) (Barcellini *et al.* 2012), (Saint-Dizier, 2014), qui permet la détection de plusieurs catégories d'erreurs dans les documents techniques par le biais d'alertes. LELIE est un système à base de règles qui permet de traiter les contraintes générales de la langue contrôlée aussi bien que les contraintes de rédaction métier. LELIE autorise la détection des erreurs en fonction de la structure de discours ainsi que la définition de degrés de gravité pour chaque type d'erreur, ce qui permet au rédacteur de planifier les corrections d'une façon adaptée à ses priorités. LELIE a été implémenté sous la forme d'un noyau qui couvre les principales erreurs rencontrées en rédaction technique de façon indépendante du domaine. L'implémentation a été réalisée sur la plateforme d'analyse <TextCoop> (Saint-Dizier, 2012). LELIE suit un modèle de

spécification linguistique déclaratif, il est conçu pour être flexible et facile à faire évoluer.

Pour illustrer les possibilités du noyau de LELIE, le tableau ci-dessous résume les principales catégories d'erreurs traitées avec les taux d'erreurs rencontrées. Les documents analysés sont un ensemble de procédures, d'exigences, de spécifications de sécurité et de manuels de produits. Les documents proviennent de trois sociétés S1, S2, S3 de secteurs industriels différents qui restent anonymes à leur demande. Les documents considérés sont des documents validés et opérationnels. 300 pages ont été considérées dans cette analyse. Le tableau ci-dessous propose une moyenne pour chaque erreur pour un document typique de 30 pages :

*Tableau 1. Résultats d'analyse du noyau de LELIE*

Type d'erreur	Nbre pour 30 pages	S1	S2	S3
Termes flous	66	44	89	49
Déverbaux	29	24	14	42
Modaux dans les instructions	5	0	12	1
Pronoms avec référence incertaine	22	4	48	2
Négation	52	8	109	9
Structures de discours complexes	43	12	65	50
Coordinations complexes ou ambiguës	19	30	10	17
Constructions N+N lourdes ou compléments de noms trop longs	46	58	62	15
Passifs	34	16	72	4
Futurs	2	2	4	1
Phrases trop complexes ou trop longues	108	16	221	24
Enumérations irrégulières	moyen	bas	élevé	moyen
Références incorrectes à d'autres sections ou figures	13	33	22	2
<b>Total erreurs</b>	<b>439</b>			

Notons que dans ce tableau, la société S3 a fait relire ses textes par une société de rédaction spécialisée, il y demeure cependant un bon nombre d'erreurs. On observe que la distribution des erreurs est très diversifiée selon les sociétés, cela tient en particulier à la complexité des textes : ceux de S2 sont par exemple nettement plus complexes que ceux de S1. Enfin, on note qu'il y a en moyenne 15 erreurs par page, ce qui fait environ une alerte toutes les 2 à 3 lignes, hors erreurs liées au métier, ce qui est très considérable et motive largement l'emploi de LELIE.

Le noyau de LELIE est conçu indépendamment d'un domaine avec un lexique général et des règles de détection génériques. Par conséquent, selon les domaines et les textes, on observe un certain nombre d'oublis ou d'alertes inappropriées. Globalement, on observe, selon les types d'erreurs, un niveau de silence qui peut aller jusqu'à 15% et un taux de bruit allant jusqu'à 30%, en particulier pour les termes flous qui sont souvent flous en fonction du contexte métier. Il est clair qu'une adaptation au domaine et aux règles de la société permet d'améliorer grandement la qualité du système et devrait le rendre acceptable dans des contextes opérationnels. Cette communication présente une démarche d'industrialisation du noyau de LELIE appliqué aux exigences.

## **1.2 La structure du noyau de LELIE**

LELIE est implémenté en Dislog, une extension à Prolog qui s'exécute sur <TextCoop>, un méta-interpréteur dédié en premier lieu à l'analyse du discours (Saint-Dizier, 2012 ; 2014). LELIE s'appuie sur Dislog de la façon suivante :

-Les structures incorrectes dans les énoncés d'exigences sont reconnues sur la base de templates qui décrivent ces structures,

-La production des alertes se fait par le biais de la réécriture : le template initial est réécrit avec l'insertion des alertes, par exemple sous forme XML, à l'endroit approprié.

-Les templates proposés par Dislog étendent le pouvoir expressif des expressions régulières : (1) en associant des structures de traits typés aux symboles préterminaux et non-terminaux, (2) en développant des symboles spécifiques, 'gaps', qui permettent d'ignorer des suites finies de mots qui n'ont pas d'intérêt pour ce template, tout en permettant de spécifier des contraintes sur ce qui est ignoré (par exemple les négations ne peuvent pas être ignorées), (3) en introduisant la possibilité de raisonnement, par exemple terminologique ou sur les connaissances du monde, pour affiner un diagnostic d'erreur, (4) enfin, en permettant d'effectuer des calculs pour produire une représentation de l'alerte qui soit précise et satisfaisante.

LELIE a une architecture linguistique modulaire, il est composé :

-d'un ensemble de templates qui décrivent les structures erronées, les autres structures étant par défaut correctes. Les règles sont relativement génériques et sont associées à des filtres pour traiter les cas particuliers (1) soit qui requièrent un traitement spécifique (2) soit qui ne sont pas des erreurs (par exemple le pronom indéfini 'it' ne pose pas de problème de référence pronominale incertaine).

-d'un ensemble de lexiques : langue générale, langue métier, marques spécifiques de certaines constructions.

A ces composants s'ajoutent des procédures de calcul et le moteur TextCoop. Le format général d'une règle de détection d'erreur est :

[patron d'erreur] → [patron de correction] {Contraintes, filtres}.

Un exemple typique de règles de LELIE est la détection de formes passives dans la proposition principale de l'exigence, où apparaît de modal typique de l'exigence (*shall*) :

*it shall be demonstrated....* doit être évité:

[pro(*it*), modal, aux(*be*), gap, verb(past participle)] → [pro(*it*), modal, <passif>, aux(*be*), gap, verb(past participle), </passif>], {}.

Cet exemple, un peu simplifié, montre comment les alertes sont insérées dans le résultat. Il n'y a pas de contraintes spécifiques ici, mais on pourrait tolérer certains verbes courants au passif, qui seraient alors filtrés (la règle ne s'applique alors pas) :

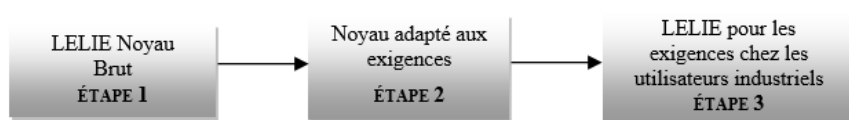
{diff(verb,[liste de verbes non concernés par l'alerte])}.

## 2. Solutions et techniques pour LELIE Exigences

Le noyau de LELIE a été conçu comme un modèle générique qui doit ensuite être adapté en fonction du sous-genre textuel considéré. C'est en particulier le cas pour l'aide à la rédaction des exigences avec leurs erreurs spécifiques. Nous considérons dans ce document des textes composés uniquement d'exigences ou bien des textes où les exigences sont clairement identifiées. L'analyse des erreurs porte uniquement sur les exigences. Nous nous focalisons sur l'anglais.

L'adaptation de LELIE aux exigences (Étape 2 de la figure 1) puis à un métier ou à une société spécifique (Étape 3 de la figure 1) se fait en deux étapes, afin de préserver une bonne organisation hiérarchique des traitements :

Figure 1. Etapes de développement de LELIE pour les exigences



### 2.1 La rédaction des exigences : outils et approches

Dans cette section, nous présentons brièvement quelques principes de rédaction définis dans les normes IEEE et les langues contrôlées, généraux tels que spécifiés dans (Weiss, 1991; O'Brien, 2003; Unwalla, 2004; Wyner, 2010; Alred *et al.* 2012; Kuhn, 2013; 2014) ainsi qu'une synthèse des consignes de rédaction développées dans les sociétés avec lesquelles nous avons développé nos expérimentations.

Les exigences se présentent souvent sous forme de listes ou bien de phrases composées identifiées et cadrées par un ID (par exemple, [REQ-SYS\_0001]) et structurées en paragraphes. Une exigence se compose d'une ou de plusieurs phrases

qui en définissent par exemple les contraintes ou les différentes facettes (Grady, 2006 ; Hull, 2011). Les règles de la rédaction qui s'appliquent aux exigences ont été identifiées à partir de normes IEEE, de recommandations spécifiques aux exigences et de principes des langues contrôlées (Normes IEEE830-1998, ISO/IEC/IEEE 29148 :2011, ISO/IEC 12207, ARP4754, ASD-STE 100<sup>1</sup>, VTT-R-01067-14<sup>2</sup>). Les contraintes de rédaction précisent la syntaxe et la sémantique ainsi que le style et le lexique que les rédacteurs doivent respecter. Les principales contraintes sont : (1) il faut éviter les termes flous ; (2) les termes utilisés doivent être consistants ; (3) il faut éviter trop de conjonctions de coordination ; (4) il faut utiliser le modal *shall* pour une exigence mandataire et *must* pour une exigence réglementaire; (5) il ne faut exprimer qu'une seule action par exigence, etc. A partir de ces règles, on peut développer un modèle linguistique de rédaction des exigences. Ce modèle s'appuie sur des ressources lexicales et des structures grammaticales typiques.

La rédaction des exigences est une tâche délicate car celles-ci doivent être précises, complètes et non-ambiguës. Le niveau d'imprécision doit être minimisé de façon à ce que le produit réalisé à partir de celles-ci ou les comportements qui en résultent soient le plus possible conforme à la demande ou aux normes. Pour ce faire, une approche à base de templates (Arora *et al.*, 2013) a été mise en place qui propose les structures de langue typiques qui peuvent être utilisées, à l'exclusion d'autres, comme par exemple :

[SUJET, MODAL, VERBE, OBJET(matériel), LIEU, BUT]

Cette approche a fait l'objet de quelques implémentations, où la conformité à ces templates est contrôlée, comme dans RAT-RQA (<http://www.reusecompany.com/>) ou Rubric (<https://sites.google.com/site/rubricnlp/>). Toutefois, dans la pratique quotidienne, en particulier lorsque les exigences sont complexes, ou pour certaines catégories d'exigences générales, ce modèle à base de templates n'est pas suivi ou ne l'est que partiellement. De plus, en termes de traçabilité et de flexibilité, il sera très difficile de suivre et mettre à jour les exigences rédigées à base de ces templates lorsque l'on modifie quelques éléments de ceux-ci. La plupart des rédacteurs souhaitent en effet avoir la liberté de s'exprimer assez librement. LELIE ne propose pas de templates a priori, il intervient a posteriori, en produisant des alertes par rapport aux contraintes de rédaction qui ont été spécifiées ou sélectionnées par le rédacteur.

Le système RAT-RQA offre des fonctionnalités dans un registre proche de notre travail, appliqué aussi aux exigences, mais avec très peu de flexibilité dans la prise en compte de l'opérateur et de la tâche. Ce système est doté d'une interface simple pour son implantation dans une entreprise (où sont essentiellement gérés les aspects lexicaux propres à la société). Ce système permet de définir des règles de rédaction propres à une entreprise (les boilerplates), mais il ne traite pas du style, de la cohésion et ne prend en compte les contraintes discursives dans les textes.

---

<sup>1</sup> Simplified Technical English, Aerospace and Defense, by Industries Association of Europe, Issue 5, 2010

<sup>2</sup> Controlled natural language requirements in the design and analysis of safety critical I&C systems, Research Report, SAREMAN Project, 2013

Le système proposé par Attempto (Fuchs, 2012) (<http://attempto.ifi.uzh.ch/site/>) implémente les principes de la rédaction contrôlée de façon générale, appliquée à l'anglais. Un langage formel a été associé (ACE, basé sur la logique et la programmation en logique) qui permet de développer des notations et des méthodes formelles en traitement de la langue naturelle. C'est donc un outil à spectre large en TALN, mais où l'adaptation à un contexte industriel particulier semble laissé ouvert à l'utilisateur, sans support méthodologique particulier.

## **2.2 LELIE dans le contexte général des exigences**

LELIE adapté aux exigences (*LELIE for Requirements*) se base sur le modèle générique du noyau présenté dans la section 1.2 et utilise ses ressources lexicales et règles de base. Ces ressources lexicales contiennent d'une part, non seulement un lexique générique comme les prépositions, les modaux et les verbes d'action mais aussi des ressources davantage orientées vers l'aide à la rédaction telles que les adjectifs et adverbes identifiés comme étant flous a priori, et d'autre part, les règles décrivant des erreurs de rédaction comme, par exemple, le repérage des compléments de noms en cascade à éviter. Ces éléments étant développés indépendamment du domaine et du type de textes (c'est-à-dire, applicable à tout type des documents techniques), il nous fallait une évaluation plus précise des performances sur ces derniers afin de déterminer les points d'amélioration. L'évaluation a été effectuée sur un corpus de 100 pages de texte qui ne contient que des exigences de trois domaines différents : télécommunication, transport, sécurité. A titre indicatif, pour la détection des mots flous, le résultat de cette évaluation a atteint à peu près à 30% de précision et 70% de bruit, et pour les pronoms personnels à référence incertaine, 58% de précision et 42% de bruit (faux positifs).

Cette analyse nous a permis d'organiser les tâches suivantes:

- a) Synthétiser les caractéristiques de la rédaction des exigences dans les standards et les normes
- b) Déterminer les erreurs les plus fréquentes dans la pratique de la rédaction
- c) Déterminer les règles cruciales et prioritaires de la rédaction à implémenter dans notre système
- d) Revoir les vocabulaires pertinents ou non pertinents dans le lexique (termes flous, termes à éviter, ambigus, etc.)

Les règles d'analyse des erreurs ont été réalisées sur la base des standards et normes de la rédaction et ont été renforcées à partir des erreurs vues dans la pratique de la rédaction lors de l'analyse de corpus (méthode manuelle, effectuée par deux rédacteurs techniques et un linguiste), composé de 25 textes d'exigence de différents domaines (télécommunication, transport, banque, écologie).

Nous présentons une partie des règles génériques en fonction de niveaux d'analyse linguistique. Au niveau du lexique, nous vérifions les emplois adaptés des termes :

- 1) Termes flous : par ex. *wherever possible, suitably, adequately*

*Ex.: Intimate contact between dissimilar metals shall be avoided wherever possible.*

Chaque terme a un degré de gravité a priori sur une échelle qualitative de flou.

- 2) Termes inutiles : par ex. *also, then, always*

Inutiles au moins dans certains contextes.

- 3) Expressions interdites ou incomplètes : par ex. « *and/or* », « TBD », « TBC », « etc. »...

Au niveau de la grammaire, nous contrôlons les formes complexes :

- 4) Utilisation de la négation et de la double négation

- 5) Emploi de pronoms à référence incertaine : par ex. *their, them, these*,

- 6) Utilisation de relatives complexes : par ex. *that x...which...*

- 7) Utilisation de conjonctions de coordination : par ex. *and, et/ou or*

*Ex.: FCV shall have a visual mechanical position indicator with the full open and full closed positions clearly and permanently marked into the valve body adjacent to the mechanical position indicator, using the words "OPEN" and "CLOSED".*

- 8) Renvoi difficile : par ex. *below, above, see...*

*Ex.: Testing shall consist of resonance search, endurance dwell and endurance sweep cycling as defined below.*

- 9) Phrases complexes: par ex. la longueur de la phrase, la présence de nombreux acronymes ou de termes techniques non définis, les structures complexes (par ex. [SUJET],[CONDITION],[ACTION]-[OBJET]).

- 10) Construction passive : par ex. imprécision sur qui doit effectuer l'action

*Ex.: it shall be demonstrated as defined in ...*

Au niveau de l'uniformité du style, nous contrôlons la structure si les exigences sont rédigées de façon homogène :

- 11) Coordination des exigences

*Ex. : REQ1 Testing shall consist of resonance search, endurance dwell and endurance sweep cycling as defined below and REQ2 shall be performed in the sequence indicated*

- 12) Régularité des temps : par ex. *shall set or set* au présent

- 13) Position structurelle: par ex. CONDITION-SUJET-ACTION-OBJET

*Ex.: [CONDITION] When signal x is received, [SUJET] the system [ACTION] shall set [OBJET] the signal x received bit [CONSTRAINT] within 2 seconds.*

- 14) Enumérations hétérogènes; par ex. contrôle sur les listes énumérées

- a. *system interaction*

- b. *user interface*

- c. *update is not applicable* : non homogène avec les 2 précédents



Ces fonctions sont adaptées à la tradition rédactionnelle des différentes sociétés. Par exemple, la fonction liée à la détection des conjonctions de coordination a six sous-catégories comme deux conjonctions dans une proposition conditionnelle, plus de trois conjonctions dans une proposition subordonnée, etc.

### 2.3 Prise en compte du métier

Pour que la détection des erreurs soit adéquate à la pratique de la rédaction industrielle, il est primordial de prendre en compte les connaissances des pratiques de rédaction du métier. Par exemple, certaines industries veulent voir la condition dans laquelle une exigence doit être effectuée toujours devant la proposition principale qui exprime l'action (<CONDITION> *Si la température est supérieure à 36°C*, <ACTION> *le système doit fonctionner...*) et/ou il peut y avoir les termes préférés (par ex. *prise de courant* doit être remplacé par *fiche à N broches*) et les termes métiers. Certains termes généralement flous comme *probable* ou *normale* peuvent être non flous lorsqu'ils sont utilisés dans un contexte industriel précis. Dans ce sens, nous avons développé les filtres qui s'appliquent selon le contexte dans lequel la règle et le lexique sont utilisés. Le mot *normale* sera détecté comme flou dans un contexte et filtré dans un autre contexte. Ces filtrages augmentent le taux de pertinence du système en permettant de minimiser les faux positifs.

Le résultat de cette prise en compte du métier est mesuré par une évaluation qualitative. Le corpus de l'évaluation est composé de 50 pages extraites de six spécifications sur trois domaines : aérospatial, télécommunications, transports. Ce corpus ne comporte que des exigences, il est validé comme étant suffisamment complexe par les rédacteurs techniques. Le corpus a été pré-annoté manuellement par le valideur-rédacteur sans regarder auparavant les annotations faites par LELIE.

Tableau 2. Résultats de l'évaluation après l'adaptation de LELIE aux exigences

Type d'erreurs évaluées	Précision	Rappel	Silence	Bruit (Faux positifs)
Termes flous	75% (295/395)	77% (295/383)	23% (88/383)	25% (100/395)
Trop de conjonctions de coordination	74% (17/23)	93% (17/18)	6% (1/18)	26% (6/23)
Trop de négations	100% (1/1)	100%	0%	0%
Exigences multiples	63% (37/59)	57% (37/65)	43% (28/65)	37% (22/59)
Expressions interdites/incomplète	88% (63/72)	72% (63/88)	28% (25/88)	13% (9/72)
Constructions passives	93% (101/109)	69% (101/146)	31% (45/146)	7% (8/109)

### **3. L'implantation industrielle**

Le noyau de LELIE dans un contexte des exigences étant réalisé, nous nous penchons à présent sur la prise en compte précise des métiers. Ce travail s'appuie sur les rédacteurs de six sociétés. Nous décrivons dans cette partie nos différentes expériences. Pour des raisons de confidentialité, les sociétés restent anonymes et les résultats présentés sont une synthèse de nos évaluations.

#### ***3.1 Analyse des spécifications de différents domaines***

##### ***3.1.1 Approche***

Les spécifications techniques envoyées par les sociétés en tant que valideurs-utilisateurs, sous une forme variée (le format .dma exporté de DOORS, Word, Excel), ont été étudiées manuellement pour analyser les structures et les termes de chaque exigence (identifiée par un ID).

Le premier objectif de cette étude a été de déterminer à quel point les règles de l'outil sont en adéquation par rapport à l'observé. Nos observations ont entraîné l'ajout de données lexicales et de filtres à LELIE en complément des premiers ajouts. Ces ajouts plus spécifiques doivent être contrôlés et sont à l'origine d'une méthode de customisation à l'étude.

Le deuxième point important a été le développement d'un lexique spécifique au domaine et au client. Dans cette étape, les termes métiers et termes préférés ont été identifiés et ajoutés dans nos lexiques spécifiques selon le domaine et le client. Le lexique générique reste stable ayant en moyenne 75% de pertinence comme le montre le résultat de l'évaluation (voir le tableau 2), cependant le lexique spécifique des termes métiers et préférés doit évoluer, en particulier, lorsque l'on applique l'outil à un nouveau domaine ou client.

##### ***3.1.2 Résultats***

Les résultats de l'analyse des erreurs dépendent largement du fait que l'on adapte ou non le lexique et les règles au contexte du projet appliqué. Notamment, la pertinence de la détection des termes flous peut être considérablement augmentée (environ 20 à 30%) par cette adaptation.

A chaque déploiement industriel, nous avons effectué une analyse statistique des résultats en matière de pertinence. Nous présentons dans le tableau suivant le taux des exigences ayant au moins une alerte de LELIE. Il s'agit d'une synthèse pondérée sur les taux obtenus à partir des résultats statistiques de six sociétés. Ces résultats (voir le tableau 3 et la figure 2) sont une indication de la pertinence de l'outil, ils constituent un ensemble d'éléments pour une évaluation indicative.

Tableau 3. Nombres des erreurs par type sur toutes les spécifications testées

Sur un ensemble de <b>4403</b> exigences analysées, <b>2708</b> exigences ( <b>62%</b> ) avec au moins une alerte de LELIE	Termes flous	1332
	Trop de conjonctions de coordination	91
	Trop de négations	29
	Expressions à éviter	715
	Construction passive	967
	Exigences multiples	325
	Termes inutiles	21
	Contrôle de la structure de l'exigence	108
	Contrôle de la partie Rationale	79
	Exigence trop longue	306
<b>Total</b>	<b>3973</b>	

Comme le montre le tableau 3, l'outil a détecté un problème en moyenne dans 62% des exigences et les taux minimal et maximal sont respectivement de 52% et 72%, ce qui est très élevé.

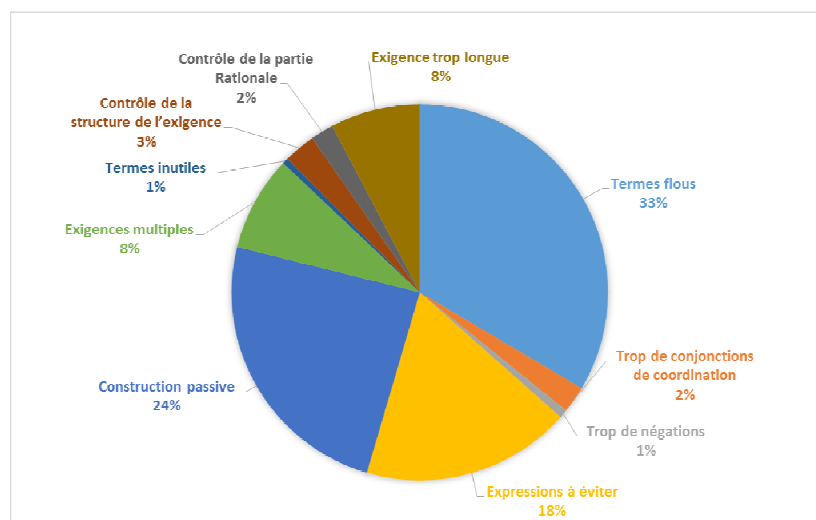


Figure 2. Distribution des erreurs par type

Ce résultat nous donne plusieurs indications par rapport à la pratique rédactionnelle des sociétés concernées : 1) le style de la rédaction (par ex. certaines spécifications majoritairement rédigées sous une forme passive), 2) le niveau de

maturité de la rédaction (par ex. certaines spécifications contenant de nombreux termes indiquant un contenu incomplet comme *To Be Determined*, moins de termes flous dans les spécifications matures que celles de tout début), 3) l'organisation syntaxique de la proposition de l'exigence, 4) le niveau discursif de l'exigence (par ex. la structure du discours comme condition, but, circonstance, illustration...).

Le résultat de certaines sociétés était plus varié que d'autres dans le sens où les spécifications d'une même société ont été rédigées chacune avec un style différent ou pas, par exemple, une spécification avec la forme passive (par ex. *shall be Verbed*), une autre avec la forme négative (par ex. *shall not*). Cela montre que la façon dont les rédacteurs rédigent leurs spécifications n'est pas homogène même au sein d'un groupe des rédacteurs ou d'une société en absence de guides spécifiques.

### 3.1.3 Retours des évaluations

Les évaluations que nous avons menées à l'occasion de différents déploiements industriels sur cet outil et les discussions avec les rédacteurs techniques peuvent être présentées selon deux perspectives complémentaires: l'analyse interne et le retour des clients.

Le premier point correspond à l'analyse des nouveaux besoins suscités par les possibilités de l'outil, cela concerne les trois aspects suivants :

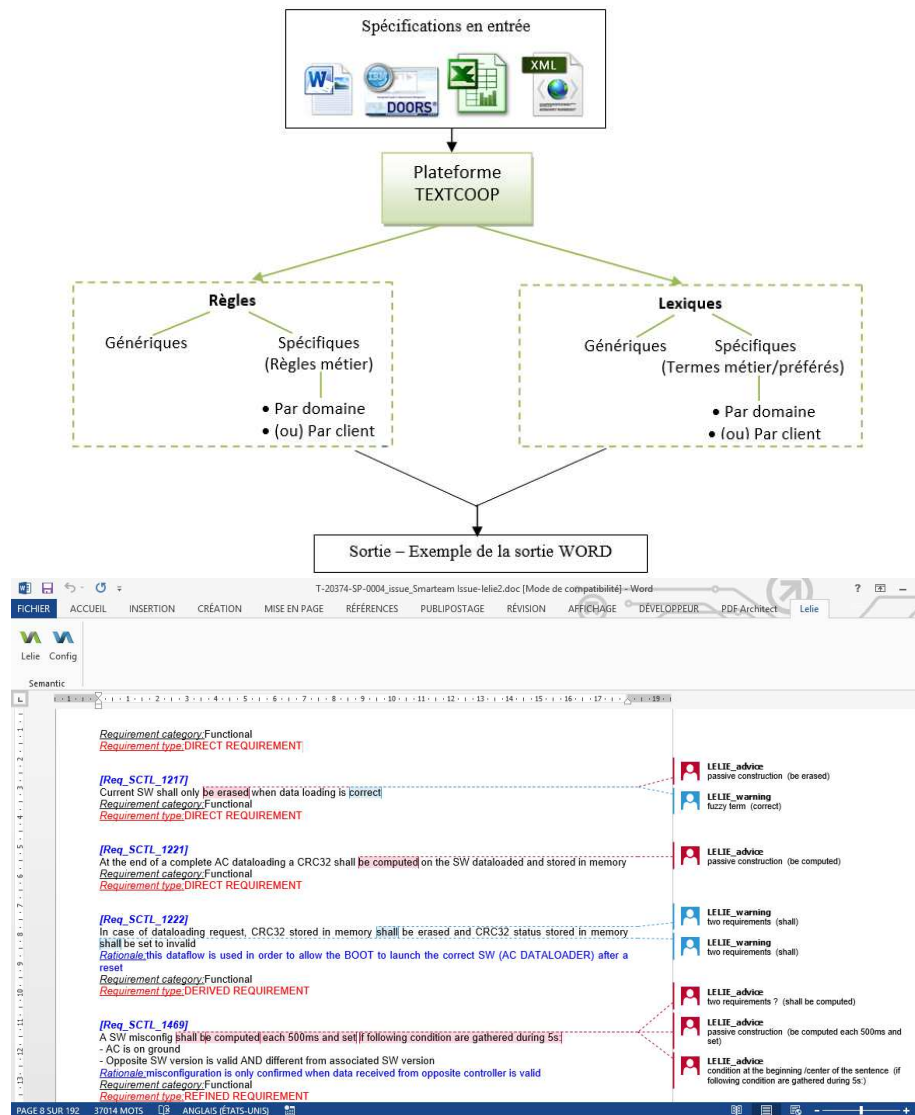
- 1) Aspect fonctionnel : nouvelles fonctions (par ex. analyse de la redondance, de la cohésion ou de la cohérence entre les exigences)
- 2) Aspect technique : amélioration d'interface, modes d'intégration des commentaires aux utilisateurs dans l'outil
- 3) Aspect ergonomie cognitive : adéquation des alertes au fonctionnement du rédacteur.

Le second point, le retour des clients, concerne davantage la possibilité d'activation/désactivation de certaines règles et vocabulaires spécifiques par leur soin. Nous observons que (1) certaines erreurs identifiées comme étant critiques sont acceptées sans problème dans la pratique quotidienne des rédacteurs de certaines sociétés et (2) certains vocabulaires identifiés comme étant problématiques dans la plupart des cas sont acceptables et compréhensibles sans difficulté par les rédacteurs de certaines sociétés.

Ces expériences industrielles nous permettent de constater que l'adaptation d'un domaine à l'autre est plus coûteuse que d'un client à l'autre sur un même domaine. Lorsque l'on aborde un nouveau domaine, le volume du lexique à ajouter peut être de 70% à 100% du lexique existant pour les termes métier. Il s'agit de développer un nouveau lexique spécifique des termes métier à un domaine. Pour les termes génériques, cela correspond à 3% à 5% de l'ajout des nouveaux vocabulaires dans le lexique générique existant, ce qui est très minime. Pour minimiser le travail d'entrées lexicales pour chaque déploiement du domaine, nous développons une plateforme d'acquisition. Nous constatons que cet outil est indispensable pour minimiser le coût de déploiements industriels et sécuriser la mise en place de l'outil chez les utilisateurs finaux. Enfin, les règles de détection existantes restent

relativement stables pour différents domaines car les exigences sont souvent rédigées selon le même schéma, indépendamment du domaine, par contre, des règles spécifiques ou des filtres peuvent être ajoutées lorsqu'une société a ses propres règles de rédaction. Nous représentons dans la figure 3 l'architecture fonctionnelle élaborée et actuellement mise en opération dans le contexte de l'ingénierie des exigences.

Figure 3. Architecture fonctionnelle de LELIE pour les exigences



Le tableau 4 donne le nombre de règles de base et de filtres contextuels ainsi que les entrées lexicales génériques et spécifiques (termes métiers), développées pour les exigences. Ceci donne une mesure des volumes pour le système complet.

**Tableau 4. Nombre de règles et volume des lexiques**

	Nbre de règles		Lexiques (entrées lexicales)	
	de base	de filtres de contextualisation	génériques	termes métier
<b>Noyau de LELIE</b>	13	1	5681	0
<b>LELIE pour les exigences</b>	37	30	8467	680

### 3.2 Problèmes et limites : vers une mémoire de correction

Les principes de rédaction de la langue contrôlée ou bien recommandés par les sociétés sont souvent ressentis comme trop rigides par les rédacteurs techniques : elles manquent de flexibilité et de prise en compte du contexte. En effet, dans *‘l’opérateur doit fermer progressivement la vanne’* l’adverbe de manière *progressivement* n’est pas véritablement flou car l’action de fermer est vue comme ponctuelle. Par contre, il l’est dans : *‘le pilote doit réduire progressivement la vitesse à 145 kts’* car cette réduction de vitesse n’est pas une opération ponctuelle. Dans les deux cas, ce n’est pas la manière qui est mise en cause mais plutôt la durée de l’action, donc sa dimension temporelle. Le contexte *vanne* versus *vitesse avion* est crucial dans la détermination d’un emploi flou ou non des adverbes de manière en général. L’énoncé *‘Ce produit ne doit pas être jeté à l’évier’* est incorrect du point de vue de la CNL dans la mesure où la négation est interdite. Cependant, dans ce cas, il est très difficile, si l’on ne connaît pas exactement l’environnement de travail de l’opérateur, de remplacer la négation par le lieu où il faut jeter le produit.

Une analyse trop directe des énoncés, sur la base d’un système de règles, quelque peu rigide, va générer des alertes incorrectes, donc beaucoup de bruit, même si on associe aux règles de détection des erreurs de nombreux filtres. Nous évoluons à présent vers un système hybride qui associe LELIE à un système d’apprentissage qui va observer les corrections faites (ou pas) par les rédacteurs lorsqu’ils examinent les alertes produites par LELIE. Cette observation va permettre d’apprendre quelles alertes sont de réelles alertes et aussi, de façon très précieuse, comment les erreurs sont corrigées. Une première expérimentation a été développée dans (Kang et Saint-Dizier, 2014) sur le suivi des termes flous. Il en est résulté des spécifications de développement d’une mémoire de correction. Très brièvement, cette expérimentation a permis de dégager les éléments suivants :

- A partir d’un certain seuil de situations non corrigées, pour un contexte donné, l’alerte n’est plus produite. La notion de contexte est définie par quatre termes (verbes, noms, adjectifs) à gauche ou à droite du terme flou et de son focus. La suspension de l’alerte à d’autres termes proches est généralisée via la référence à la terminologie du domaine,

- Les corrections effectuées sont mémorisées. Une généralisation ou un ensemble de généralisations (Daelemans *et al.*, 2005) sont réalisées et proposées aux rédacteurs : l'objectif est de faire un choix unique sur l'une de ces généralisations, et de l'appliquer ensuite systématiquement dans un contexte donné, afin de garantir une bonne homogénéité des corrections. La décision est prise par un administrateur ou collectivement par médiation.
- Des propositions de correction issues de ces généralisations sont proposées. Les corrections sont structurées sur deux niveaux : un patron général de correction qui contient des éléments sous-spécifiés : les recommandations de correction contextuelles, qui sont instanciées en fonction du contexte.

Par exemple, pour un déterminant flou tel que 'a few' appliqué à un nom *Noun* :  
 [a few Noun] → [less than X Noun].

Le patron de correction propose de remplacer *a few* par une expression non floue: moins de N, où N dépend du contexte. Par exemple, 'N=12' dans *take-off a few knots above V1* → *take-off less than 12 knots above V1*, avec

Contexte = [focus: knots, autres: take-off, above V1].

De même pour le cas de *progressivement*, un patron général de correction est induit: [*progressively* Verb(durative)] → [*progressively* Verb(durative) in X(time)].

Où une variable contextuelle indiquera la durée, par exemple *3 minutes* pour la réduction de vitesse de l'avion dans l'exemple donné en 3.2.

Nous allons poursuivre cette expérimentation avec d'autres catégories d'erreurs complexes à régler telles que la négation ou les phrases trop complexes. Ce système de mémoire de correction interviendra alors à la suite des règles de LELIE :

- (1) pour filtrer les alertes qui n'en sont pas, améliorant en cela le confort du rédacteur, ainsi, pour les termes flous, nous sommes passés de 30% de bruit à environ 12%,
- (2) pour proposer des éléments de correction qui restent toujours à l'appréciation des rédacteurs. Ceux-ci apprécient les alertes, mais sont en fait très demandeurs de solutions de correction.

#### 4. Conclusion et perspectives

Nous avons présenté dans ce document les résultats d'une expérience de déploiement industriel à partir du noyau de LELIE, prototype universitaire. Un outil de relecture automatisée des exigences a été développé via différentes étapes d'études en adaptant le modèle générique du noyau de LELIE au contexte des exigences puis à celui des utilisateurs. La dernière version de LELIE pour les exigences est toujours en cours d'évolution dans le cadre du déploiement industriel. Dans l'étape suivante de notre travail, cette version plus finalisée sera évaluée et comparée aux résultats de la première évaluation présentés dans le tableau 2.

Outre ces résultats nous réfléchissons au développement de plusieurs éléments complémentaires: nouveaux types d'erreurs suscités par le dialogue avec les

utilisateurs, par exemple les structures N+N en anglais (Garnier, 2011), interface IHM plus avancée, acquisition lexicale automatisée et éditeur de règles, outil de mise au point pour le linguiste et intégration dans d'autres services (suites éditoriales). D'un point de vue plus fondamental, la notion de mémoire de correction nous paraît essentielle. Nous prévoyons aussi de développer des outils d'analyse de la cohérence entre exigences, mais ceci demeure très complexe.

## Références

- Alred G.J. *et al.* (2012). *Handbook of Technical Writing*. St Martin's Press, New York.
- Arora C. *et al.* (2013). Automatic Checking of Conformance to Requirement Boilerplates via Text Chunking: An Industrial Case Study, *7<sup>th</sup> ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM2013)*. Baltimore, MD, USA.
- Barcellini F *et al.* (2012). Risk Analysis and Prevention: LELIE, a Tool dedicated to Procedure and Requirement Authoring, *LREC*, Istanbul.
- Daelemans W. et van Der Bosch A. (2005). *Memory-Based Language Processing*, Cambridge.
- Fuchs N.E. (2012). First-Order Reasoning for Attempto Controlled English, In *Proceedings of the 2<sup>nd</sup> International Workshop on Controlled Natural Language (CNL2010)*, Springer.
- Garnier, M., (2011) Correcting errors in N+N structures in the production of French users of English, *EuroCall*, Nottingham.
- Grady J.O. (2006). *System Requirements Analysis*, Academic Press, USA.
- Hull E. *et al.* (2011). *Requirements Engineering*, Springer.
- Kang J. et Saint-Dizier P. (2014). Towards an Error Correction memory to Enhance Technical Texts Authoring in LELIE, *CNL'14*, Galway, Ireland, August.
- Kuhn T. (2013). A Principled Approach to Grammars for Controlled Natural Languages and Predictive Editors. *Journal of Logic, Language and Information*, 22(1).
- Kuhn T. (2014). A Survey and Classification of Controlled Natural Languages. *Computational Linguistics*, 40(1).
- O'Brien S. (2003). *Controlling Controlled English. An Analysis of Several Controlled Language Rule Sets*. Dublin City University report.
- Saint-Dizier P. (2012). Processing Natural Language Arguments with the <TextCoop> Platform. *Journal of Argumentation and Computation* 3(2).
- Saint-Dizier P. (2014). *Challenges of Discourse Processing: the case of technical documents*. Cambridge Scholars, UK.
- Unwalla M. (2004). *AECMA Simplified English*, <http://www.techscribe.co.uk/ta/aecma-simplified-english.pdf>.
- Weiss E. H. (2000). *Writing remedies. Practical exercises for technical writing*. Oryx Press.
- Wyner A. *et al.* (2010). *On Controlled Natural Languages: Properties and Prospects*.