

Vers la définition de nouvelles langues contrôlées pour limiter le « risque langagier »

Anne Condamines
CLLE, Université de Toulouse, CNRS, UT2J
anne.condamines@univ-tlse2.fr

RESUME

Le propos de cet article est basé sur plusieurs études réalisées depuis une dizaine d'années à CLLE-ERSS, sur les langues contrôlées (CNLs). Le principal constat est que les CNLs ne sont pas toujours adaptées et utilisables. Par ailleurs, l'impact réel de leur mise en œuvre sur l'amélioration de la « readability » a été très peu évalué. L'article dresse un panorama des problèmes associés aux CNLs et propose de nouvelles pistes de constitution. Dans cet objectif, certaines méthodes de TAL et de psycholinguistique pourraient être mises en œuvre pour améliorer les CNLs existantes ou en proposer de nouvelles.

ABSTRACT

Toward the Definition of New Controlled Natural Languages in order to Prevent Risks Related to Language Use.

The content of this paper is based on several studies carried out over the last decade in CLLE-ERSS lab, on controlled natural languages (CNLs). It finds that they are not always adapted and usable. Another point is that their impact on readability has not been widely measured. The paper gives an overview of the problems linked to CNLs and proposes new directions for designing them. In this aim, NLP and psycholinguistic methods could be used to improve existing CNLs or to propose new ones.

MOTS-CLES : Compréhensibilité, Langues contrôlées, Lisibilité, Tests psycholinguistiques.

KEYWORDS: Comprehensibility, Controlled Natural Language, Readability, Psycholinguistic tests

1 Introduction

Le « risque langagier » est lié au fait que, les langues n'étant pas des moyens de transmission de l'information parfaits parce qu'elles comportent des ambiguïtés, de la polysémie du flou etc., il peut exister une différence entre ce que le locuteur pense avoir transmis et ce que l'interlocuteur comprend réellement. Dans la vie quotidienne, ce décalage ne nuit pas à la communication. Mais, dans des situations elles-mêmes à risque (parce que des vies humaines peuvent être concernées ou de très fortes sommes d'argent engagées), cet écart peut avoir de graves conséquences. Les langues contrôlées (CNLs) ont pour objectif de limiter ce risque en proposant des recommandations d'expression.

Cependant, la mise en œuvre des recommandations des CNLs existantes, si elles donnent assez souvent de bons résultats, pose aussi un ensemble de problèmes. L'un de ces problèmes vient du fait que la définition des langues contrôlées s'est inscrite dans la suite des travaux sur la lisibilité (« readability ») existant depuis le début du XX^{ème} siècle, sans que le recours à la réalité des usages dans des situations restreintes ait été interrogé. Un autre problème est que l'amélioration de la compréhension grâce à la mise en œuvre des langues contrôlées existantes, qui est leur principal objectif, n'a été que rarement évalué.

Les travaux menés sur la terminologie et les CNLs à l'ERSS depuis une dizaine d'années, dans le cadre de différentes thèses ou études (Condamines, 2008), (Condamines, 2009), permettent de proposer de nouvelles pistes de constitution des CNLs, qui mettent en œuvre les méthodes développées en TAL et en psycholinguistique.

2 Origines et caractéristiques des langues contrôlées

2.1 Origine

La réflexion sur les langues contrôlées est apparue dans les organisations après la guerre de 39-45 (Lux, 1998). La première langue contrôlée a été élaborée chez Caterpillar au début des années 1970. Mais, si l'objectif de ces langues était nouveau (permettre une meilleure communication dans les entreprises, en particulier entre partenaires n'ayant pas la même langue maternelle), l'idée d'une langue restreinte favorisant les échanges entre non-natifs n'était pas nouvelle. On la trouvait déjà dans les travaux sur le Basic English dans les années 1930. La notion de « readability » entendue comme « the ease of understanding or comprehension due to the style of writing » (DuBay, 2004,3) est sans doute encore plus ancienne et a entraîné le développement de formules supposées mesurer cette lisibilité, dont la pertinence est toutefois souvent remise en question (Klare, 1976), (Bruce et al.,1981). De fait, toutes les langues contrôlées s'inspirent de ces travaux pourtant déjà anciens. Que ce soit chez Caterpillar, à l'ASD (Aerospace and Defense, anciennement AECMA (Association Internationale des Constructeurs de Matériel Aérospatial)) ou chez INCOSE (International Council on Systems Engineering) pour ne prendre que quelques exemples, on retrouve de grandes similitudes dans les recommandations : elles oscillent entre suggestions générales « Use correct grammar », prescriptions très strictes : « Use definite articles » ou proscriptions « Avoid the use of the oblique symbol » (Ces trois recommandations sont extraites des recommandations de INCOSE). Pour un linguiste, la seule lecture de ces recommandations soulève des interrogations tant elles semblent parfois inapplicables voire inadaptées.

2.2 Caractéristiques

Kuhn (Kuhn, 2014), propose quatre caractéristiques pour décrire toutes les CNLs (en tout cas celles destinées à une utilisation par un humain), orales ou écrites :

- Elles sont basées sur une langue naturelle.
- Le lexique, la syntaxe et/ou la sémantique sont plus restreints que dans la langue naturelle.

- Les « propriétés naturelles » de la langue de base sont maintenues pour permettre une compréhension intuitive.
- Ce sont des langues consciemment construites.

Remarquons que, parmi ces caractéristiques, seule la quatrième permet de distinguer les langues contrôlées des sous-langages. En effet, le lien avec la langue naturelle et la notion de restriction (lexicale, syntaxique ou sémantique) sont constants dans les caractéristiques proposées pour décrire les sous-langages. La seule spécificité des langues contrôlées serait donc qu'elles sont construites et proposées par des organismes ayant vocation à mettre en place un contrôle de l'usage langagier (Kittredge, 2005). Il y aurait ainsi des normes descriptives, qui se mettent en place spontanément et inconsciemment dans toute communauté discursive, et des normes prescriptives, consciemment définies par des organismes à qui ce rôle a été confié. On pourrait résumer notre objectif général par le souci de répondre à la question : comment rendre compatible ces deux types de normes ? La réponse à cette question permettrait de proposer des normes langagières visant à encadrer les échanges verbaux tout en ne contraignant pas trop ces échanges afin de ne pas trop les éloigner de la spontanéité des usages.

3 Problèmes

La mise en œuvre des langues contrôlées existantes est loin d'être toujours faisable ni même efficace. Différents types de problèmes peuvent apparaître.

3.1 Des règles inapplicables

Les règles proposées par les langues contrôlées sont parfois difficilement applicables. Pour ne prendre qu'un exemple, dans le travail qu'il réalise pour sa thèse sur l'analyse de la rédaction des exigences au CNES, Maxime Warnier a montré que la consigne de Incose : « Repeat nouns in full instead of using pronouns to refer to nouns in other requirement statements », si elle était réellement appliquée, alourdirait considérablement les phrases en générant des répétitions qui, loin d'améliorer la compréhension, la rendrait souvent plus difficile (Condamines et Warnier, 2014).

3.2 Des règles qui ne prennent pas en compte l'intention de communication

C'est peut-être le principal problème et pourtant le moins décrit. A l'idée d'une langue contrôlée s'associe l'idée que peuvent exister des situations de communications dans lesquelles seule est visée la transmission d'informations. Or, l'intention de communication n'est jamais totalement absente dans ces situations, même lorsqu'elles sont très contraintes. Voici deux exemples de l'émergence de ces aspects dans la mise en œuvre des CNLs.

Dans sa thèse réalisée avec un financement de l'ENAC (Ecole Nationale de l'Aviation Civile), S. Lopez a étudié la mise en œuvre réelle de la « phraséologie » (qui est la langue contrôlée définie par l'OACI (Organisation de l'Aviation Civile Internationale) pour les communications pilote/contrôleur) (Lopez, 2013). Un des résultats de son étude est que les formules de politesse sont souvent bien plus nombreuses que ce qui est préconisé par l'OACI. Deux explications peuvent être proposées pour expliquer ce phénomène. Tout d'abord, il s'agit d'une communication entre deux

personnes qui doivent collaborer pour faire en sorte que l'avion atterrisse dans les meilleures conditions. Elles doivent donc établir une relation de confiance qui passe par des formules conviviales. L'autre explication possible tient à ce que la situation de communication est inégale. En effet, c'est le contrôleur qui a le pouvoir d'attribuer telle ou telle piste d'atterrissage, à tel ou tel moment, et donc d'augmenter ou de diminuer le temps de vol. Or, les pilotes dépendent d'une compagnie qui les pousse à gagner le plus de temps possible. Ils ont donc intérêt à entretenir de bonnes relations avec les contrôleurs et peuvent avoir tendance à ajouter des formules de politesse.

Voici un autre exemple. Une des règles récurrentes dans les CNLs concerne le fait de ne pas utiliser la voix passive parce qu'elle peut s'avérer trop imprécise en permettant la non-énonciation de l'acteur de l'action. Dans la rédaction des exigences (thèse de M. Warnier), il s'avère que les rédacteurs veulent parfois rester assez vagues. En effet, les exigences qu'ils rédigent ont une valeur contractuelle et, trop de précision pourrait, en cas de conflit, se retourner contre la société qu'ils représentent (des cas de ce type nous ont été rapportés, au CNES par exemple). C'est sans doute pour cette raison qu'ils utilisent très peu la voix passive (comme le veut la recommandation des CNLs) mais utilisent fréquemment le pronom *on* (et, plus précisément, le pronom *on* suivi d'un verbe au futur), lui-même très peu précis quant à l'agent censé réaliser une action (par exemple *On vérifiera que les tables sont reçues en intégralité*). Ils mettent donc en œuvre une stratégie de contournement des recommandations qui aboutit à des structures langagières qui semblent suivre à la lettre la règle mais qui en dénaturent l'esprit.

Dans les deux cas évoqués, les choix langagiers ne sont bien sûr pas faits consciemment ; ils sont liés à l'intégration par les locuteurs de la situation de communication dans laquelle ils se trouvent et des contraintes qu'ils y subissent. La présence de ces régularités langagières, inconsciemment mises en place en lien avec la situation extra-linguistique, caractérise précisément ce que l'on appelle le genre textuel (voir 4.1.1).

3.3 Des règles dont l'efficacité est peu évaluée

Hormis les formules de lisibilité évoquées en 2., qui ne sont basées que sur des mesures quantitatives concernant la longueur des mots et des phrases, et dont la pertinence est controversée, très peu de travaux ont cherché à évaluer l'efficacité de l'utilisation des CNLs. Cette efficacité concerne le fait que le destinataire comprenne ce que le locuteur a voulu dire (c'est-à-dire souvent réponde par une action attendue) et ce, en général, dans le temps le plus restreint possible. Cependant, on peut trouver quelques travaux sur l'utilisation de la norme AECMA, et plus récemment, des travaux réalisés sur le Controlled Language for Crisis Management (CLCM) (Temnikova, 2012) qui ont mis en œuvre des tests de compréhension pour mesurer cette efficacité. Ils ont révélé que les CNLs n'étaient pas toujours aussi efficaces qu'on pourrait le souhaiter. (Jahchan *et al*, à paraître), voir ci-dessous.

4 Propositions

Pour pallier les difficultés rencontrées par la mise en œuvre des langues contrôlées, il semble possible d'améliorer, voire de construire, de nouvelles langues contrôlées visant à accroître la compréhension. Deux voies possibles sont présentées ici.

4.1 Propositions appuyées sur les apports du TAL

Dans le but de les voir utilisées dans un grand nombre de situations de communication, les langues contrôlées ont été construites avec l'idée de ne pas prendre en compte des situations de communication particulières et donc de proposer des règles générales susceptibles de convenir à la plupart des situations. Or, avec le recul de plusieurs dizaines d'années de mise en œuvre et avec le développement des méthodes en TAL et en psycholinguistique, il semble maintenant possible de développer des CNLs qui intègrent le contexte d'utilisation et adaptent les préconisations à ces situations.

4.1.1 Construire des langues contrôlées à partir d'attestations réelles

Il se trouve que, dans de nombreux cas, les documents techniques sont rédigés sans le recours à des CNLs. Cela ne signifie pas que ce mode de rédaction « spontané » n'obéisse pas à des règles mais il ne s'agit pas de règles prescrites mais de règles inconscientes, propres à un genre textuel (ou à un sous-langage). On le sait en effet, des régularités linguistiques dans les échanges se mettent en place dans des situations de communications récurrentes. Ces deux aspects (réurrence de la situation de communication et régularités linguistiques, le plus souvent inconscientes) sont les éléments qui caractérisent un genre textuel, dont on pense aussi qu'il favorise une économie cognitive : lorsqu'on s'attend à une forme particulière, l'effort cognitif à fournir est moindre qu'avec une forme inédite.

Dans une vision ascendante, appuyée sur les méthodes de la linguistique de corpus, on peut envisager d'utiliser les textes produits dans une situation et donc obéissant à des caractéristiques d'un genre pour identifier les structures récurrentes (les phraséologies). On peut alors faire l'hypothèse que ces régularités peuvent constituer le matériau de base d'une langue contrôlée adaptée au genre en question. C'est l'étude que nous sommes en train de mener avec le CNES sur un corpus d'exigences (spécifications de besoins), constitué des exigences de deux projets spatiaux et d'un projet informatique. Deux types d'approches peuvent être mises en œuvre pour repérer les régularités syntactico-sémantiques propres à un corpus. A partir de ce corpus étiqueté, des méthodes de statistiques de repérage des motifs récurrents peuvent être mises en œuvre (ce type de méthode est utilisée dans l'outil SDMC par exemple (Quiniou et al., 2012)) ou bien l'approche statistique peut être mise en œuvre par comparaison du corpus à l'étude avec un corpus dans la langue de référence (approche présentée dans (Temnikova et Bretonnel Cohen, 2013)). Nous sommes en train de travailler au dépouillement des résultats fournis par SDMC (Sequential Data Mining under Constraints) sur notre corpus d'exigences en les filtrant de deux façons (Warnier et Condamines, 2015) :

- En ne conservant que les structures comportant des termes du genre (et non du domaine), c'est-à-dire en excluant des résultats obtenus à partir d'un extracteur de termes-candidats les termes du domaine (fournis par le CNES) afin de ne conserver que les candidats-termes susceptibles de relever du genre « exigences ». Par exemple, nous avons supprimé de la liste des termes du domaine, fournie par le CNES comme *égalesisation radiométrique* ou *Banc Opérationnel Système Satellite*, mais conservé des termes-candidats comme *durée de la manoeuvre*, *paramètre de programmation*, *rapport d'expertise*, *confidentialité*....
- En retenant les structures contenant un des verbes les plus fréquents (*devoir*, *permettre*, *fournir*, *définir*, *contenir*...) et un candidat-terme que nous avons conservé comme relevant du genre « exigences ».

Nous avons par exemple pu repérer les structures suivantes :

- [Det N1 + permettre au présent + à det N2 + de (V + nominalisation déverbiale)]

Le Segment Sol de Programmation PHR permet à un utilisateur d'afficher les données opérationnelles.

- [Det N + permettre au futur + de (V + nominalisation déverbiale)]

le DUPC permettra de modifier localement les paramètres du calcul.

Dans ces deux exemples, une des questions que pose la caractérisation du genre « spécification de besoins » est celle de l'expression de l'injonction, qui caractérise la situation de communication, comme l'indique le nom « exigences ». Selon les phrases, elle s'exprime soit à l'aide d'un modal (*devoir, pouvoir*), soit à l'aide du futur, soit même à l'aide du seul présent, comme c'est le cas dans le premier exemples ci-dessus, le présent étant utilisé alors pour décrire le système tel qu'on attend qu'il soit. Notons que l'impératif (a priori le mode le plus adapté à l'injonction) n'est jamais utilisé. La meilleure forme adaptée à l'expression de l'injonction dans cette situation précise fait actuellement l'objet d'une enquête d'acceptabilité au CNES.

4.2 Partir des informations à transmettre et proposer différentes formes langagières

La particularité des langues contrôlées est qu'elles sont censées se mettre en œuvre dans des situations contraintes, surtout du point de vue des informations à transmettre, en lien avec les actions attendues par la transmission langagière.

Pour certaines situations, dans lesquelles une catégorisation des informations à transmettre peut être circonscrite a priori, on pourrait envisager d'élaborer toutes les formes langagières possibles pour transmettre ces types d'information afin de les évaluer et de voir laquelle/lesquelles est/sont la/les mieux adaptée(s). Il n'est pas impossible d'imaginer que, dans une perspective assez similaire à celle de l'extraction d'information, un programme de TAL, partant d'une forme supposée canonique (car proche de la structure cognitivo-sémantique, le *frame*), par exemple, dans le domaine de la circulation routière, [sujet prédicat lieu] : *un obstacle est présent au kilomètre 45* puisse proposer l'ensemble des variantes linguistiques (lexicales et syntaxiques) possibles, dont l'efficacité pourrait être testée dans un second temps (par exemple : *obstacle kilomètre 45 ; kilomètre 45 : obstacle ; Attention : obstacle présent kilomètre 45etc...*) .

Cette façon de procéder aurait un second mérite : faciliter le traitement automatique des textes rédigés avec cette langue contrôlée, voire, faciliter leur formalisation. Cette dimension formelle est par exemple très présente dans les travaux qui sont menés sur la rédaction des exigences (ingénierie des exigences). La récurrence des informations à véhiculer est aussi parfois utilisée dans la définition de gabarits (boilerplates) à remplir ; cette fonctionnalité a été introduite par exemple dans l'outil Doors développé par IBM, censé assister la rédaction des exigences. Mais, dans le cas des exigences, seule une petite partie des informations à transmettre peut s'exprimer sous cette forme alors que, dans d'autres cas (annonces sur les autoroutes, alarmes dans les cockpits d'avion...), la quasi-totalité des informations pourraient être traitée de cette façon.

Les méthodes de TAL à mettre en œuvre doivent ainsi être adaptées en fonction de la nature des informations à transmettre (récurrentes ou non, brèves ou non) et de la situation de communication.

4.3 Propositions appuyées sur la mise en œuvre des méthodes de la psycholinguistique

L'évaluation de la compréhensibilité passe nécessairement par la mise en place de tests psycholinguistiques. Dans les faits, il s'agit d'évaluer différentes formes linguistiques possibles afin de voir laquelle est la mieux adaptée pour transmettre l'information censée être transmise. Les tests - doivent être faits dans des situations les plus proches possibles des situations écologiques. Deux études sont en cours sur ces aspects.

4.3.1 Exemple 1 : interroger l'alternative nominalisation/verbe

Plusieurs études ont montré que les langues spécialisées utilisent plus souvent les nominalisations déverbales que la langue générale (Condamines et Picton, 2014). On a pu faire l'hypothèse que cette sur-utilisation était liée à une volonté de donner une dimension objective au discours, en utilisant une forme inattendue mais plus « autonome » que les verbes pour désigner des actions ou des états. En effet, la forme nominale est souvent considérée comme la forme langagière la plus achevée et en quelque sorte la plus simple, en particulier parce qu'elle peut fonctionner hors discours, sur le mode de l'étiquetage.

On retrouve dans certaines langues contrôlées, probablement inspirées par les langues spécialisées, des préconisations concernant l'utilisation de la forme nominale plutôt que la forme verbale. Ainsi, sur le site [muegge.cc](http://www.muegge.cc/index.htm) (<http://www.muegge.cc/index.htm>) qui propose des règles pour faciliter la traduction, on trouve cette recommandation :

N'écrivez pas :

Comment installer votre scanner

Mais

Installation du scanner :

Or, on sait aussi que la nominalisation a souvent une couverture sémantique plus importante que celle du verbe (*construire* renvoie uniquement à une action alors que *construction* renvoie aussi au résultat de cette action, sous forme tangible). Dans le même ordre d'idée, une spécificité des nominalisations déverbales vient de ce que, contrairement à ce qui se passe pour le verbe, la présence de tous les arguments n'est pas indispensable pour qu'elle fonctionne, ce qui peut introduire une part d'incomplétude, alors que c'est beaucoup moins facile pour le verbe.

On peut dès lors s'interroger sur les préconisations les plus adaptées concernant l'expression langagière d'une action : forme verbale ou forme nominale ? Une étude est en cours d'élaboration sur ce sujet.

4.3.2 Exemple 2 : interroger la « simplification » dans les langues contrôlées

Très souvent, les langues contrôlées sont considérées comme des simplifications de la langue générale (on parle ainsi de *Simplified English*). Dans les recommandations d'écriture, la concision (à entendre comme le raccourcissement de la forme linguistique) est généralement associée à l'amélioration de la compréhension :

Concision is an aid to *Comprehensibility*, and is therefore subsumed by it (Incese).

If you have a choice, use the shortest and simplest name (ASD).

Concrètement, il s'agit souvent de supprimer des éléments syntaxiques, considérés comme superfétatoires et consommateurs de temps de lecture. On trouve ainsi fréquemment des syntagmes nominaux sans préposition ni déterminant ou encore des prédicats sans la présence de tous les arguments attendus. Notons que ces structures correspondent à la notion de fonctionnement « déviant » souvent évoqué pour décrire les sous-langages (Grishman et Kittredge, 1986). Dans les deux cas, la régularité de la situation de communication est supposée pallier la suppression d'éléments sémantico-syntaxiques : les interlocuteurs sont censés partager suffisamment de connaissances sur le domaine et sur la situation de communication pour comprendre des structures elliptiques, qui sont pourtant potentiellement ambiguës (par exemple, *observation satellite*, fréquent dans les corpus du CNES, peut être équivalent soit à *observation du satellite* (avec *satellite*, objet ou locatif), soit à *observation par satellite*. Ces hypothèses (et d'autres) ne sont pas sans risque et ont été très peu évaluées. Toutefois, Simpson (1976) a montré, dans une étude déjà ancienne réalisée à la NASA, que des phrases sans ellipse étaient comprises plus rapidement (à l'oral) que les phrases avec ellipse (par exemple *the fuel pressure is low* vs *fuel low*). La thèse en cours de N. Jahchan (financement CIFRE Airbus) porte précisément sur l'évaluation du lien entre concision de la forme et efficacité de la compréhension (Jahchan *et al.*, à paraître).

4.3.3 Quels tests ?

La mise en place de tests psycholinguistiques pour vérifier la compréhensibilité d'une structure linguistique fait appel à des compétences à la fois linguistiques (pour définir les formes langagières à évaluer) et psycholinguistiques, pour le choix et la mise en place des tests eux-mêmes.

Deux types de tests sont envisageables : les tests d'acceptabilité directe (préférence spontanée pour telle ou telle structure, en situation générale en comparaison avec une situation écologique) et les tests basés sur la prise en compte d'indices physiologiques (mouvements de la pupille pour l'oculométrie ou activation cérébrale pour l'électroencéphalogramme (EEG)).

Pour ce qui est de l'oculométrie, il semble que, si cette méthode paraît assez fréquemment utilisée dans des études concernant la traduction (Göpferich et al., 2008) (voir aussi, International Conference on Eyetracking and Applied Linguistics, <https://iceal.uni-mainz.de/programme>), peu de travaux ont fait appel à des tests de ce type pour valider des fonctionnements linguistiques. Notons toutefois l'utilisation de l'oculométrie pour valider des phénomènes de polysémie (Hollard, 2010) ou encore, et cela nous intéresse au premier chef, la mise en œuvre de l'oculométrie par O'Brien pour évaluer les langues contrôlées, mais plutôt à nouveau avec une orientation vers les aspects traductionnels (O'Brien, 2010).

Pour ce qui concerne l'EEG, son utilisation semble exceptionnelle dans l'analyse des corpus spécialisés. Signalons toutefois l'utilisation des méthodes d'imagerie cérébrale par Faber et son équipe (Faber et al., 2014). Ils ont mis en évidence que, lors d'une tâche d'association de mots avec des termes plus génériques, les experts ne mettent pas en œuvre la même zone du cerveau que les non-experts. Les experts activent, entre autres, l'amygdale, qui est la zone des émotions. Ce résultat pourrait être exploré pour voir comment cet aspect émotionnel interfère dans la connaissance et la communication expertes.

5 Conclusion

Un peu délaissée ces dernières années aussi bien en TAL qu'en linguistique, la question des langues contrôlées pourrait trouver un regain d'intérêt par la mise en œuvre parallèle et/ou complémentaire des méthodes en TAL (pour élaborer des recommandations linguistiques plus proches des usages réels et pour fournir une aide à la rédaction) et des méthodes en psycholinguistique (pour repérer quelles structures linguistiques sont les plus efficaces et donc à privilégier).

Je remercie Maxime Warnier pour sa relecture attentive de cet article.

Références

- ASD-STE 100, Simplified Technical English, 2007.
- BRUCE B., RUBIN A., STARR K. S. (1981). Why Readability Formulas Fail. *IEEE Transactions on Professional Communication*, PC-24, 50-52.
- CONDAMINES A. (2010). Variations in Terminology. Application to the Management of Risks Related to Language Use in the Workplace. *Terminology*, 16(1), 30-50.
- CONDAMINES A. (2008). Peut-on prévenir le risque langagier dans la communication écrite? *Langage et société*, 125, 77-97.
- CONDAMINES A., PICTON A. (2014). Étude du fonctionnement des nominalisations déverbiales dans un contexte de déspecialisation. *Congrès Mondial de linguistique Française (CMLF)*. Berlin, 19-23 juillet 2014.
- CONDAMINES A., WARNIER M. (2014). Linguistic Analysis of Requirements of a Space Project and their Conformity with the Recommendations Proposed by a Controlled Natural Language. *Fourth Workshop on Controlled Natural Language (CNL 2014)*, 20-22 août 2014, Galway
- DUBAY W., H. (2004). *The Principles of Readability*. Costa Mesa, CA: Impact Information, 2004 <http://www.impact-information.com/impactinfo/readability02.pdf>
- FABER BENÍTEZ P., VERDEJO J., LEÓN P., REIMERINK A., GUZMÁN G. (2014). Neural Substrates of Specialized Knowledge Representation: An fMRI study. *Revue française de linguistique appliquée*, 19(1), 15-32.
- GÖPFERICH, S., JAKOBSEN, A. L., MEES I. M., HRSG. (2008): *Looking at Eyes: Eye-Tracking Studies of Reading and Translation Processing*. Copenhagen Studies in Language 36. Copenhagen: Samfundslitteratur.
- GRISHMAN R., KITTREDGE R. (1986). *Analyzing Language in Restricted Domain*. Hillsdale, London :Lawrence Erlbaum Associates, Publishers
- HOLLARD S. (2010). Interprétation de textes polysémiques : une étude expérimentale appuyée sur l'oculométrie, *Glossa* n° 109, 16- 41.
http://glossa.fr/pdfs/109_20101022114839.pdf
- INCOSE (INTERNATIONAL COUNCIL ON SYSTEMS ENGINEERING. (2001). Guide for Writing Requirement. Version 1.
- JAHCHAN, N., CONDAMINES A., CANESSON E, à paraître. *To What Extent Does Text Simplification Entail a More Optimized Comprehension in Human-oriented CNLs?* Soumis à CNL 2016.
- KLARE G. R. 1976. A second look at the validity of readability formulas. *Journal of Reading Behavior*, 129-152.
<http://jlr.sagepub.com/content/8/2/129.full.pdf>
- KUHN T. (2014). A Survey and Classification of Controlled Natural Languages. *Computational Linguistics*, 40(1), pp.121-170.
- LOPEZ S. (2013). *Norme(s) et usage(s) langagiers : le cas des communications pilote-contrôleur en anglais*. Thèse de l'Université Toulouse 2.

- LUX V. (1998). *Elaboration d'un Français Rationalisé Etendu Modulaire (FREM) pour les Manuels de Maintenance d'Aéronef*. Thèse de Doctorat en Linguistique de l'Université Paris 7.
- O'BRIEN S. (2010). Controlled Language and Readability. In Gregory M. Shreve and Erik Angelone (eds) : *Translation and Cognition*.143-165.
- QUINIOU S., CELLIER P., CHARNOIS T., LEGALLOIS D. (2012). What about Sequential Data Mining Techniques to Identify Linguistic Patterns for Stylistics? In A. Gelbukh (Ed.), *Computational Linguistics and Intelligent Text Processing* (166–177). Springer Berlin Heidelberg.
- SIMPSON C.A. (1976). Effects of Linguistic Redundancy on Pilot's Comprehension on Synthesized Speech, *Twelfth Annual Conference on Manual Control*, Nasa Technical Memorandum. 294-308.
- TEMNIKOVA I. (2012). *Text Complexity and Text Simplification in the Crisis Management Domain*. Ph.D. thesis, University of Wolverhampton.
- TEMNIKOVA I., BRETONNEL COHEN K. (2013). *Recognizing Sublanguages in Scientific Journal Articles through Closure Properties*. In the *12th Workshop on Biomedical Natural Language Processing (BioNLP 2013)*, held in conjunction with ACL 2013, August 8th-9th, 2013, Sofia, Bulgaria.
- WARNIER M., CONDAMINES A. (2015). A Methodology for Identifying Terms and Patterns Specific to Requirements as a Textual Genre Using Automated Tools. Proceedings of *International Conference "Terminology and Artificial Intelligence "*. 4-6 November, Granada, Spain, 183-190.