

# Évaluation des langues contrôlées pour limiter le risque langagier dans la rédaction des exigences : exemple sur un projet du CNES

Maxime Warnier (CLLE-ERSS)  
maxime.warnier@univ-tlse2.fr



# Contexte de l'étude

- Demande du service Assurance Qualité du CNES : comment améliorer la rédaction des **exigences** ?  
*“conditions or capabilities that must be met or possessed by a system or system component to satisfy a contract, standard, specification, or other formally imposed document”*  
[IEEE 1990]
- Nombreuses exigences, organisées en spécifications, écrites en langue naturelle (français/anglais)
  - Logiciels de gestion des exigences : DOORS, Reqtify, etc.
  - Mais aucune aide à la rédaction

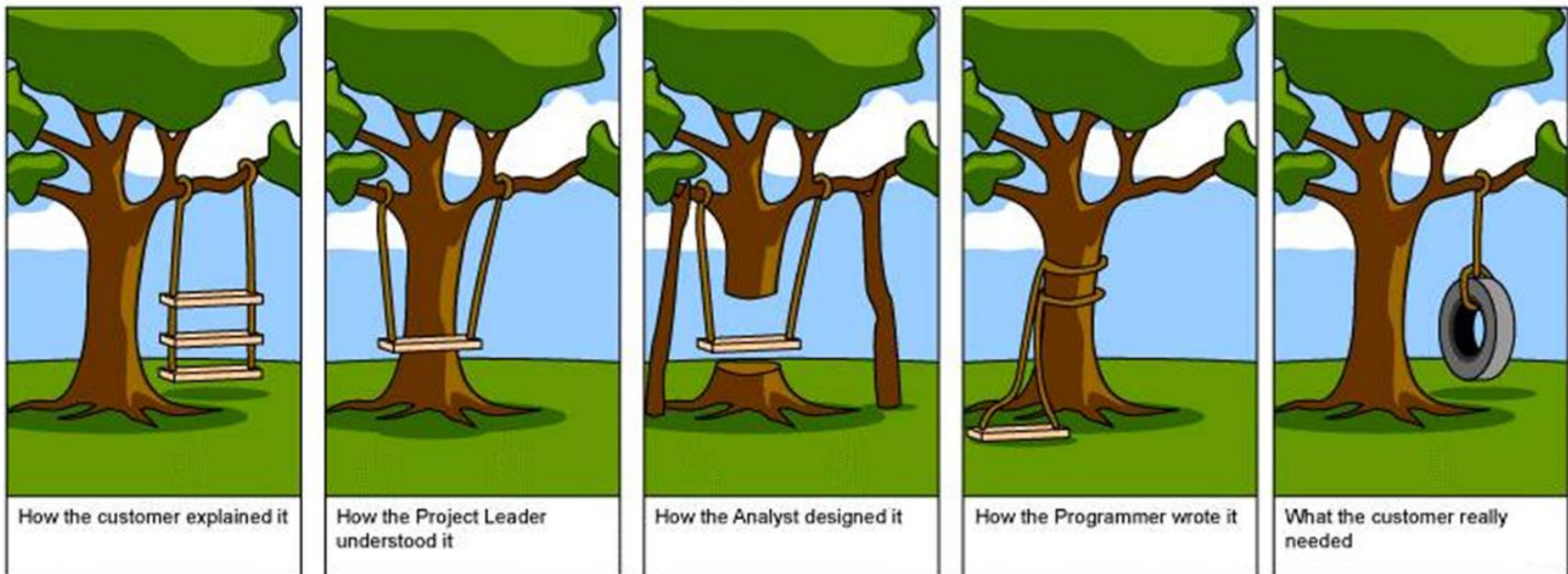
# Dimension langagière du risque

- La langue n'est pas un moyen de transmission de l'information parfait (à l'oral comme à l'écrit)
  - Ambiguïté [= plusieurs interprétations possibles]
    - Lexicale (« où sont les jumelles ? »)
    - Syntaxique (« elle a trouvé son fils prodigue »)
    - Référentielle (« Pierre a ramené Paul chez lui »)
  - Imprécision
    - « la tâche devra être effectuée en quelques secondes »
  - Incomplétude
- « *Si l'on prend la forme linguistique comme un wagon que l'on charge d'un certain sens dont le destinataire prendra possession, on se trompe* » (Gardin, 1995)

# Dimension langagière du risque

- Les exigences sont critiques (obligation contractuelle)
  - Surtout pour les projets de grande envergure (plus d'exigences, plus d'intervenants, durée plus longue, etc.)
- Conséquences possibles d'une mésinterprétation :
  - Retards
  - Augmentation des coûts
  - Litiges
  - Accidents

# Dimension langagière du risque



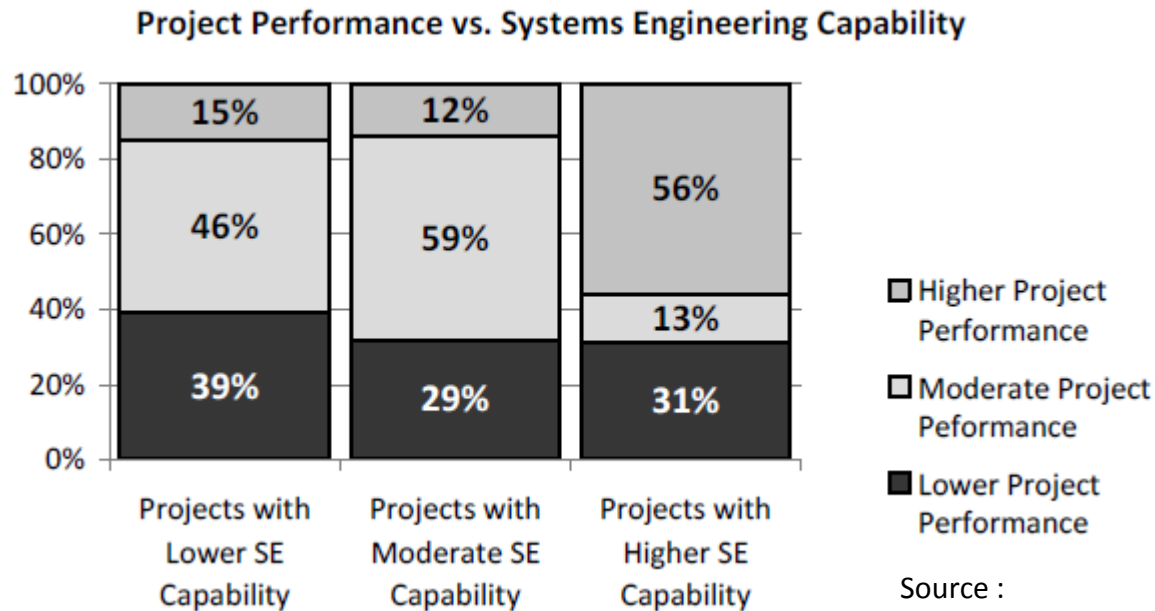
“Tree swing cartoon”

# Dimension langagière du risque

- Le langage naturel reste pourtant
  - le moyen le plus expressif pour spécifier ses besoins
  - facile à utiliser par toutes les parties prenantes
  - inévitable (pour certaines étapes de la spécification)
- L'objectif est donc un moyen d'exprimer les besoins...
  - aussi naturel et simple que possible
  - qui limite autant que possible les risques langagiers

# Dimension langagière du risque

- L'ingénierie des exigences est un facteur de réussite...



- ... la bonne rédaction des exigences aussi !
  - Travail « en amont »

# Les langues contrôlées

- Les langues contrôlées (LC) pour éliminer les risques langagiers
  - Basic English (1930)
  - Caterpillar Fundamental English (1971)
- Selon Kuhn (2014), une LC :
  1. est basée sur une langue naturelle
  2. **pose des restrictions sur le vocabulaire, la syntaxe et la sémantique**
  3. reste suffisamment proche de la langue source
  4. est un langage consciemment défini



# Les langues contrôlées

- Il existe déjà des LC pour la rédaction technique, mais, selon nous :
  - elles sont parfois **inapplicables** (trop contraignantes)
  - elles sont parfois **inefficaces** (pas assez strictes)
  - elles **manquent de rigueur** (incohérentes)
  - > Elles ne sont pas adaptées aux pratiques existantes
- LC établies par des experts du domaine (ingénieurs), pas des experts de la langue (linguistes)
  - Connaissances différentes mais précieuses

# Les langues contrôlées

- Pas de LC imposée aux ingénieurs du CNES
- Nous supposons quand même une possible influence des LC existantes, due à l'expérience des rédacteurs
- LC analysées (en anglais) :
  - *ASD Simplified Technical English* (2007)
    - Standard de l'aérospatial (documentation de maintenance)
  - *Guide for Writing Requirements* d'INCOSE (2011)
    - Orienté rédaction des exigences, multidisciplinaire
    - S'inspire largement de normes et guides de rédaction technique antérieurs (récurrence des règles)

# Règles : ASD STE

- **“Keep procedural sentences as short as possible (20 words maximum)”**

*The maximum length of a procedural sentence is 20 words. (Refer to Section 8 for the rules about word count.)*

# Règles : INCOSE

- **“Avoid combinators”**  
*Combinators are words that join clauses together, such as 'and', 'or', 'then', 'unless'. Their presence in a requirement usually indicates that multiple requirements should be written.*
- **“Repeat nouns in full instead of using pronouns to refer to nouns in other requirement statements”**  
*Pronouns are words such as 'it', 'this', 'that', 'he', 'she', 'they', 'them'. When writing stories, they (sic.) are a useful device for avoiding the repetition of words; but when writing requirements, pronouns should be avoided, and the proper nouns repeated where necessary.*

# Critique des règles

- Ces règles
  - sont **imprécises** (pas de définition, liste non exhaustive)
  - sont très **restrictives**
  - sont mal justifiées
  - manquent de cohérence

# Méthodologie

- Linguistique de corpus :
  - Corpus en français, spécifications Pléiades
    - 1 142 exigences
    - 53 000 mots
    - > **données attestées**
  - Outils d'analyse
- « Diagnostic » des exigences du CNES :
  - Les règles des LC sont-elles (indirectement) respectées ?
  - Objectif : se baser sur l'existant pour proposer des règles plus adéquates

# Analyse quantitative : longueur des phrases

	Nombre de phrases	Nombre de phrases longues	Longueur moyenne des phrases
<b>Exigences</b>	4859	350 <b>(7,20 %)</b>	11
<b>Manuel</b>	3456	591 <b>(17,10 %)</b>	15
<b>Journal</b>	2201	839 <b>(38,10 %)</b>	24

Phrase la plus longue du corpus :

*« Si la différence (en valeur absolue) entre les dates de fin de lecture de deux fichiers, lus sur tranche de COME M - canal TMI i et sur tranche de COME N - canal TMI j, est inférieure à OPS\_DELAI\_INTER\_FIN\_LEC secondes, alors il est interdit d'enchaîner (lecture enchaînée) par la lecture de la tranche de COME N sur le canal i et de la tranche de COME M sur le canal j. »*

# Analyse quantitative : pronoms et conjonctions

	<b>Pronoms</b>	<b>Conjonctions coordination</b>	<b>Conjonctions subordination</b>	<b>Total conjonctions</b>
<b>Exigences</b>	986 <b>(1,86 %)</b>	882 <b>(1,66 %)</b>	365 <b>(0,69 %)</b>	1247 <b>(2,35 %)</b>
<b>Manuel</b>	1554 <b>(2,93 %)</b>	1455 <b>(2,75 %)</b>	442 <b>(0,83 %)</b>	1897 <b>(3,58 %)</b>
<b>Journal</b>	2710 <b>(5,11 %)</b>	1274 <b>(2,40 %)</b>	579 <b>(1,09 %)</b>	1853 <b>(3,50 %)</b>



# Revue d'exemples : conjonctions

- **Indispensables**

- *Le générateur de TCH vérifiera **que** la valeur du champ PHASE est comprise entre 0 **et**  $FREQ\_DIV - 1$ .*

- **Justifiées**

- *pour  $n=2$  la loi de la taille est respectée de fait **mais** le test "FIFO vide" reste nécessaire*

- *Les champs  $SM\_ID$  **et**  $FM\_ID$  seront extraits à partir de la BDS*

- **Indésirables**

- *Les demandes sont saisies sur le FOS **et** le logiciel ARPE gère les conflits entre les demandes Spot, Hélios et Pléiades*

# Revue d'exemples : conjonctions

- Certaines conjonctions sont obligatoires
- Certaines conjonctions sont utiles
  - elles évitent des répétitions
  - elles rendent l'instruction plus claire
- Certaines conjonctions sont inutiles
  - La conjonction de coordination « et » ne devrait pas être utilisée pour lier deux propositions indépendantes sans élément commun (sujet, verbe, complément)

# Revue d'exemples : pronoms

- **Indispensables**

- *// ne sera pas utile de vérifier ce paquet "vide"*

- **Justifiés**

- *Le paquet ne sera généré que s'il est activé par le LVC.*
  - *Cette TC a également pour effet d'effacer le précédent contenu du buffer FDTM si **celui-ci** était non vide.*

- **Indésirables**

- *// calculera aussi, a une fréquence paramétrable (ordre de grandeur 1 mois), la moyenne de mise en œuvre*

# Revue d'exemples : pronoms

- Certains pronoms sont obligatoires
- Certains pronoms sont utiles
  - ils évitent des répétitions
- Certains pronoms sont problématiques
  - Un pronom personnel peut être utilisé s'il a un (et un seul) antécédent possible dans l'exigence

# Conclusion

- **Objectif** : proposer des règles de rédaction adaptées aux pratiques réelles du CNES qui limitent au maximum les risques langagiers
- **Analyse**
  - de règles existantes
  - de données attestées
    - recherche de régularités linguistiques

# Références

- ASD: Simplified Technical English. International specification for the preparation of maintenance documentation in a controlled language (Issue 4) (2007)
- Bünzli, A., Höfler, S.: Controlling Ambiguities in Legislative Language. Controlled Natural Language: Second International Workshop. 21–42 (2012)
- Condamines, A.: Variations in Terminology. Application to the Management of Risks Related to Language Use in the Workplace. Terminology. 16, 30–50 (2010)
- Condamines, A., Warnier, M.: Linguistic Analysis of Requirements of a Space Project and Their Conformity with the Recommendations Proposed by a Controlled Natural Language. Controlled Natural Language: 4th International Workshop. 33-43 (2014)
- Fuchs, N., Schwitter, R.: Controlled English for Requirements Specifications. IEEE Computer Special Issue on Interactive Natural Language Processing (1996)
- International Council on Systems Engineering: Guide for Writing Requirements (Version 1) (2011)
- Kuhn, T.: A Survey and Classification of Controlled Natural Languages. Computational Linguistics. 40, 121–170 (2014)
- O'Brien, S.: Controlling Controlled English. An Analysis of Several Controlled Language Rule Sets. Proceedings of EAMT-CLAW. pp. 105–114 (2003)