



# Ethique conséquentialiste et traitement automatique des langues : une typologie de facteurs de risques adaptée aux technologies langagières

Anais Lefevre, Jean-Yves Antoine, Willy Allegre

## ► To cite this version:

Anais Lefevre, Jean-Yves Antoine, Willy Allegre. Ethique conséquentialiste et traitement automatique des langues : une typologie de facteurs de risques adaptée aux technologies langagières. Atelier Ethique et TRaitement Automatique des Langues (ETeRNAL'2015), conférence TALN'2015, Jun 2015, Caen, France. Atelier ETeRNAL'2015, conférence TALN'2015, pp.53-66, 2015, Actes de la 1e Ethique et TRaitement Automatique des Langues (ETeRNAL'2015), Caen (France). <[http://www.atala.org/taln\\_archives/TALN/TALN-2015/index.html](http://www.atala.org/taln_archives/TALN/TALN-2015/index.html)>. <hal-01170630>

**HAL Id: hal-01170630**

**<https://hal.archives-ouvertes.fr/hal-01170630>**

Submitted on 2 Jul 2015

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Ethique conséquentialiste et traitement automatique des langues : une typologie de facteurs de risques adaptée aux technologies langagières

Anais Lefevre<sup>1</sup>, Jean-Yves Antoine<sup>1</sup>, Willy Allegre<sup>2</sup>

(1) Laboratoire d'Informatique de l'Université de Tours, 3 place Jean Jaurès, 41000 Blois

(2) Laboratoire d'Electronique, CMRFF de Kerpape, BP 78, 56275 Ploemeur

[anais.lefeuvre@univ-tours.fr](mailto:anais.lefeuvre@univ-tours.fr), [Jean-Yves.Antoine@univ-tours.fr](mailto:Jean-Yves.Antoine@univ-tours.fr), [wallegre@kerpape.mutualite56.fr](mailto:wallegre@kerpape.mutualite56.fr)

**Résumé.** Cet article présente une typologie de facteurs de risques concernant les technologies numériques et plus particulièrement les technologies langagières. Son objectif est d'offrir une grille d'analyse pour une évaluation critique des recherches et applications du TALN dans une démarche éthique conséquentialiste.

### Abstract.

**Consequentialist ethics and NLP: a typology of risk factors suitable to language technologies.**

This paper details a typology of risk factors that should concern digital technologies and more specifically NLP. It aims at providing an evaluation grid for an ethical assessment of researches and applications.

**Mots-clés :** Ethique, risque, facteur de risques, vulnérabilité, criticité, risque individuel, risque sociétal, TAL.

**Keywords:** Ethics, risk analysis, risk factor, vulnerability, criticality, individual risk, societal risk, NLP.

## 1 Pour une éthique conséquentialiste des technologies langagières

L'informatisation de la société et sa mise en réseau ont le plus souvent suscité des discours enthousiastes sur l'émergence d'une société de la connaissance, le fonctionnement non hiérarchisé d'internet, vision moderne d'un village global par excellence conduisant à une intelligence et une prise de décision collectives. Ce consensus sociétal est désormais battu en brèche par de multiples travaux en sciences sociales (Jarrige 2014). La critique la plus visible concerne le respect de la vie privée sur les réseaux sociaux, le droit à l'oubli numérique, et plus généralement la question de l'émergence d'une société du contrôle permise par ces techniques. Ces questions surviennent à un moment où le TALN a acquis une maturité suffisante pour permettre par exemple une fouille de données intelligente dans de grands flux d'informations. Jusqu'à récemment, les technologies langagières pouvaient sembler moins concernées par des questionnements éthiques que les biotechnologies ou les nanotechnologies. L'association entre analyse automatique intelligente de la langue, masses de données (*big data*) et informatique ubiquitaire requiert désormais que notre communauté scientifique s'interroge sur son objet de recherche et sur ses applications.

Les réflexions éthiques en TALN se sont jusqu'à présent concentrées surtout sur l'anonymisation des données personnelles dans les corpus. En France, cette question est traitée de longue date par la réglementation (loi *Informatique et Libertés* de 1978), hétéro-régulation normative relevant de décisions étatiques. A l'opposé, l'éthique relève d'une auto-régulation non normative, fondée des choix et jugements de valeurs collectifs. Des comités d'éthique ont progressivement été mis en place. Ils ont avant tout un rôle de recommandation et de conseil, même si l'éthique influence désormais le droit par l'intermédiaire de la jurisprudence. La CERNA (Commission de réflexion sur l'éthique de la Recherche en sciences et technologies du Numérique d'ALLISTENE - ALLiance des Sciences et TEchnologies du NumériquE) joue précisément ce rôle dans le cadre des sciences et technologie du numérique, en se positionnant à l'interface entre les recherches du domaine et de leurs applications industrielles. Elle a émis un premier jeu de recommandation sur l'éthique de la recherche en robotique (CERNA 2014). Le TALN est toutefois relativement absent de ses activités. Dans cet article, nous aimerions précisément apporter quelques éléments méthodologiques pour la mise en place d'une réflexion éthique en TALN.

Une réflexion éthique peut se baser sur deux principaux courants de pensée contemporains. D'une part, l'éthique conséquentialiste suit une approche téléologique qui consiste à se focaliser non pas sur des principes mais sur les conséquences de nos actions. Ce point de vue guidait l'utilitarisme de Jeremy Bentham ou John Stuart Mills, pour qui toute action est justifiée par ses effets positifs sur le plus grand nombre. L'éco-éthique contemporaine cherche au contraire à réduire les effets néfastes de nos actions. Les principes de précaution et de responsabilité, théorisés entre autre par Hans Jonas (1990), relèvent précisément de cette logique.

D'autre part l'éthique déontologique promeut le respect de principes moraux pour régir nos actions, la réflexion éthique devant porter sur l'établissement de ces principes. Par exemple, Rawls (1987) et sa théorie de la justice proposent une logique contractualiste entre personnes libres et rationnelles : celle-ci établit des principes de fonctionnement de la société, qui sont jugés éthiques si leur processus amont d'élaboration a été équitable.

La réflexion qui est présentée dans cet article relève d'une étude conséquentialiste des effets négatifs des recherches et applications en TALN. Cette démarche téléologique nous semble en effet la plus adaptée pour engager une prise de conscience dans un champ de connaissances n'ayant amorcé que de manière embryonnaire une étude réflexive sur ses pratiques. Une méthodologie d'analyse en termes de risques, positifs ou négatifs cette fois, existe dans des domaines tels que la finance par exemple. A l'instar de l'analyse du risque industriel, nous privilégierons ici une attention aux impacts jugés négatifs dans un objectif d'amélioration des recherches menées en TALN, dont on ne doit bien sûr pas négliger les promesses. En pratique, nous proposons de conduire une analyse de risque telle qu'envisagée dans les autres domaines technologiques, sur les technologies langagières que notre communauté concourt à développer. Cette analyse évaluative nous semble d'autant plus nécessaire que nos sociétés postmodernes ne sont désormais plus régies uniquement par la question du partage des richesses mais également par celle du risque technologique (Beck 2001).

La démarche que nous proposons se place dans une perspective d'amélioration des pratiques de recherche par la mise en place d'une réflexion éthique. Comme toute analyse de risque, elle nécessite l'élaboration de protocoles d'évaluation des effets induits par les technologies que nous produisons. Par la connaissance experte de son domaine d'étude, le chercheur en TALN est idéalement placé pour jouer un rôle de lanceur d'alerte sur des problématiques dont il peut être parfois le seul à percevoir le risque. Cela ne remet aucunement en cause le fait qu'à terme, les protocoles d'expérimentation appropriés devront être conçus en collaboration avec les spécialistes de chaque type d'impact.

Dans un premier temps, nous allons présenter le cadre classique de l'analyse de risque et voir dans quelle mesure il pourrait s'appliquer au TALN. Nous nous focaliserons ensuite sur la notion de facteur de risques, que nous mettrons en regard d'une première classification de types de risques qui a été inspirée par nos propres travaux sur l'utilisation des technologies langagières pour l'aide aux personnes handicapées. Cette classification sera ensuite affinée sous la forme d'une typologie à 5 niveaux qui forme une grille d'analyse pour l'étude du risque en TAL, et potentiellement toute technologie numérique. Cette typologie sera illustrée au regard de diverses applications du TAL.

## 2 Qu'est-ce que le risque ?

Historiquement, les premières tentatives de modélisation du risque remontent aux XVII<sup>ème</sup> et XVIII<sup>ème</sup> et sont le fait de mathématiciens (Huygens, Bernoulli, Pascal) qui s'interrogeaient sur l'incertitude liée à la notion de risque. La révolution industrielle va profondément renouveler cette notion : alors que le risque était jusque-là lié dans les consciences aux calamités naturelles par essence inévitables, celles-ci vont laisser la place aux catastrophes et crises environnementales liées au processus industriel. Dans nos sociétés modernes, le risque est donc le fait principal des activités humaines où d'où la définition par certains d'une nouvelle ère dénommée *anthropocène* (Crutzen et al. 2007).

Il en résulte l'émergence d'une société réflexive du risque (Beck 2001), où le développement du risque n'est acceptable par la population que dans la mesure où l'on cherche à l'évaluer (rapport bénéfice/risque) et le gérer. C'est dans ce cadre que l'on voit apparaître un ensemble de réglementations et de normes liées à la définition du risque et à sa maîtrise. Le référentiel ISO Guide 73 [10] sur le vocabulaire du risque ne lie pas ce dernier à une menace mais, comme Huygens, à l'incertitude d'un événement. Le risque y est défini comme "*l'effet de l'incertitude sur l'atteinte des objectifs*", une note précisant bien que "*un effet est un écart, positif et/ou négatif, par rapport à une attente*". Cette définition permet d'intégrer le risque financier et économique. Toutefois, dès qu'on en revient au risque industriel, sanitaire ou environnemental, ce sont bien les conséquences néfastes des processus qui sont mises en avant. Ainsi, l'Union Européenne définit comme suit la notion de risque grave pour la santé publique (Union Européenne 2006) :

- risque : probabilité qu'un événement se produise (on retrouve ici la définition ISO)
- risque potentiel grave pour la santé publique : une situation dans laquelle il existe une forte probabilité pour qu'un danger grave provoqué par un médicament (í ) affecte la santé publique
- grave : (í ) signifie un danger qui pourrait entraîner la mort, mettre en danger le patient, nécessiter une hospitalisation, entraîner une invalidité (etc í )

Pour décrire le risque dans cette perspective, nous proposons de suivre un cadre normatif standard, comme par exemple les normes européennes (EN 292-1 et EN 1050) relatives aux risques ayant incidence sur la santé humaine. Dans ce type de démarche, le risque est modélisé comme l'association de trois concepts :

- le **facteur de risque**, qui caractérise l'élément ou le processus susceptible de causer un risque, donc d'être la cause d'une situation indésirable. La question que nous nous posons donc est de savoir si les technologies langagières doivent être considérées comme des facteurs de risque. Ceci, en ayant conscience que toute technologie complexe peut constituer a priori une source plurifactorielle de risques variés (Beck 2001).
- la **criticité**, qui combine l'impact du risque (son effet ou sa gravité, pour reprendre le règlement européen détaillé plus haut) avec sa probabilité d'occurrence. La question que l'on se pose ici est l'évaluation de l'impact des technologies que nous développons. Cette évaluation peut être expérimentale (étude statistique sur une population de test) ou subjective et introspective (retour d'expérience d'experts, par exemple).
- la **vulnérabilité**, qui revient d'une part à décrire l'objet du risque, à savoir l'élément qui le subit (ce peut être aussi bien un écosystème, un individu ou l'ensemble de la société), et d'autre part ses conséquences (par exemple, la mort ou l'invalidité dans l'exemple du règlement européen cité précédemment).

A notre connaissance, les technologies langagières ont très rarement fait l'objet d'une analyse de risque ó voir toutefois (Kaplan 2014) pour une exception notable. Afin d'amorcer une telle démarche, nous proposons ici de nous focaliser sur les dimensions de vulnérabilité et de facteur de risque, afin déjà d'identifier quelles applications du TALN peuvent présenter des conséquences potentiellement néfastes. La question de la criticité est beaucoup plus délicate à aborder dans l'immédiat, car elle nécessite la mise en òuvre d'expérimentations ou de suivis utilisateurs qui peuvent être potentiellement très lourds (large cohorte de sujets ou étude longitudinale sur une longue durée). Elle ne saurait donc être abordée que sur les facteurs de risques pour lesquels une vulnérabilité critique est caractérisée.

Cette vulnérabilité devient rapidement cruciale dès lors qu'on touche des personnes déjà fragilisées : c'est par exemple le cas des personnes handicapées, pour lesquelles sont développés des systèmes d'assistance et de suppléance de plus en plus efficaces. Le travail qui est présenté ici a ainsi été initié dans le cadre du RTR (Réseau Thématique Régional) *Risques* de la région Centre s'intéressant à l'impact des aides techniques au handicap. Un des objectifs de ce RTR sera précisément de soutenir des études expérimentales de criticité sur des technologies d'assistance qui pour certaines impliquent des traitements linguistiques. Les travaux du RTR ont conduit à une première esquisse de classification du risque (Antoine et al. 2014) qui sert de base à la typologie arborescente présentée ci-après (§3).

### 3 Une première classification de risques et de leur vulnérabilité

Les deux premiers niveaux de la typologie de risques que nous présentons ici relèvent avant tout de la vulnérabilité (Figure 1). Le premier niveau correspond à l'objet de vulnérabilité en distinguant les atteintes à l'individu de ceux qui relèvent d'une dimension sociétale. Le risque est ensuite sous-catégorisé suivant 5 grandes classes d'impacts identifiés.

On classera en *risque individuel* tout risque dont l'objet de vulnérabilité se limite à l'individu (utilisateur d'un système technique ou autre). Dans le cadre de l'aide technique au handicap, on peut ainsi imaginer que l'objet de vulnérabilité soit par exemple un aidant. Le risque individuel est ensuite découpé en trois classes d'impacts principales :

- Le **risque physique** caractérise une atteinte à l'intégrité physique de l'individu par suite de l'usage d'un système technique. Il affecte le corps d'une personne et correspond à des blessures/dégradations, des traumatismes et/ou des handicaps physiques supplémentaires liés à l'utilisation du système.
- Le **risque cognitif** porte sur une altération dommageable de certaines fonctions cognitives du fait de l'utilisation d'un système technique. Dans le cadre de l'aide au handicap, les thérapeutes sont particulièrement attentifs à ce type de risque, une assistance trop importante pouvant par exemple induire une régression cognitive entraînant une perte d'autonomie en l'absence de dispositif d'assistance. De même, il a été montré que l'usage régulier de moteurs de recherche avait un impact sur nos stratégies de mémorisation à long terme (Sparrow et al. 2011). Nous verrons plus loin que le TAL est directement concerné par ce type de risque.
- Le **risque psychique** se traduit par une perturbation des affects, des réactions ou de la perception de la réalité, qui peut être accentuée par l'usage de nouvelles technologies. On peut par exemple citer le stress induit chez certaines personnes par leur usage. Un exemple relevant du traitement de la parole concerne les synthèses vocales qui utilisent la voix désormais perdue d'un patient (cas de maladies neurodégénératives). On manque encore de recul sur l'introduction de cette pratique, mais nos discussions avec des praticiens montrent que l'adoption de cette voix à forte charge émotionnelle, puisque porteuse forte d'identité, est tout sauf anodine d'un point de vue psychologique.

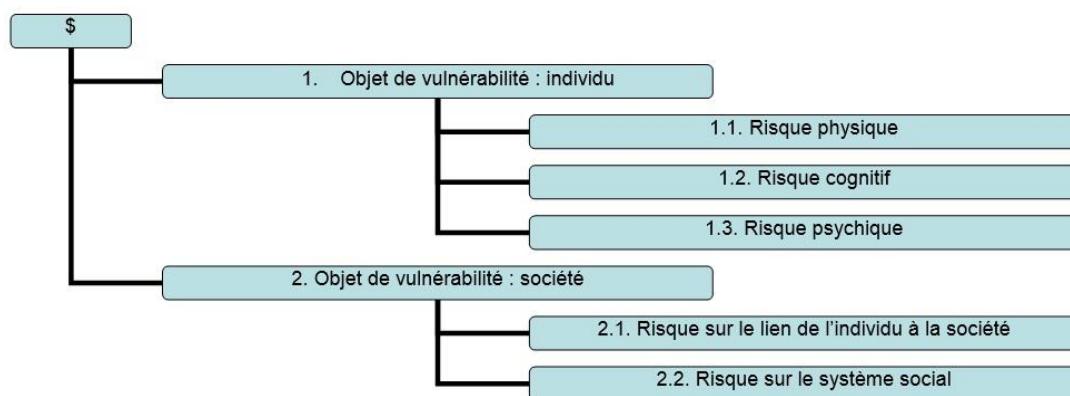


FIGURE 1 : Deux premiers niveaux de notre typologie de risques : objet de vulnérabilité et classe d'impact

Le *risque sociétal* met en jeu autant les relations sociales des individus que le système social en lui-même. Il s'agit donc d'impacts qui dépassent l'utilisateur seul d'un système technique, et qui sont souvent sous-évalués. Nous distinguons précisément deux types de risques sociétaux :

- Le risque d'altération des **relations de l'individu à la société** qui concerne aussi bien la modification du lien social que le rapport de l'individu à la société par l'intermédiaire du respect de droits et libertés individuelles. On peut ainsi citer l'impact des technologies numériques sur le droit du travail, le respect de la vie privée, mais aussi les modifications des liens inter-individus avec l'utilisation massive des réseaux sociaux.
- Le risque de modification du **système social**, dans sa dimension politique, économique ou culturelle. Pensons par exemple au vote électronique, ou aux questions d'invisibilité du handicap dans une société numérique.

Cette présentation des premiers niveaux de notre typologie n'a pas fait appel, à dessein, à des exemples relevant des technologies langagières. Nous espérons en effet que cette classification présente un degré de généralité qui lui permet de s'appliquer à d'autres technologies numériques. Nous allons maintenant présenter en détail chaque élément de notre typologie en l'illustrant cette fois avec des applications relevant du TALN, afin de mener une évaluation critique des applications proposées par le TAL dans un contexte où leur diffusion va aller croissant dans les années à venir.

## 4 Une classification typologique des facteurs de risques adaptée au TALN

La classification des facteurs de risques que nous avons introduite dans le paragraphe précédent a été présentée dans le cadre de la journée d'études « Ethique et TAL » organisée par l'ATALA (Antoine et al. 2014). Si elle fournit un cadre général que nous espérons utile pour une analyse du risque du TAL, elle reste toutefois trop générale pour permettre la mise en place d'un diagnostic éthique face à une thématique de recherche précise. C'est pourquoi nous avons continué à affiner cette classification pour proposer une caractérisation du risque suivant une typologie pouvant présenter jusqu'à cinq niveaux hiérarchiques de caractérisation. La typologie que nous proposons pourrait certainement répondre en partie aux interrogations concernant d'autres technologies numériques (big data et analyse décisionnelle, réseaux etc.). Dans cette section, nous allons toutefois chercher à illustrer notre typologie en nous focalisant uniquement sur le domaine des technologies langagières.

### 4.1 Risque physique

Le risque physique concerne assez peu une technologie comme le traitement des langues qui a rarement pris sur l'environnement physique. Il ne doit toutefois pas être négligé. Notre typologie distingue ici les facteurs de risques concernant l'utilisateur d'un système technique de ceux entraînant un dommage sur l'environnement physique. Dans ce dernier cas, on peut citer l'exemple d'un fauteuil roulant autonome piloté par commande vocale qui heurterait un mur ou un meuble suite à une erreur de commande ou de reconnaissance. Tout facteur de risques correspond à un nœud terminal de notre typologie. Ici, il s'agit du nœud 1.1.2 (figure 2).

Un impact physique direct sur l'utilisateur a été caractérisé dans le cas de l'utilisation de la reconnaissance vocale dans les centres logistiques. Afin de permettre un travail mains libres, les préparateurs de commande sont guidés dans leur mission grâce à un dialogue oral homme-machine. Ce mode de gestion entraîne une densification du travail qui peut entraîner une augmentation des lombalgies ou des troubles musculo-squelettiques (INRS 2009). Notre typologie

distingue deux types d'impacts sur l'utilisateur : les atteintes physiologiques internes (modification hormonale par exemple) et les atteintes physiques externes correspondant précisément aux effets néfastes que nous venons d'évoquer.

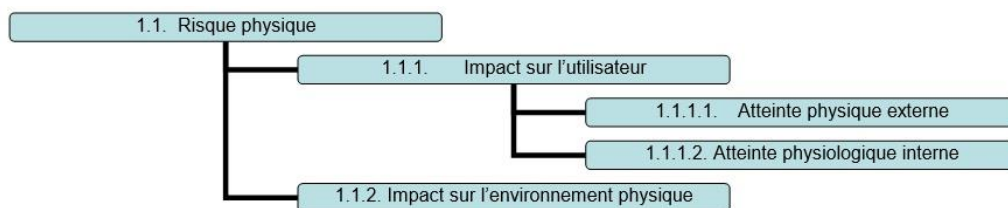


FIGURE 2 : Sous-typologie des risques physiques

## 4.2 Risque cognitif

Bien que notre communauté scientifique se soit peu penchée sur la question par le passé, le risque cognitif concerne de manière significative les technologies langagières. Nous l'avons dit dans la section précédente, ce risque porte sur une altération dommageable de certaines fonctions cognitives ou de l'état cognitif général de l'utilisateur. Comme le montre la figure 3, le premier sous-niveau de classification du risque cognitif concerne précisément cette distinction.

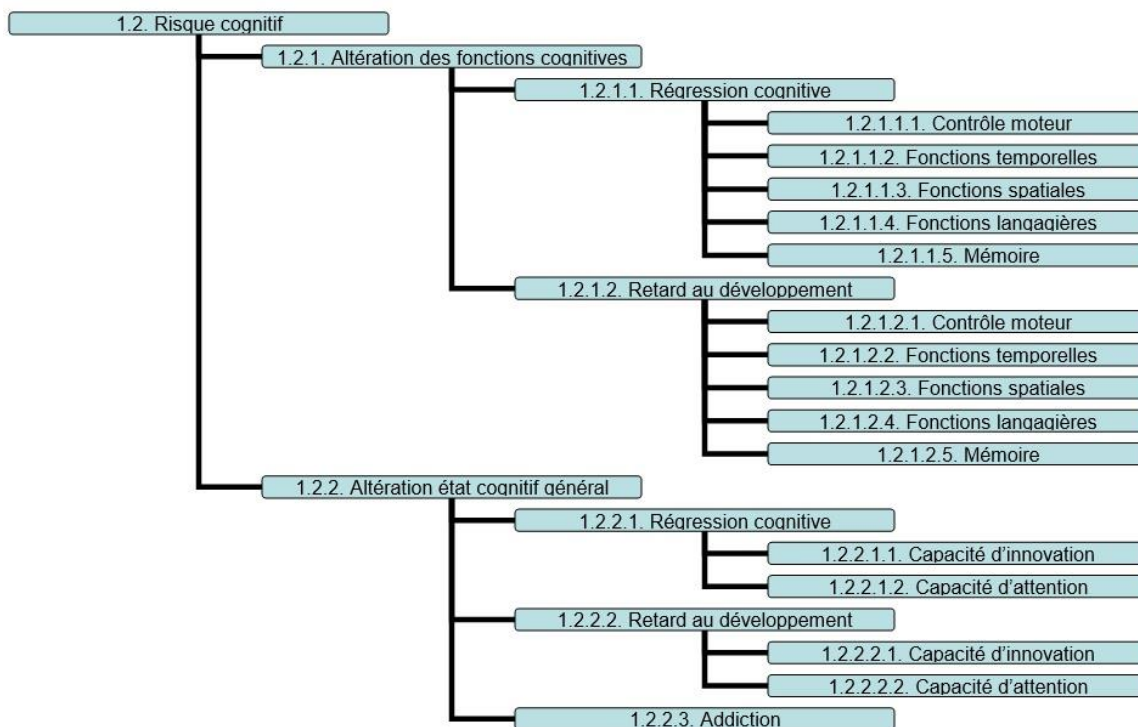


FIGURE 3 : Sous-typologie des risques cognitifs

Le niveau suivant distingue la régression cognitive (déjà vue plus haut), qui entraîne la perte partielle ou totale d'une fonction cognitive du fait de l'usage d'une technologie, du retard au développement qui se traduit par une altération de l'acquisition de cette fonction du fait de l'usage d'une technologie lors des stades de développements cognitifs infantiles. Ces deux types d'altérations sont ensuite différenciés suivant la fonction cognitive concernée :

- Contrôle moteur, mémorisation, fonctions liées aux représentations spatio-temporelles ou au langage, etc. dans le cas de l'altération des fonctions cognitives
- Fonctions plus globales telles que la capacité d'attention ou d'innovation dans le cas des atteintes à l'état cognitif général de l'individu

Cette structuration s'inspire directement de la classification CIF des fonctions mentales globales et spécifiques proposée par l'Organisation Mondiale de la Santé (OMS 2001). La CIF peut d'ailleurs servir de cadre à une généralisation ou un raffinement (niveaux hiérarchiques supplémentaires) de notre typologie. S'ajoute enfin la question des addictions cognitives qui ont un impact général sur le fonctionnement cognitif. Toutes ces atteintes ne

concernent pas le TALN. Nous allons toutefois donner quelques exemples de risques liés aux technologies langagières qui montrent l'importance d'une réflexion sur les impacts cognitifs des technologies que nous élaborons.

#### **Altération des compétences langagières : prédiction de mots ou auto-complétion pour l'aide à la saisie de texte**

Ce type d'applications relève d'une modélisation statistique du langage ou d'une consultation de lexique dans le cas des systèmes les plus simples. Il concerne aussi bien l'aide à la communication pour personnes handicapées que la saisie de texte sur dispositifs mobiles (Antoine 2011). Ces techniques sont généralement évaluées à l'aune des bénéfices qu'elles permettent en termes de vitesse de saisie, voire parfois de développement cognitif. Par exemple, il a été observé que la prédiction lexicale avancée du système d'aide à la communication Sibylle (Wandmacher et al. 2008) induisait un accroissement des productions langagières chez des enfants infirmes moteurs cérébraux, et que l'augmentation des interactions langagières qui en résultait favorisait le développement cognitif de certains enfants. En outre, leurs enseignants de l'école intégrée au centre de Kerpape ont remarqué une baisse notable des fautes d'orthographe commises. Une analyse plus fine demande toutefois d'étudier l'impact de ces technologies d'un point de vue rapport bénéfices/risques. La question que se posent en effet orthophonistes et éducateurs est de savoir si cette aide augmente la maîtrise du système de la langue, ou si elle ne masque au contraire pas un abandon de cette capacité au profit du système. Ce risque demande à être évalué en termes de criticité, ce qui demanderait la mise en place de tests de compétences linguistiques longitudinaux. Ce risque se classe sous le n° 1.2.1.2.4. de notre typologie, soit suivant le chemin: Individu > Risque Cognitif > Retard au développement > Fonctions langagières

Ces aides à la saisie et à la composition de texte se retrouvent dans les correcteurs orthographiques et grammaticaux comme dans toutes les formes d'aide à la traduction désormais disponibles. Qui parmi nous n'a pas ressenti une forme de perte de maîtrise de ses compétences langagières due à l'abandon facile à ces aides techniques ? Pour paraphraser le philosophe Bernard Stiegler (Stiegler 2015), l'homme augmenté (par les technologies langagières) est un homme diminué (en termes de compétences langagières). Nous sommes ici en présence d'une situation de régression cognitive dont il conviendrait d'évaluer la criticité (impact en termes de compétences langagières).

**Mémoire altérée : texte numérique et abandon de l'écriture cursive** Si, comme l'a montré l'exemple précédent, on imagine aisément que l'utilisation d'un système de traitement automatique du langage peut influencer sur les fonctions langagières d'un utilisateur, cet impact peut également concerner d'autres fonctions cognitives. C'est le cas en particulier des fonctions de mémorisation, pour lesquelles les exemples d'influence des technologies numériques se multiplient. Nous avons déjà cité plus haut le "Google effect" relevé par (Sparrow et al. 2011). Cet effet est d'autant plus insidieux qu'il n'est pas ressenti par les sujets, qui tendent au contraire à surestimer l'importance des connaissances mémorisées lorsqu'ils utilisent un moteur de recherche (Fisher et al. 2015). Certaines études nous conduisent à nous demander si un tel impact négatif ne peut pas être également induit par la saisie numérique de texte au détriment de l'écriture cursive. Suite à l'annonce de l'abandon de l'apprentissage scolaire de l'écriture cursive aux Etats-Unis et en Finlande, de nombreux psychologues du développement ont fait remarquer que la mobilisation des aires motrices lors de l'écriture favorise la mémorisation (Longcamp 2003). La prévalence de l'écriture numérique que favorisent les technologies langagières peut donc induire un risque mémoire dont il serait intéressant d'évaluer la criticité. Ce risque concerne ici encore un éventuel retard au développement classé au n° 1.2.1.2.5.

**Capacité d'innovation et création linguistique : aide à la saisie de texte** Les systèmes d'aide à la saisie basés sur la suggestion et l'auto-complétion proposent des items lexicaux, ou lexico-syntaxiques qui sont puisés au sein de ressources linguistiques. Le choix de l'item suggéré se conforme le plus possible à un idéal linguistique : dans le cas d'un système probabiliste, l'idéal est défini par le nombre, dans le cas d'un système symbolique, c'est une norme qui fait loi par exemple. Quelle place accorder dès lors à la créativité linguistique (par le détournement par exemple), et serait-il intéressant de l'évaluer ? Depuis l'invention de l'écriture jusqu'à l'arrivée de la télévision, on assiste à l'épanouissement de telles tendances directives étudiées par la sociolinguistique (voir par exemple à la revue électronique Glottopol pour une approche des politiques linguistiques et l'impact des média sur les pratiques discursives), il semble donc intéressant de mettre en évidence ce risque que nous avons placé à l'indice 1.2.2.2.1.

### **4.3 Risque psychique**

Le risque psychique se traduit par une perturbation de l'état psychologique de l'individu de par l'usage d'un système technique. Ces altérations peuvent concerner les affects, les réactions aux situations, la perception de la réalité etc. Leurs effets peuvent être temporaires, comme lors d'un choc émotionnel ou bien plus durables comme dans le cas d'un renforcement d'état dépressif. Le premier niveau de sous-catégorisation de notre typologie du risque psychique, proposée en figure 4, repose précisément sur une analyse de la vulnérabilité en termes de durée temporelle (durable ou temporaire). Nous distinguons ensuite le risque suivant la variable psychologique qui est affectée chez l'individu.



Dans le cas d'un impact psychique durable, il nous semble envisageable de nous baser sur les différentes classifications internationales de troubles mentaux (OMS 2006, APA 2003) telles que celles proposées par l'OMS (CIM-10) et l'Association Américaine de Psychiatrie (DSM-IV-TR). Bien que parfois controversées, ces classifications s'accompagnent de critères diagnostiques qui peuvent être utiles pour une étude en criticité d'un facteur de risque. Ces classifications concernent toutefois le plus souvent des formes sévères de troubles psychiques. On peut se demander si les technologies langagières constituent des facteurs de risques d'une telle criticité. Ces classifications doivent donc avant tout être considérées comme des lignes directrices d'analyse. Notre réflexion sur les technologies langagières et/ou d'aide au handicap nous a toutefois conduits à caractériser des effets potentiels se rapprochant de troubles de l'humeur (CSM-10 F32 à F38), de troubles anxieux ou de la personnalité (CSM-10 F60).

Nous avons tenu par ailleurs à distinguer des effets temporaires que nous avons observés lors de nos travaux sur l'aide à la communication pour les personnes handicapées : fatigue psychique, situations de stress ou chocs émotionnels. Ce type d'impact ne doit pas être négligé au titre de son caractère transitoire : par exemple, la fatigue psychique dû à l'effort cognitif demandé par l'usage de tels systèmes est une cause très fréquente d'abandon de leur usage. Cet impact se retrouve en situation de travail avec le guidage par commande vocale des préparateurs logistiques (INRS 2009).

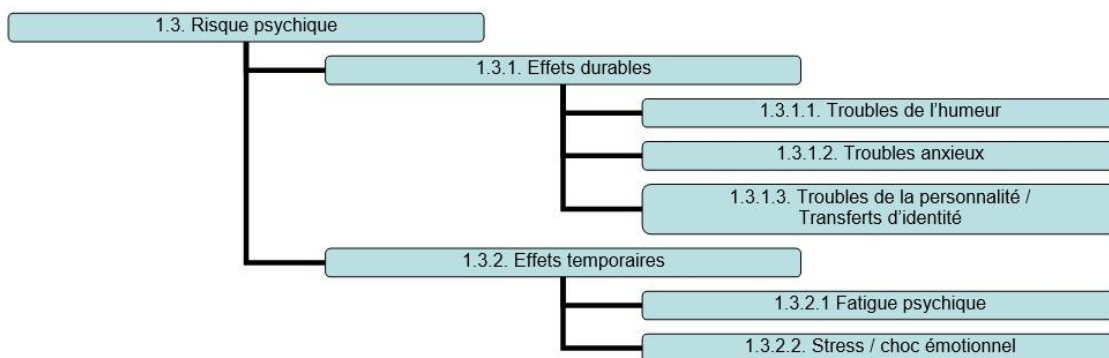


FIGURE 4 : Sous-typologie des risques psychiques

Quelques exemples vont nous permettre d'illustrer l'application de cette sous-typologie sur des applications relevant du traitement des langues naturelles ou du traitement de la parole.

**Synthèse vocale, Agents Communicationnels Artificiels (ACA) : choc émotionnel et troubles de l'identité** La synthèse vocale à partir du texte a atteint désormais un degré de naturalité suffisante pour permettre des applications grand public de plus en plus confondantes. On la retrouve ainsi pour les annonces dans certains réseaux de transports en commun. C'est son utilisation à des fins personnalisées qui interroge toutefois en priorité. On pense aux synthèses vocales qui utilisent la voix désormais perdue d'un patient dans le cas de maladies neurodégénératives. Nos discussions avec des praticiens montrent que l'adoption de cette voix à forte charge émotionnelle, puisque porteuse d'identité, est tout sauf anodine d'un point de vue psychique. Deux situations semblent être à contrôler avec attention :

- l'annonce au patient qu'il va perdre sa voix et qu'il faut s'y préparer en procédant à son enregistrement : cette séance de recueil peut induire un choc émotionnel critique que la technologie peut renforcer (risque 1.3.2.2).
- l'usage ultérieur de la voix recueillie avec un logiciel d'aide à la communication, qui peut entraîner des effets durables en termes d'identité (risque 1.3.1.3). Les premiers retours d'expérience suggèrent que ce facteur de risques concerne, en termes de vulnérabilité, la personne handicapée comme son entourage (famille, aidants).

Ces troubles psychiques liés à la synthèse de parole ne peuvent être ignorés : ils nous semblent équivalents à ceux, bien documentés, liés aux greffes d'organes (Triffaux et al. 2002). Et ce d'autant qu'ils concernent un trait de personnalité aussi important que le visage, dont la greffe fait l'objet de questionnements éthiques (Colas-Benayoun et al. 2006).

Ce risque psychique peut être généralisé à toutes tentatives d'imitation du vivant, parmi lesquels les travaux relevant de l'interaction affective. Dans le domaine du TALN, ils concernent par exemple la détection des émotions ainsi que les recherches en dialogue homme-machine sur les agents conversationnels animés (ACA). Si les ACA destinés au grand public restent encore assez frustrés (voir par exemple Laura, du site EDF particuliers : [www.bleuciel.edf.com](http://www.bleuciel.edf.com)), leur inspiration reste anthropocentrée. On peut s'interroger sur l'impact psychique que pourrait avoir à l'avenir l'échange avec un agent artificiel plus convaincant (phénomènes de transferts, attentes trop fortes dans l'interaction, etc.). Sur un sujet proche, la CERNA vient précisément de produire des recommandations portant sur la pertinence de l'imitation du vivant dans le domaine de la robotique (CERNA 2014). Ses propositions sont proches de nos préoccupations, comme le montre par exemple la recommandation IVI-1 sur l'utilité au regard des finalités :

(í ) Dans les cas où l'apparence ou la voix humaines sont imitées, le chercheur s'interrogera sur les effets que pourrait avoir cette imitation, y compris hors des usages pour lesquels le robot est conçu.

#### 4.4 Risque sur le lien social et sur le rapport de l'individu à la société

Du point de vue des applications grand public, c'est ainsi plus les situations de vulnérabilité sociétale qu'individuelles qui ont retenu l'attention des chercheurs en TALN. Cette réflexion s'inscrit avant tout dans les interrogations contemporaines sur la société de l'information numérique créée par Internet : la question de l'anonymisation des corpus, qui relève de fait de la loi Informatique et Liberté du 6 janvier 1978, revêt par exemple un caractère encore plus sensible avec les recherches portant sur les réseaux sociaux.

Toutefois, l'impact sociétal du TALN dans notre société numérique couvre bien d'autres dimensions qui sont le plus souvent ignorées. La première classe de risques que nous avons tenu à identifier concerne l'altération des rapports que les individus entretiennent dans la société : lien social inter-individu médié par les outils numériques, relation avec une technologie de plus en plus anthropocentrée, et enfin respect des droits et libertés individuelles que la société définit par l'intermédiaire de la réglementation (droit du travail, propriété intellectuelle, respect de la vie privée, liberté d'expression et d'information, etc.). La sous-typologie de risques qui en résulte est décrite figure 5. Nous allons illustrer à l'aide de quelques exemples d'applications langagières nécessitant analyse.

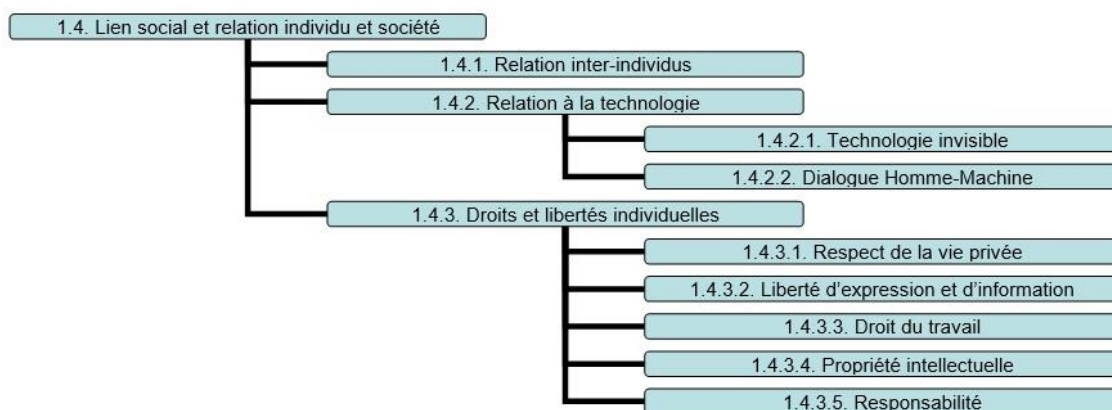


FIGURE 5 : Sous-typologie des risques sur le lien social et le rapport de l'individu à la société

**Synthèse de parole, agents conversationnel et interaction affective : altération du lien social** ó Nous avons identifié plus haut les risques psychiques que présentent la tentation d'utiliser les technologies langagières et/ou robotiques à des fins d'imitation du vivant. Lorsqu'un agent virtuel a pour finalité de se substituer à un être humain dans l'échange, il existe un risque potentiel de voir l'utilisateur s'enfermer dans des relations purement virtuelles ou fortement médiées par la technologie dans lesquelles la conscience de la distinction entre humanité et intelligence artificielle peut être brouillée. Ici encore, ce risque a été étudié par la CERNA, dont nous rappelons les recommandations (CERNA 2014) :

- Si une ressemblance quasi parfaite est visée, le chercheur doit avoir conscience que la démarche biomimétique peut brouiller la frontière entre un être vivant et un artefact (recommandation IVI-2).
- Pour les projets (í ) qui ont trait au développement de la robotique affective, le chercheur s'interrogera sur les répercussions éventuelles de son travail sur les capacités de socialisation de l'utilisateur (IVI-3)

On observe que ces impacts potentiellement négatifs concernent aussi bien l'individu (risque 1.3.1.3 évoqué plus haut) que son rapport à autrui : ce second risque est classifié 1.4.1 dans notre typologie. Au fur et à mesure des progrès futurs de l'interaction homme-machine, il est probable qu'un risque spécifique devra être identifié, tant une relation d'ordre nouveau devra être identifiée. Nous l'avons classifié au n° 1.4.2.2 de notre typologie.

**Synthèse vocale, génération automatique de texte : invisibilité de la technologie** ó Les exemples précédents de brouillage entre être humain et artefact concernent l'interaction directe entre individus. D'une manière plus générale, il est important de s'interroger sur notre rapport à la technologie. Quelle place devons-nous lui accorder dans nos sociétés techniques ? Dans quelle mesure redéfinit-elle la notion d'humanité ? Pour aborder ces questions, l'utilisateur doit avoir une conscience claire des interventions de la technologie dans son quotidien. Cette question de l'invisibilité de la technologie concerne la synthèse vocale, déjà évoquée. Il faut également évoquer l'intervention de la génération

automatique de texte dans des domaines aussi variés que la gestion automatique de la relation client et la rédaction automatique de compte-rendus boursiers ou de résultats électoraux. Ces interventions sont si sensibles que le quotidien *Le Monde* a tenu à les expliciter à l'occasion de la dernière consultation départementale (Le Monde 2015). Nous classons ces problèmes d'invisibilité sous le n<sup>o</sup> ud 1.4.2.1 de notre typologie.

**Détection d'auteur sur les réseaux sociaux : respect de la vie privée** ó Le respect de la vie privée est certainement un des droits individuels sur lesquels la population est la plus sensibilisée avec l'émergence des réseaux sociaux. Cela n'empêche toutefois pas de nombreux chercheurs en TALN et recherche d'information de mener des recherches sur la recherche d'auteurs non identifiés explicitement (Stamatatos 2009). Historiquement, ces recherches en AA (*Authorship Authentication*), issues de la stylistique quantitative, ont eu une justification scientifique (attribution d'auteur à des manuscrits littéraires anciens) puis juridique (détection de plagiat). Leur application à la fouille sur réseaux sociaux pose au contraire des problèmes éthiques qui recommanderaient de s'interroger sur l'impact de ces recherches : doit-on pouvoir lever l'anonymat sur un post publié sous un pseudonyme ? Est-il légitime d'utiliser comme preuve juridique ces techniques à la précision perfectible ? Ces atteintes à la vie privée sont classées au n<sup>o</sup> ud 1.4.3.1 de la typologie.

**Myriadisation et TALN : droit du travail, propriété intellectuelle** ó Le droit du travail (n<sup>o</sup> ud 1.4.3.3. de notre typologie) est également une dimension qui ne doit pas être oubliée dans le cadre d'une réflexion éthique. Ici, la technologie n'entre dans la réflexion qu'à la marge : depuis la révolution industrielle, le statut de la machine dans l'activité a été réglementé par la loi. Du fait de l'émergence d'un TAL centré sur les données faisant un large usage de ressources langagières, c'est sur nos pratiques quotidiennes que nous devons nous interroger. Le recours à un travail parcellisé et myriadisé (*microsourcing* et *crowdsourcing* en anglais) pour l'obtention de telles ressources doit ici être questionné. Dans leur analyse critique d'Amazon Mechanical Turk, (Sagot et al. 2011) montre ainsi qu'une part non négligeable des Turkers y réalise une véritable activité dissimulée représentant une part significative de leurs revenus. Ce travail dissimulé concerne directement le TAL, puisque les tâches de transcription et de traduction y sont fréquentes. Cette question concerne également les GWAP (*Game With A Purpose*) dédiés à la production de ressources linguistiques tels que JeuxDeMots (Lafourcade et Lebrun, 2014) et Zombilingo (Fort et al. 2014), dont les concepteurs insistent sur l'importance que l'activité des joueurs reste strictement ludique et par conséquent non rémunérée.

Ces activités et plus généralement la constitution de toute ressource linguistique posent en outre des questions de propriété intellectuelle identifiées comme le n<sup>o</sup> ud 1.4.3.4 de notre typologie. Si la législation sécurise ces questions, il est essentiel d'adopter de bonnes pratiques sur la documentation des licences, l'identification des propriétaires et des intervenants dans la constitution pour pouvoir sécuriser ces dernières (Couillault et Fort 2013, Baude et al. 2006).

**Agents virtuels communicants : responsabilité individuelle** ó Le statut des machines considérées doit être réévalué suite à l'émergence du TAL, d'une part à cause de son impact sur les évolutions des conditions de travail tel que décrit dans la section précédente, mais aussi par l'effet des productions langagières de ces machines. Dans une société où le lien social est lui aussi mécanisé par des plateformes de mise en relation, l'introduction d'agents conversationnels autonomes est une pratique déjà très répandue<sup>1</sup> et une véritable économie se développe sans que les questions de responsabilités qu'elles induisent n'aient été discutées. Les actes performatifs nuisibles tels que la diffamation, l'insulte, etc. sont déjà encadrés dans la sphère publique, mais quel statut juridique doit être accordé à un agent conversationnel personnel à la source de telles données ? Si la jurisprudence ou la loi attribue la responsabilité à l'agent, alors le créateur est dédouané de responsabilité, ce qui fait des agents une arme performative redoutable, tandis que si le créateur est tenu responsable, alors il est possible qu'il soit injustement attaqué pour simple mauvais réglages. Se pose dès lors aussi la question de la répartition équitable de la responsabilité dans le cas où l'agent a été créé par plusieurs personnes. Ces situations nourrissent un débat juridique actuel sur la nécessité ou non de l'attribution d'une personnalité juridique aux robots et agents conversationnels (Robolaw 2014, Bensoussan 2015). Elles relèvent du risque 1.3.4.5 de notre typologie.

## 4.5 Risque sur le système social

Enfin, le dernier type de risques concerne la modification de l'état de la société dans sa globalité et dans au moins une de ses composantes. Nous pensons ainsi à l'influence des technologies numériques sur la prise de décision politique. Ainsi, les aides techniques aux personnes handicapées peuvent limiter leur visibilité dans l'espace politique, avec pour conséquence éventuelle une limitation des politiques d'aides à leur égard et à celle des aidants dans nos sociétés

<sup>1</sup> Le nombre de « bots » utilisés pour la publicité sur les plateformes de rencontres a décuplé ces derniers temps, et même si le test de Turing n'est pas encore passé, le « bot » Ava a réussi le test de Tinder (<http://www.adweek.com/adfreak/tinder-users-sxsw-are-falling-woman-shes-not-what-she-appears-163486>).

budgétairement contraintes. La question de l'influence de l'automatisation sur le marché de l'emploi et l'économie est documentée depuis les débuts de la révolution industrielle. Nous verrons dans les exemples ci-dessous que cette question ne peut être éludée par les technologies langagières. Enfin, l'impact socio-culturel du TALN associé aux grandes masses de données disponibles sur les réseaux, ne peut plus être ignoré. En particulier, il convient de s'interroger sur l'influence des technologies langagières sur le système même de la langue. La figure 6 ci-dessous résume ces différents types de risques que nous allons une nouvelle fois illustrer à l'aide de quelques exemples.

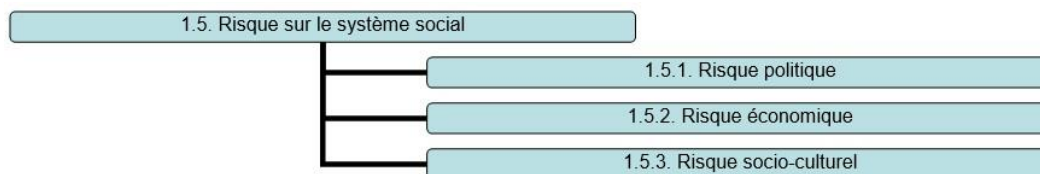


FIGURE 6 : Sous-typologie des risques sur le système social dans sa globalité

**Indexation automatique, fouille de texte et recherche d'information : risque politique et surabondance de l'information** ó Au sein de notre société numérique connectée, les technologies langagières ne sont généralement pas des facteurs de risques primaires, mais des facteurs aggravant de criticité auxquels il convient de prendre garde. Face à une évolution rapide d'une rareté vers une surabondance nocive de l'information (Ganascia 2013 ; Mariani 2013 :14), le TALN a un rôle important à jouer : l'indexation automatique de contenus, la fouille de texte et la recherche d'information peuvent aider l'utilisateur à donner du sens à son immersion dans de larges flux d'information et de connaissances. Mais elles peuvent également participer à l'illusion trompeuse d'une maîtrise de ces grandes masses de données, comme le suggèrent les expérimentations (cf § 4.2) sur le sentiment erroné de connaissance que développent les utilisateurs de moteurs de recherche (Fisher et al. 2015). Ainsi, la mise en œuvre de robots logiciels pour la rédaction de textes journalistiques, déjà citée, où bien de moteurs de filtrage automatiques nous semblent jouer contre un contrôle conscient de nos stratégies personnelles de contrôle et de sélection de l'information. Ce risque politique est catégorisé au n° 1.5.1. de notre typologie.

**Technologies langagières et économie** ó Les technologies langagières sont de plus en plus concernées par l'économie de la société numérique de l'information. Leur impact économique ne peut donc être ignoré. A titre d'exemple (Sagot et al. 2011) fait une revue des études qui ont été menées sur le coût des ressources créées avec Amazon Mechanical Turk. Dans un autre contexte, la recherche d'information a créé "*la bourse des mots*", algorithme permettant à un annonceur de proposer une enchère sur un mot tapé dans un moteur de recherche afin de lui affecter la publicité la plus appropriée (à la fois la plus de chance d'être cliquée et la plus de chance d'être tapée en premier lieu). Cette économie linguistique très étudiée par Kaplan (2014) conforte une disparité entre langues richement et faiblement dotées. Ce risque (au n° 1.5.2 de notre classification) est abordé par d'autres auteurs tels que Enguehard et al. (2014).

**Technologies langagières et système de la langue** ó L'usage des technologies langagières influence l'évolution de la langue, tout comme le fit jadis le passage de l'oral à l'écrit : d'un certain point de vue, à chaque bond technologique en matière de supports<sup>2</sup>, le système linguistique dans son ensemble est impacté. Par exemple la sémantique et les mécanismes référentiels à l'œuvre dans l'usage de "*technomorphèmes*" tels que les hyperliens, les hashtags dans les discours numériques commencent à susciter l'intérêt des analystes du discours (Paveau 2014). Frédéric Kaplan (2014) montre par ailleurs que le choix de l'anglais comme langue pivot entre deux autres idiomes par Google Translate se traduit tout d'abord par un biais culturel qu'il faudrait étudier. Mais qu'en outre, les textes produits ainsi automatiquement (et leurs erreurs), peuvent être prises à tort comme ressources primaires et participer à une forme de "*créolisation numérique*". Au-delà d'un simple enrichissement lexical nécessaire à désigner une nouvelle réalité, on assiste à un genre d'hybridation de la langue entre productions naturelles et productions artificielles pour lesquels l'appareillage méthodologique de l'analyste se doit d'évoluer. De même, on pourrait examiner l'impact de la simplification de texte, domaine du TAL qui propose de remédier provisoirement aux difficultés lexicales d'apprenants ou de souffrants de pathologies dans le but de les accompagner vers une compétence langagière de meilleure qualité. Son principe est de remplacer au sein d'un texte un item lexical par exemple dont la complexité est jugée trop poussée, par un concurrent sur son axe paradigmatique dont la complexité est inférieure. Cette méthode permettrait de rendre accessibles à l'utilisateur diverses productions discursives dont des œuvres littéraires par exemple. On peut chercher à évaluer l'impact de cette pratique sur l'intégrité et la portée des discours, et à terme sur l'évolution du système linguistique dans son ensemble. Ces modifications sont regroupées sous le n° 1.5.3.

<sup>2</sup> On extrapole un peu ici l'acception de technologies du TAL pour considérer tout l'appareillage numérique utile à la production linguistique : SMS, tweet, forum, etc.

## 5 Conclusion

Dans cet article, nous avons cherché à mener une réflexion éthique sur les conséquences de la diffusion de plus en plus importante des technologies langagières. Cette réflexion s'articule avec une volonté d'évaluation des systèmes développés par notre domaine de recherche. Les technologies langagières ouvrent des possibilités d'augmentation de l'humain de plus en plus présentes dans le quotidien, certaines étant directement issues d'une recherche d'aide aux personnes handicapées (auto-complétion par exemple). Cette situation nous plonge directement dans les questionnements autour du posthumanisme et du transhumanisme, ayant respectivement pour principe la réparation puis l'augmentation de l'humain (Kleinpeter 2013). En ce qui concerne le langage, nous tenons à rappeler que les technologies permettant une augmentation de la compétence langagière devrait être mises en perspective avec celles qui les ont précédées, à savoir l'écriture, l'imprimerie, ou encore la télévision. C'est donc dans cette perspective que nous cherchons à savoir en quoi l'augmentation proposée par ces outils reste une forme d'amélioration et à limiter les effets pervers de l'utilisation massive de ces derniers par une maîtrise étroite des risques qui lui sont liés.

Ce que nous avons tenté de montrer est que, comme toute autre technologie, les technologies langagières, ne sont pas des objets neutres mais ont un impact individuel et sociétal sur lesquels le chercheur doit s'interroger. Une pratique éthique de recherche en TALN ne peut donc se résumer aux questions importantes d'anonymisation qui ont le plus souvent concentré l'attention de la communauté, mais concerner plus généralement les risques psychologiques, cognitifs, sociétaux induits par ces technologies.

Une conclusion préliminaire à ce travail est la nécessité de classer le facteur de risques d'une part et de classer les impacts d'autre part. Nous avons ainsi proposé une typologie de risques qui permet de guider une évaluation éthique des technologies langagières. Nous avons vu qu'un facteur de risques peut impliquer plusieurs impacts de vulnérabilité de notre typologie (des capacités cognitives d'un locuteur au système linguistique dans son ensemble). Cette analyse des facteurs de risques et de leur vulnérabilité associée doit maintenant être complétée, quand cela le nécessite, d'une étude expérimentale de leur criticité. C'est ce que nous nous envisageons de faire désormais dans le cas des technologies langagières d'aide au handicap, ceci dans le cadre du RTR Risques de la région Centre.

## Remerciements

Nous tenons à remercier Christian Toinard (LIFO, U. Orléans) pour sa participation à la réflexion sur les risques induits dans le domaine de l'aide au handicap et Adrien Granger pour avoir partagé son expertise en matière de "chatbots" et ses réflexions sur la responsabilité engagée par ceux-ci (§ 4.4, sous-section sur les agents virtuels).

## Références

- ANTOINE J.-Y. (2011) Prédiction de mots et saisie de requêtes sur interfaces limitées : dispositifs mobiles et aide au handicap, In. Bellot P. (Ed). *Recherche d'information contextuelle, assistée et personnalisée*. Hermès, Paris. 273-298.
- ANTOINE J.-Y., LEFEUVRE A., ALLEGRE W. (2014). Pour une réflexion éthique sur les conséquences de l'usage des NTIC: le cas des aides techniques (à composante langagière ou non) aux personnes handicapées. *Journée ATALA "Ethique et TAL"*, novembre 2014.
- ANTOINE J.-Y., LABAT M.-E., LEFEUVRE A., TOINARD C. (2014b) Vers une méthode de maîtrise des risques dans l'informatisation de l'aide au handicap. Actes *Envirorisk2014*. Bourges.
- APA (2003) *DSM-IV-TR, Manuel diagnostique et statistique des troubles mentaux*, Elsevier Masson, Paris 2003.
- BAUDE O. ET AL (2006) *Corpus oraux, guide des bonnes pratiques*. Presses Universitaires d'Orléans, CNRS Editions.
- BECK U.(2001). *La société du risque*, Flammarion, Champs/essais.
- BENSOUSSAN A. (2015) Faut-il des lois pour nous protéger des robots ? *L'Expansion*. 4 avril 2015.
- CERNA.(2014). Ethique de la recherche en robotique. Rapport n° 1 de la CERNA. CERNA ó ALLISTENE . Section IVI : limitation du vivant et l'interaction affective et sociale avec les humains.
- COLAS-BENAYOUN M.D., FIDELE G., FAVRE J.D. (2006), De la défiguration à la transfiguration : la greffe d'un visage est-elle la solution ? *Annales Médico-Psychologiques*, 16(8), 687-691.
- COULLAULT A., FORT K. (2013) Charte éthique et big data : parce que mon corpus le vaut bien ! Acte colloque *Linguistique, Langue et Parole : statuts, usages et mésusages*. Strasbourg.



- CRUTZEN P.J., STEPHEN W., MC NEILL J. (2007), The Anthropocene: are Humans now Overwhelming the Great Forces of Nature? *Ambio*, 36, 614-621.
- ENGUEHARD C, MANGEOR M. (2014). Favorisons la diversité linguistique en TAL. *Journée ATALA "Éthique et TAL"*.
- FISHER, M., GODDU, M. K., & KEIL, F. C. (2015). Searching for Explanations: How the Internet Inflates Estimates of Internal Knowledge. *Journal of Experimental Psychology: General*. Advance online publication (30 mars 2015) consultee sur : <http://dx.doi.org/10.1037/xge0000070>.
- FORT K., GUILLAUME B., CHASTANT H. (2014) Creating Zombilingo, a Game With A Purpose for dependency syntax annotation. Prof. *Gamification for Information Retrieval Workshop (GamifIRø4)*, Amsterdam, Pays-Bas.
- GANASCIA J.G. (2013) L'initiative Onlife de la commission européenne. Audition auprès de la CERNA-ALLISTENE, 18 mars 2013. Consulté le 30/03/2015 sur : [http://cerna-ethics-allistene.org/digitalAssets/31/31320\\_Onlife.pdf](http://cerna-ethics-allistene.org/digitalAssets/31/31320_Onlife.pdf)
- INRS (2009) *Fiche pratique de sécurité ED 135. Préparation de commande guidée par reconnaissance vocale*.
- ISO (2009) ISO Guide 73:2009(fr): management du risque : vocabulaire. Consulté le 20/11/2014 sur <https://www.iso.org/obp/ui/fr/#iso:std:44651:fr>
- JARRIGE F. (2014) *Technocritiques : du refus des machines à la contestation des technosciences*. La Découverte.
- JONAS H. (1990) *Le principe responsabilité*. Le Cerf, Paris.
- KAPLAN F. (2014) Linguistic Capitalism and Algorithmic Mediation. *Representations* 127 (1): 57663.
- KLEINPETER E. (Dir). (2013) *L'humain augmenté*. CNRS Editions, coll. "Les essentiels d'Herms".
- LAFOURCADE M., LEBRUN A. (2014) Éthique et construction collaborative de données lexicales par des GWAPs (quelques leçons tirées de l'expérience JeuxDeMots). Actes journée d'étude "Éthique et TAL" de l'ATALA. Paris.
- LE MONDE (2015) Des robots au « Monde » pendant les élections départementales ? Consulté le 24/03/15 : <http://makingof.blog.lemonde.fr/2015/03/23/des-robots-au-monde-pendant-les-elections-departementales-oui-et-non/>
- LONGCAMP A. (2003) Etude comportementale et neuro-fonctionnelle des interactions perceptivo-motrices dans la perception visuelle de lettres. Notre manière d'écrire influence-t-elle notre manière de lire? Thèse U. Aix-Marseille II.
- MARIANI J. (2013) Pour une éthique de la Recherche en Sciences et Technologies de l'Information et de la Communication. Consulté le 30/03/2015 sur : [http://www.lina.univ-nantes.fr/IMG/pdf/COMETS\\_Mariani.pdf](http://www.lina.univ-nantes.fr/IMG/pdf/COMETS_Mariani.pdf)
- OMS (2006) *Classification statistique internationale des maladies et des problèmes de santé connexes/ International Statistical Classification of Diseases and Related Health Problems*. Organisation Mondiale de la Santé. Chapitre 5 : troubles mentaux et du comportement.
- OMS (2001) *Classification internationale du fonctionnement, du handicap et de la santé*. Organisation Mondiale de la Santé. Chapitre 1 : fonctions mentales. Consultée le 2/04/2015 sur : <http://apps.who.int/classifications/icfbrowser/>
- PAVEAU M.-A. (2014) « L'alternative quantitatif/qualitatif à l'épreuve des univers discursifs numériques », *Corela* [En ligne], HS-15 | 2014, mis en ligne le 15 octobre 2014, consulté le 17/04/2015 sur : <http://corela.revues.org/3598>
- RAWLS J. (1987) *Théorie de la justice*. Le Seuil, Paris.
- ROBOLAW Project (2014) *D6.2 Guidelines for Regulating Robotics*, consulté le 20/05/2015 sur <http://www.robolaw.eu/>
- SPARROW B., LIU J., WEGNER D.M. (2011) Google effects on memory: cognitive consequences of having information at our fingertips. *Science*, 333, 776-778
- STATAMATOS E. (2009) A survey of modern authorship attribution methods. *Journal of the American Society for Information Science and Technology*, 60(3), 538-556.
- STIEGLER B. (2015) S'augmenter ou se diminuer ? *Entretien BiTS S02E20 : les promesses du transhumanisme sont-elles réalisables ?* ARTE TV, magazine BiTS. Diffusé le 12 mars 2015. <http://creative.arte.tv/fr/bits-trans-human>
- TRIFFAUX J.M., MAURETTE J.L., DOZOT J.P., BERTRAND J. (2002) *Troubles psychiques liés aux greffes d'organes*. Editions Scientifiques et Médicales. Elsevier. Postprint consulté le 30/03/2015 sur : <http://hdl.handle.net/2268/80452>
- UNION EUROPEENNE (2006) Ligne directrice concernant la définition d'un risque potentiel grave pour la santé humaine ou animale ou pour l'environnement dans le cadre de l'article 33, paragraphes 1 et 2, de la directive 2001/82/CE. *Journal officiel de l'Union Européenne*, C133/6. 8.6.2006.
- WANDMACHER T., ANTOINE J.-Y., DEPARTE J.-P., POIRIER F. (2008) Sibylle, an assistive communication system adapting to the context and its user. *ACM Transactions on Accessible Computing*. 1(1). pp. 1-30.