

Le continuum entre « instrumentation » et « instrumentalisation » du sens

The continuum from « instrumentation » to « instrumentalization » of meaning

M. Caterina MANES GALLO

Laboratoire CSD – JE 23 85 TELANCO, France
maria-caterina.manes-gallo@u-bordeaux3.fr

Résumé. L'hypothèse proposée dans cet article est que l'étude des problèmes soulevés par la conception de systèmes informatiques plus « anthropocentrés » ne peut être isolée de l'analyse ni des retombées épistémologiques du travail théorique sous-jacent, ni des retombées cognitives et/ou psychosociologiques que la présence de ces dispositifs détermine chez les humains. Après un rapide tour d'horizon sur la naissance et le développement du secteur de recherche sur la communication Humain/Machine, sont mises en « dialogue » deux orientations de recherche différentes. Leur thème fédérateur concerne l'analyse, le recueil, l'accès et la conservation de corpora en langue naturelle à l'oral et à l'écrit. Cette réflexion polyphonique sur le « phénomène » des NTIC propose de repenser la « question du sens » et de ses variations.

Mots-clés. Cognition, représentation des connaissances, instrumentation du sens, instrumentalisation du sens.

Abstract. In this article a working hypothesis for the design of geared towards human-computer systems is proposed. According to this hypothesis, the problems brought up by the design of such systems have not to be studied separately from : firstly, the epistemological consequences of the theoretical work underlying their conception, and secondly, the cognitive and psychosociological effects of their presence on human's activities. A synthetic overview of the human/machine research field is outlined and two different trends are compared and contrasted. Their unifier theme concerns the analysis, the collection, the access and the preserving of oral and written natural language corpora. The purpose is to justify how thinking about the new information technologies in a polyphonic way, may help to rethink the « question of meaning » and its variations.

Keywords cognition, knowledge representation, meaning instrumentation, meaning instrumentalization.

1 Introduction

Au cours des trente dernières années, l'élargissement progressif des applications des Nouvelles Technologies de l'Information et de la Communication (NTIC) a contribué au développement de dispositifs d'interface plus adaptés à une *logique d'utilisation*. Ensuite, à partir des années 90, l'apparition d'Internet et la conséquente diversification des possibilités d'interaction humain/machine ont

rendu de plus en plus évident que la conception de dispositifs informatiques capables de *partager du sens* se trouve au carrefour entre avancée technologique et recherche fondamentale. D'où l'intérêt de faire état des nouvelles synergies (entre sciences humaines et sociales et sciences de l'ingénieur) qui reconfigurent la question du sens, notamment par rapport à son *instrumentation* et à son *instrumentalisation*.

Les recherches interdisciplinaires mises en dialogue dans cet article concernent la médiation du dispositif informatique dans l'analyse, le recueil, l'accès et la conservation de corpora en langue naturelle, à l'oral et à l'écrit. Un état des lieux sur les problèmes, les enjeux ainsi que les solutions existantes est esquissé.

L'objectif est de montrer, d'un côté, que la conception de dispositifs conviviaux pour l'accès à l'information doit inévitablement prendre en compte les retombées cognitives et/ou psycho-sociologiques générées par l'utilisation de ces mêmes dispositifs. D'un autre côté, il apparaît aussi que le travail théorique (et/ou méthodologique), sous-jacent à l'« instrumentation » du sens langagier, propose une conceptualisation de ce dernier qui peut intéresser la réflexion épistémologique en sciences humaines et sociales.

2 La question du sens : l'apport des sciences cognitives

À partir des années 70 et 80, la réflexion philosophique a contribué à réduire à de plus justes proportions à la fois les ambitions de l'Intelligence Artificielle (Dreyfus, 1979) et le monopole de la *conception du sens*, issue de la métaphore computationnelle (Searle, 1980 ; 1983). Simultanément on a assisté à la naissance du nouveau programme de recherche de Science Cognitive, ou des Sciences Cognitives. Cette double appellation renvoie d'un côté, à la détermination qualifiante de l'étude sur la *cognition-connaissance*, considérée(s) comme fondement(s) du processus d'attribution du sens, et de l'autre aux connotations différentes attribuées à cet objet, selon le regard porté par les disciplines engagées dans son étude (Ganascia, 1996).

Dès sa fondation, ce programme de recherche a été articulé sur la triade esprit - cerveau - ordinateur. Le but était de regrouper des chercheurs issus d'horizons disciplinaires parfois très éloignés mais tous impliqués, sous des angles différents, dans l'étude de la *connaissance* ou de la *cognition* conçue comme un objet scientifique. Il ne s'agissait pas de circonscrire un savoir, en visant une définition abstraite de son contenu. Il s'agissait, au contraire, de donner un fondement scientifique, à la fois formel et empirique, à un objet abstrait analysable à partir de perspectives très différentes. Ces perspectives concernent les sources individuelles de l'*objet connaissance* (système neuronal et/ou mental), ses origines sociales (la culture), son support (l'ordinateur), et enfin le vecteur principal qui en permet la circulation au sein du corps social (la langue naturelle).

Le caractère spécifique de l'objet *cognition-connaissance* est d'être finalisé à la résolution de problèmes nouveaux dont on ne connaît pas *a priori* la procédure de résolution (par exemple, répondre à une requête d'information, résumer un récit, gagner une partie d'échecs, faire un diagnostic,...). Cet objet fonde la capacité d'inférer ou de dériver de nouvelles expressions symboliques (ou données de connaissance) plus ou moins complexes, à partir de données déjà connues (mots, formules logiques, concepts d'objets ou d'événements, règles de savoir-faire, ...). D'où la centralité accordée à la *représentation des connaissances* comme fondement de l'*intelligence* (à la fois *naturelle* et *artificielle*) considérée comme le produit d'un calcul (Colombetti, 1985).

Les Sciences Cognitives ne couvrent pas l'ensemble des champs du savoir. Les sciences de la terre n'y sont pas impliquées. Au cours de sa première période de vie, approximativement jusqu'aux années 90, elle a qualifié un paradigme méta - théorique (dit de manipulation symbolique) commun à l'ensemble des chercheurs qui étudient des classes de phénomènes différents, mais qui impliquent tous, directement ou indirectement, la présence de l'humain soit comme objet (neurologie, psychologie, anthropologie, linguistique, ...), soit comme modèle (informatique, sciences de l'information, ...), soit encore comme abstraction (philosophie). L'hypothèse de fond étant que, indépendamment de leur nature, les prestations des systèmes *intelligents* (humain, voire animal, ou computationnel) peuvent être analysées et modélisées à partir de la définition de règles de calcul sur des expressions symboliques.

Ce paradigme théorique a été déterminant pour la définition de l'*hypothèse fonctionnaliste*, sous-jacente au développement de la psychologie des processus cognitifs. L'identification des processus cognitifs en termes de processus de traitement de l'information, en cautionnant la scientificité de l'étude sur l'activité mentale, l'affranchissait aussi de l'analyse de l'activité neuronale qui en permet l'émergence. Actuellement, après plus de vingt ans, l'essor incontestable des neurosciences cognitives a réduit l'importance heuristique du paradigme de manipulation symbolique. Ce dernier a été dépassé par l'émergence et la consolidation du paradigme connexionniste.

Cette *première révolution* des Sciences Cognitives s'est faite en partie au prix d'un délaissement de la problématique de la représentation des connaissances. Délaissement qui, au sein des études sur l'activité mentale, contribue à une (dangereuse) monopolisation de la notion de *cognition* en sens neuro-computationnel (cf. par exemple, l'argumentation développée par Guy Tiberghien (1999)).

Dans ce qui suit je restreindrai délibérément mon propos à l'apport du programme de Science Cognitive dans le cadre du paradigme dit de *manipulation symbolique*. Comme on le verra par la suite, c'est au sein de ce paradigme que se sont développées les recherches qui sont à l'origine de la distinction entre traitement des connaissances et gestion de l'information.

3 L'orientation des recherches : quelques repères

D'abord, l'apparition de la micro-informatique et la conséquente diffusion d'ordinateurs mono-poste à un public plus vaste que les seuls professionnels a contribué à diversifier les problèmes de conception. Notamment, la nécessité de concevoir des interfaces qui ne fassent pas porter à l'utilisateur tout le poids de l'interaction (Rouault, 1992). Ensuite, à partir des années 90, l'apparition du Web et la mise à disposition de technologies pour de nouveaux champs d'application, comme l'imagerie médicale ou la messagerie électronique, ont imposé la nécessité de repenser ce que signifie pour l'acteur humain « partager du sens » par l'intermédiaire d'un dispositif informatique. Ces deux événements ont impulsé deux orientations de recherche différentes.

Historiquement le problème du partage du sens de la part du dispositif s'est imposé assez vite aux sciences de l'ingénieur. La solution a été de développer des dispositifs de calcul autonomes, ou interfaces, qui idéalement doivent assurer la fonction d'intermédiaire entre l'humain et les logiciels de gestion de l'information (e.g. les moteurs de recherche sur le Web). En tant qu'intermédiaire, l'interface garantit la traduction des tâches, que veut réaliser l'humain, en un code intégrable

par ces logiciels. Les modes d'expression de ces tâches peuvent être multiples (graphiques, parole, gestuel, écrit). Le mode plus ancien et plus familier pour l'humain est la langue naturelle.

D'où le problème de définir ce que signifie pour un dispositif « comprendre le sens » construit par ce mode de « mise en forme » de l'information. Un mode de « mise en forme » qui pour l'interface correspond à des chaînes de caractères dépourvues de sens. Les recherches en communication Humain/Machine, visant à faire évoluer l' *intelligence linguistique* des modules d'interface, posent à nouveaux frais la nécessité de distinguer information et connaissance. Les apports des disciplines, ayant vocation à s'intéresser à l'objet *cognition* chez les humains, sont sollicités. Notamment la psychologie et les sciences du langage.

La crise économique ayant changé la donne, le consommateur devient plus exigeant par rapport au confort d'usage à la simplicité d'utilisation du produit final. Bref, il est plus sensible à l' « utilisabilité » d'un système. D'où l'importance accordée à son adéquation en regard des capacités cognitives, motrices et sensorielles de l'acteur humain (Calvary, 2002). Le tournant interdisciplinaire de la communication Humain/Machine rejoint les préoccupations des Sciences Cognitives avec lesquelles elle partage désormais l'habitude à réfléchir sur la fécondité des tentatives de transferts et d'adaptations des « savoirs » élaborés en sciences humaines.

Les « savoirs » (résultats expérimentaux, modèles théoriques, analyses empiriques) élaborés en psychologie et en sciences du langage se sont révélés particulièrement précieux pour la définition des connaissances nécessaires à la conception de dispositifs informatiques plus adaptés à l'acteur humain. L'effort pour produire ces « savoirs », s'appuyant sur le monopole exercé par la métaphore de l'activité mentale comme stratification de processus de traitement de l'information, a contribué largement à déterminer les orientations des recherches en psychologie générale ; à titre d'exemple, dans les années 70 et 80, les études sur la compréhension des récits ou sur le raisonnement quotidien (Johnson-Laird, 1983), ou plus récemment, la modélisation des processus cognitifs de base et des méta processus, sous-jacents à la capacité de sélectionner des informations pertinentes, à partir de bases de données informatisées (Dinet et Rouet, 2002) et l'incidence de l'âge, de la profession et de la culture sur les stratégies métacognitives mises en oeuvre dans l'activité de recherche d'information informatisée (Rouet et Tricot, 1998).

Dans le cadre des travaux sur le Traitement Automatique de la Langue Naturelle Ecrite (TALNE), s'est imposée l'idée que le mode langue naturelle loin d'être un moyen d'expression plus ou moins transparent, pouvant être abordé par le biais du tout automatique, devait au contraire être considéré comme un *vecteur de communication* modélisable de façon opératoire (Nicolle, 2003). L'effort pour rendre les dispositifs d'interface plus adaptés à l'opaque polysémie du discours langagier a abouti à la définition de modélisations opératoires, fondées sur des résultats théoriques élaborés en linguistique. La réalisation de dictionnaires automatisés des synonymes qui, à partir d'une approche *continuiste* du sens, prennent en compte la polysémie de chaque terme, selon ses co-textes d'emplois (Ploux et Victorri, 1998), ou les réalisations logicielles (Beust et Perlerin, 2003), basées sur la *Sémantique interprétative* de François Rastier (Rastier, 1991), et finalisées au repérage et à l'analyse automatique des « métaphores », constituent des exemples récents.

En revanche, le problème du partage du sens de la part de l'humain s'est imposé plus tardivement, comme conséquence de la diffusion du « phénomène NTIC ». L'application de ces technologies dans un nombre toujours croissant de secteurs d'activités (médecine, recherche d'information, communication

interpersonnelle, ...) projette en premier plan la question de la signification attribuée par les humains à la présence d'un « outil-partenaire » auquel ils sont à la fois associés et confrontés pour la réalisation de leurs objectifs. En d'autres termes, l'élargissement des applications de la numérisation de l'information fait émerger des questionnements nouveaux sur les pratiques mises en œuvre et les significations attribuées à ces dispositifs dans la phase de leur utilisation. Les enjeux théoriques ne sont pas les mêmes : il s'agit de conceptualiser les effets du « phénomène NTIC ».

Par ce glissement de la « question du sens », l'utilisateur devient un usager. L'accent n'est plus (ou pas seulement) sur les problèmes de conception finalisés à l'adaptation du système à son partenaire humain, mais sur les nouvelles modalités de communication que ces dispositifs déterminent entre les usagers eux-mêmes, ou les usagers et la société environnante (Rouault et Miège, 2003).

Une dimension différente de l'écart entre information et connaissance s'impose désormais. Une dimension qui nécessite aussi de nouveaux regards, en provenance des sciences de l'information et de la communication (Fondin, 2001) ou de la psycholinguistique (Vivier, 2001). Ce second aspect de la question du sens demeure d'autant plus actuel à cause des catégorisations globalisantes et non moins réductrices que les NTIC semblent inspirer. A titre d'exemple on peut citer la notion « d'interactivité ». Le flou sémantique qui entoure cette notion est proportionnel à la fréquence d'utilisation du terme correspondant, employé pour rendre compte de phénomènes communicationnels aussi éloignés que l'échange conversationnel et l'usage d'environnements multimédias, ou la consultation de page Web et la messagerie électronique, ou encore le transfert d'informations entre dispositifs en réseau et les expériences de « fiction participative » en salles de cinéma. D'où l'intérêt de définir la signification implicite qu'acquière les termes d'*interaction* et d'*interactivité* selon qu'ils soient appliqués à l'échange entre humains, à la communication Humain/Machine, ou à l'échange entre humains par l'intermédiaire d'un dispositif de transmission de l'information (Guéneau, 2005).

La conception de dispositifs « anthropocentrés » vise l'adéquation du système d'interface au processus de sémiotisation et/ou au fonctionnement cognitif de l'agent humain. En revanche, si l'attention est portée sur les effets de la diffusion du « phénomène NTIC », l'analyse sera focalisée sur comment l'humain *s'adapte* (e.g. les conditions de référence dans les échanges par messagerie électronique (Vivier, 2001)) - *ne s'adapte pas* (e.g. les différents types d'effets de redondance sur l'apprentissage à partir de documents multimédia (Le Bohec et Jamet, 2005)) - ou *détourne* l'intermédiation du dispositif informatique dans la communication avec un autre humain. L'analyse des caractéristiques langagières des courriers électroniques attestent, par exemple, la mise en œuvre de stratégies discursives hybrides qui ne sont complètement assimilables ni aux modalités d'expression immédiate, typiques de l'oral, ni aux canons plus normés de l'expression écrite accomplie. Elles se situent entre les deux. L'étude de deux corpora de mails, recueillis respectivement auprès d'une population d'universitaires (Panckhurst, 1999) et auprès du personnel d'un service informatique » (Cusin-Berche, 1999), a permis de mettre en évidence que ce processus d'hybridation affecte à la fois les stratégies énonciatives et la mise en page des textes électroniques finalisés à la communication interpersonnelle « instrumentée ».

Bien qu'elle prenne en compte les mêmes acteurs, l'orientation des recherches sur la communication Humain/Machine/Humain, reconfigure la « question du sens » selon une direction différente. Une direction qui vise à prendre en considération les retombées cognitives, psychosociologiques ou émotionnelles des activités réalisées grâce à l'intermédiation informatique. D'où l'intérêt de la

proposition de Jacques Anis, visant à distinguer la *médiatisation* informatique de la *médiation* informatique. L'utilisation des deux termes permet en effet de différencier la *fonction* et l'*effet* du dispositif. Il s'agit de marquer la différence entre sa fonction d'*intermédiaire* dans le transfert de l'information, de son effet comme *médiateur*. Un médiateur qui, en mode langue naturelle écrite, induit indirectement des modifications à la fois sur les formes de communication (e.g. les requêtes d'information formulées par mots clés) et sur les genres de discours de l'agent humain. Par exemple, dans l'espace virtuel et dématérialisé de la communication électronique instantanée (les « chats »), la nouvelle sémiolinguistique des représentations de l'Ego permet au locuteur de se conceptualiser de façon différente, à la fois face à lui-même et face à son interlocuteur (Anis, 2001).

Dans ce qui précède j'ai essayé d'explicitier pourquoi les deux dimensions de la conception des NTIC et de leur diffusion contribuent aussi à reconfigurer les synergies entre sciences humaines et sciences de l'ingénieur, notamment par rapport à l'approfondissement de la « question du sens » et de ses variations. Des synergies articulées à la fois sur l'« instrumentation » du sens et sur celle de son « instrumentalisation »¹.

Dans ce qui suit, l'articulation de cette double problématique sera illustrée à partir de la « mise en dialogue » de différentes recherches interdisciplinaires (dont une bonne partie sont présentées dans ce numéro spécial), renvoyant soit à la *médiatisation* informatique, soit à sa *médiation* dans l'accès, l'analyse ou le recueil de corpora en langue naturelle écrite ou orale.

4 De l'« instrumentation » du sens à son « instrumentalisation »

En s'inscrivant dans une optique de collaboration entre sciences humaines et sciences de l'ingénieur, Matthieu Quignard (2007), dans ce numéro spécial, focalise son attention sur les apports de deux dispositifs - DAMOCLÈS et DREW - dans le développement des recherches en psychologie cognitive. Notamment, des recherches focalisées sur l'étude des changements cognitifs des partenaires d'un dialogue argumentatif, dans une situation d'apprentissage du genre résolution de problèmes.

Au niveau théorique le problème est de réussir à « pister » les inférences et les échanges épistémiques, mis en œuvre au cours du débat. Ces inférences, orientées vers la compréhension du problème posé à partir de la construction de nouvelles connaissances, sont considérées d'un point de vue cognitif comme essentielles pour l'élaboration d'une solution de la part des apprenants. Les deux interfaces (le *chat* et un éditeur de graphes) qui assurent le recueil des données, contraignent d'une côté les opérations communicatives autorisées au cours de l'échange argumentatif (par exemple la rédaction écrite). Mais d'un autre côté, elles permettent aussi simultanément de repérer la dynamique argumentative à partir de laquelle les deux partenaires explorent *conjointement et de façon critique le problème posé*. Par rapport à la question du sens, le modèle de communication déterminé par l'outil informatique inaugure un terrain expérimental nouveau au sein duquel les dispositifs d'interface jouent un rôle central pour rendre compte du déroulement et de l'organisation de l'activité collaborative. L'importance de leur rôle se déploie à deux niveaux différents.

Au niveau méthodologique, ils fournissent la possibilité de recueillir de façon automatique, et donc de mesurer, deux types de données : verbales (le *chat*) et non-

¹ Je remercie Jean-Pierre Courtial pour m'avoir suggéré la dénomination de cette opposition.

verbales (les graphes) ; bien que, comme le souligne Matthieu Quignard, ces deux types de données restent difficilement comparables. Tandis que au niveau théorique, les données que ce terrain expérimental permet de construire contribuent à l'élaboration de nouvelles hypothèses ; notamment sur le rôle fondamental joué par la confrontation dialectique, pour sensibiliser l'apprenant aux données du problème à résoudre, à ses erreurs de compréhension et/ou aux limites de ses connaissances.

Mais comme l'indique Daniel Luzzati dans ce numéro spécial (Luzzati, 2007), il est important aussi de réfléchir sur les limites et les retombées épistémologiques de l'« instrumentation » du sens, notamment en sciences du langage. La question liminaire autour de laquelle s'organise sa contribution pose le problème du décalage entre la dimension éminemment dynamique du dialogue oral spontané et sa transformation en un objet statique lorsqu'il devient objet d'étude. Tout en soulignant l'utilité incontestable des outils d'archivage et d'indexation automatique, véritables « microscopes » pour l'étude des processus langagiers qui sous-tendent l'échange à l'oral, Daniel Luzzati nous met aussi en garde contre l'illusion de pouvoir rendre « complètement automatisable » tous les aspects de cette expérience langagière. L'illusion de réussir une fois pour toute à définir, simuler et commenter un processus aussi *insaisissable et fluctuant* que le dialogue oral spontané, dont le sens va bien au delà de la simple manipulation de signes. L'avantage inestimable des outils pour la numérisation et la transcription automatique des corpora est de rendre possible la circulation et/ou le partage des données et la confrontation des analyses et des approches théoriques qui les déterminent. Mais cet avantage risque de se transformer en dangereux faux-semblant, si on oublie que ces outils traitent seulement des descriptions de processus réels. Sans compter que la mise à disposition de ces outils, au sein de la communauté pluridisciplinaire qui s'y intéresse, crée inévitablement un problème de compatibilité et de diffusion des logiciels de traitement et de codage de l'information.

Cependant, l'apport des recherches finalisées à l'« instrumentation » du sens n'est pas seulement de fournir (à terme) des outils qui rendent *visibles* des phénomènes jusque là non observables de façon systématique. Plus précieuse encore est l'analyse des pratiques que ces outils génèrent.

D'après cette perspective, dans ce numéro spécial, Franck Cormerais (2007), dans le cadre du projet collectif du *Codex Atlanticus*, s'interroge sur les présupposés heuristiques mobilisés pour résoudre ce qu'il définit *l'identité trouble* du livre électronique. Ces présupposés sont définis par l'auteur à partir de deux hypothèses, visant à identifier d'un côté l'origine du sens et de l'autre la possibilité d'existence du texte. Pour ce qui concerne le premier point, la non immédiateté du sens renvoie l'identification de son origine à l'« instrumentation » par laquelle on a accès à un certain contenu (e.g. le langage ou l'écriture). Tandis que la possibilité d'existence du texte suppose toujours, et dépend donc, des instruments mobilisés pour l'activité éditoriale qui le produit. D'où la possibilité, offerte par la numérisation de l'information, d'envisager deux types complémentaires de *corpus* textuel : le corpus *ex-post* (i.e. textes objet de fouille) et le corpus *ex-ante* (i.e. texte produit par l'édition électronique). Cette diversification typologique est à l'origine de ce que Franck Cormerais définit *l'hyperlecture appropriative*, dont les caractéristiques interdisent de l'assimiler à une simple recherche d'information. Dans l'architecture du *Codex Atlanticus* cette lecture appropriative, ou écriture de sa propre lecture, est *mise en action* par des formats éditoriaux différents, qui constituent autant d'instruments de lecture-écriture. Des instruments qui transforment la complexe activité de compréhension-explication en une activité de coopération avec le livre, et dont la mise en œuvre dépend des choix du lecteur, i.e. des chemins de

consultation/interprétation choisis par le lecteur à partir du « trajectif » fourni par le support.

Le dénominateur commun de ces trois recherches est de mettre en évidence pourquoi la conception de dispositifs pour l'« instrumentation » du sens ne peut pas être disjointe de l'analyse à la fois de leurs retombées cognitives sur le déploiement des activités humaines, et des pratiques inédites mises en œuvre dans la réalisation de ces mêmes activités. D'après cette nouvelle orientation du travail interdisciplinaire, l'analyse des « effets de sens » déclenchés par ces « outils-partenaires » chez l'utilisateur et/ou chez le chercheur, contribue à identifier les modes de conceptualisation du sens, induits par son « instrumentation ».

Pour ce qui concerne l'utilisateur-usager, la « médiation » du dispositif ne contribue pas seulement à l'émergence de nouvelles pratiques, elle favorise aussi l'élaboration plus ou moins consciente de nouveaux imaginaires. On peut faire l'hypothèse que les *imaginaires globalisants* induits par l'irruption d'internet dans notre univers quotidien (Guichard, 2001) ne sont pas indépendants des modalités d'accès à l'information imposées par cet outil. Actuellement, le rôle monopolisateur accordé au syntagme nominal force l'utilisateur à effectuer son activité de recherche indépendamment d'un réel échange avec l'interface. La progression de l'activité de recherche ne permet pas une *négociation conversationnelle* (satisfaisante) avec l'interface. Si l'utilisateur est insatisfait il n'a qu'à reformuler sa requête : i.e. taper une autre séquence de mots clés.

On pourrait souhaiter dans un futur (peut-être pas immédiat) une évolution du niveau de *convivialité* de l'interface qui en fasse un partenaire plus interactif, un partenaire qui *ait l'air* de participer à la progression du discours avec l'utilisateur ou qui ressemble plus à un *dialogueur*. La difficulté (voire l'impossibilité) de déterminer à l'avance les objectifs (souvent confus) et les compétences des utilisateurs, couplée avec le souci d'une évolution *conviviale* de l'outil informatique, suggèrent de libéraliser le format du mode d'interrogation sous forme propositionnelle et pas exclusivement nominale. Les avantages de cette libéralisation en termes de satisfaction de l'utilisateur sont évidents. Du calcul du sens à partir d'une expression nominale on pourrait passer à un calcul du sens à partir d'une expression propositionnelle. Ce qui revient à dire que pour élargir la convivialité de l'interface, il est nécessaire d'abord de définir une modélisation des connaissances linguistiques adaptées au calcul du sens d'une requête formulée sous forme propositionnelle.

La réalisation d'un *dialogueur* informatique à *coopération propositionnelle* passe inévitablement par l'intégration dans l'interface d'une connaissance sur d'autres formes langagières. Notamment les constructions prédicatives. Pour l'interface il s'agit de repérer, à partir de la formulation linguistique, quels sont les buts implicites sous-jacents à une requête. Par exemple si elle vise l'obtention d'une information déclarative (e.g. les caractéristiques d'un objet ou ses modalités de fonctionnement) ou d'une information procédurale (e.g. comment réaliser une démarche, les procédures à suivre pour atteindre un certain objectif) (Chanet, 2002).

Dans la partie qui suit, je montre comment les contraintes qui pèsent sur le travail théorique de *représentation des connaissances* à fort ancrage langagier, dans la conception d'interfaces « conviviales », contribuent aussi à focaliser l'attention du chercheur sur de « nouveaux observables ». Notamment, des observables qui puissent rendre compte du sens évoqué par la mise en fonctionnement des constructions prédicatives dans un discours.

5 Quelques retombées épistémologiques de l'« instrumentation du sens »

La nécessité d'une libéralisation du format d'interrogation imposé à l'utilisateur est d'autant plus cruciale dans l'accès informatisé à de gros corpus textuels, dans le cadre de la Recherche d'Information en plein texte. Le rôle d'intermédiaire de l'interface est double. D'un côté elle doit repérer les buts implicites sous-jacents à la requête ; d'un autre côté elle doit aussi trouver des données textuelles qui satisfassent au moins en partie ces objectifs. L'élément qui assure ces deux fonctionnalités est la *représentation des connaissances*. C'est-à-dire la définition de modèles de connaissances linguistiques adaptés et/ou adaptables à l'« instrumentation » du sens ; notamment le sens actualisé par la mise en fonctionnement de la langue naturelle dans un discours. Par la définition de ces données métalinguistiques, on vise à améliorer la fonction d'*intermédiaire* de l'interface, en prenant en compte le décalage « cognitif » par rapport à son partenaire humain, pour lequel le mode langagier constitue un mode de communication très familier (bien que non nécessairement conscient) pour construire, co-construire et partager du sens. En revanche pour l'interface, ce sont ces modèles qui guident à la fois l'extraction du sens implicite des requêtes en entrée et l'indexation automatique du contenu des documents textuels.

Les données métalinguistiques représentées dans le système concernent les différents niveaux de fonctionnement (morpho-syntaxique, sémantique, pragmatique) des formes langagières dans un discours. Elles déterminent et circonscrivent l'étendue de(s) sens(s) calculables par le dispositif informatique, lui permettant, par exemple, d'interpréter l'inscription sémiotique des tâches ou des objectifs sous-jacents aux requêtes. Dans le cadre de la Recherche d'Information en plein texte, les connaissances représentées dans l'interface sont à l'origine du degré d'efficacité de son processus d'appariement entre la requête de l'utilisateur et des portions de texte pertinentes.

Cette théorisation à finalité opérationnelle, contrainte par les logiques de la formalisation algorithmique, comporte des retombées épistémologiques intéressantes. Etant fondée sur la nécessité de délimiter de façon non ambiguë le niveau de granularité des phénomènes langagiers pris en compte (Rouault et Manes Gallo, 2003), ce type de théorisation constitue aussi une *heuristique* féconde afin d'approfondir l'analyse du sens langagier, en évitant le recours circulaire soit à sa dimension référentielle soit à la dimension cognitivo-mentale qui en permet l'appréhension de la part de l'acteur humain (Manes Gallo, 2001). Dans ce cadre, les contraintes méthodologiques de l'« instrumentation du sens » orientent le travail théorique vers une plus grande explicitation des relations entre la valence re/présentationnelle et la valence communicationnelle de la mise en fonctionnement de la langue naturelle dans un discours.

La fonction de la langue naturelle est en partie celle de *re/présenter* la réalité phénoménale concrète ou des réalités imaginaires ou abstraites que l'acteur humain peut conceptualiser ou dont il peut avoir expérience. Mais cette *valence re/présentationnelle* est *figurée*, elle ne correspond pas à l'attestation d'une signification linguistique propre. Toute unité linguistique actualisée dans un discours à la fois *convoque* pour son fonctionnement d'autres unités et, en synergie avec ces dernières, la même unité *évoque* certains effets de sens (Victorri, 1999 ; 2003). D'où la dimension dynamique de la construction du sens par le langage et son opacité polysémique.

Dans ce cadre, la langue naturelle est un système de *re/présentation* autonome à la fois du type de réalité extralinguistique qu'il permet d'*évoquer* et de l'activité

mentale du sujet qui le met en œuvre à la fois en compréhension et en production. C'est dans le creuset de cette indépendance du sens construit par le discours, que les contraintes liées à l'« instrumentation » du sens deviennent une méthodologie précieuse pour approfondir, par exemple, quels sont les facteurs co-textuels qui contribuent à la polysémie d'une construction prédicative et/ou quelle est la fonction de cette forme, notamment par rapport à la progression du discours (Rouault et Manes Gallo, 2003).

Travailler dans le cadre de la dissymétrie cognitive entre l'humain et la machine présuppose un travail de modélisation qui force le chercheur à se décentrer le plus possible de sa familiarité avec la langue naturelle. Se décentrer notamment de « l'évidence » de la variabilité des réalités extralinguistiques que les utilisations différentes d'une même forme permettent d'évoquer. Par exemple, l'emploi transitif du verbe « prendre » renvoie souvent à l'idée de « saisir quelque chose » (*prendre une cigarette, prendre un studio en ville*), tandis que l'emploi intransitif peut évoquer l'idée soit de « intervenir en faveur de quelqu'un ou de quelque chose » (*prendre parti, prendre position*), soit de « atteindre un certain effet » (*le feu prend, la mayonnaise prend, le vaccin prend*), soit encore de « actualiser l'existence et/ou la validité de quelque chose » (*les décisions européennes prennent effet par cette notification*). L'enjeu est ici de reconstruire en quoi consiste cette « évidence » en faisant le plus possible abstraction des connaissances encyclopédiques qui contribuent à appréhender cette variabilité. Il s'agit de mobiliser des objets théoriques qui permettent de rendre compte des glissements de sens mais dans les termes d'un ensemble limité de règles qui soient généralisables et traduisibles en un algorithme. La réflexion théorique est donc articulée sur deux contraintes opposées qui fixent le niveau de granularité auquel étudier les phénomènes langagiers. La reconnaissance automatique de la polysémie d'un terme dépend de la prise en compte de la séquence dans laquelle il figure, notamment l'énoncé élémentaire, hors connecteurs. Cette unité du discours permet de repérer les facteurs co-textuels qui contribuent aux glissements de sens d'un même terme. Dans le cas des verbes, ces facteurs co-textuels correspondent, par exemple, au rang du ou des compléments régis.

Focaliser la réflexion théorique sur les facteurs co-textuels de *mise en fonctionnement* discursive de la langue naturelle, et qui contribuent aux glissements de sens d'une même forme, implique aussi de se doter d'un modèle métalinguistique qui permette une représentation homogène de cette polysémie.

Le schéma métalinguistique, sous-jacent au modèle de sémantique énonciative exposé dans la partie suivante, est une adaptation de la notion de *lexis* élaboré par Antoine Culioli (1990). L'adoption de ce schéma métalinguistique permet de reconduire les glissements de sens des prédications verbales à quatre facteurs principaux : 1) le schéma syntaxique qui actualise le prédicat dans le discours, 2) le *type de procès* auquel renvoie la construction prédicative, selon le schéma syntaxique actualisé, 3) les caractéristiques « sémantiques » des compléments régis, 4) les opérations d'énonciation qui accompagnent la construction prédicative (e.g. les déterminants qui introduisent les actants qui encadrent le verbe, les modalités associées au verbe, la temporalité, ...).

6 Un exemple

Le modèle de sémantique énonciative, que je présente ici synthétiquement, s'inscrit dans une problématique de TALNE. Il a été défini comme composante d'un système de *représentation des connaissances* finalisé au traitement pragmatique. Notamment, un système qui permette d'extraire et de reconnaître à partir des

données textuelles à la fois *ce dont on parle* dans un discours et *comment on en parle* (Berrendonner et Rouault, 1991). Le modèle de sémantique énonciative prend en compte la reconnaissance du sens implicite d'énoncés élémentaires (hors connecteurs). Les connaissances représentées visent à permettre au système de reconnaître quelle est la direction de prédication des constructions verbales, afin de déterminer comment s'articule la progression du discours. C'est-à-dire quels sont les objets de discours qui sont affectés de nouvelles propriétés, au fur et à mesure que le discours avance. Dans ce cadre il s'agit de repérer quelle est la direction de prédication d'un verbe. Notamment, si la prédication induit l'attribution implicite d'une nouvelle propriété à son actant de droite (i.e. l'argument régi), comme dans *après un long combat l'armée prend la ville la ville est prise*, ou bien si elle associe de façon explicite une nouvelle propriété à l'actant de gauche, comme dans *le mandat prend fin le 1^{er} janvier*.

La direction de prédication des constructions verbales est approfondie en référence aux *types de procès* lexical auxquels elles renvoient et aux facteurs co-textuels qui transforment ce dernier en induisant son glissement vers un autre *type de procès*. Dans les deux exemples précédents, « prendre » dénote dans le premier cas une situation dynamique qui aboutit à un accomplissement qui correspond à l'attribution implicite à l'argument régi d'une nouvelle propriété. Tandis que dans le second cas, la locution « prendre fin », qui implique un emploi intransitif, contribue à la transformation du processus en un état, en partie aussi à cause de l'occurrence du déverbal en position de sujet grammatical.

Très synthétiquement, le « contrôle » de la *polysémie des verbes* actualisés dans un discours, se fait à partir de données de connaissance, concernant les deux facteurs suivants : 1) le *type de procès* lexical et 2) les règles de transformation du *type de procès* lexical.

Par rapport au premier point, le *type de procès* lexical auquel renvoie le verbe est déterminé en référence au schéma syntaxique de base C0 – V – C1 pour les verbes transitifs et C0 – V pour les verbes intransitifs. Une prédication peut renvoyer alternativement à une situation « statique vs. dynamique » ou à une situation « dynamique évoquant un accomplissement vs. n'évoquant pas d'accomplissement » (Gallo et Rouault, 1992).

Pour ce qui concerne le second point, les transformations du *type de procès* lexical peuvent être induites alternativement ou simultanément par :

- les schémas syntaxiques alternatifs associés au même lexème verbal (e.g. *manquer X* vs. *manquer de X*, *prêter quelque chose* vs. *prêter quelque chose à quelqu'un*) ;
- les caractéristiques sémantiques des arguments qui occupent les différentes places syntaxiques (e.g. le cas de *l'instrument* qui concerne le sujet grammatical et le cas des processus *inchoatifs* qui dépendent souvent des caractéristiques de l'argument régi). Par exemple, l'emploi transitif du verbe « prendre » renvoie souvent à une situation dynamique susceptible d'un accomplissement (*prendre une cigarette*, *prendre un studio en ville*), mais « l'imperméabilité » de l'argument visé peut actualiser le caractère inchoatif de la situation dynamique évoquée (*prendre la rue à gauche*), soit la transformer en situation stative (*les décisions européennes prennent effet par cette notification*) (Manes Gallo, 2003).
- les opérations d'énonciation qui accompagnent l'actualisation de la construction prédicative. Notamment, les déterminants associés aux arguments qui encadrent le verbe dans le discours (e.g. *les états membres prennent toute mesure générale ou particulière propre à assurer* vs. *les états membre prennent toutes les dispositions utiles pour...*) (Galatanu et Manes Gallo, 2007), les compléments non régis qui lui sont associés. Par exemple, le complément d'agent associé à l'état « prendre

fin » le transforme en un état résultat du processus sans accomplissement, placé comme sujet grammatical et lexicalisé par une nominalisation : *la participation de l'état membre concerné prend fin par démission volontaire, l'existence de la commission temporaire d'enquête prend fin par le dépôt de son rapport*.

Globalement, d'après le modèle, l'approfondissement du jeu entre ces trois facteurs co-textuels est essentiel si l'on veut rendre compte de l'articulation entre *comment on parle* d'un objet de discours et ce qu'on communique à son sujet (Manes Gallo, 2003).

7 De la « question du sens » et de ses variations : pour une conclusion provisoire

Pour terminer. Dans ce qui précède, la « mise en dialogue » de contributions focalisées soit sur l'« instrumentation » soit sur l'« instrumentalisation » du sens ne vise pas à mettre en concurrence deux conceptions du sens. Il s'agit au contraire de maintenir en tension ces deux dimensions de la « question du sens » inaugurées par les NTIC. L'idée de fond étant que l'« instrumentation » du sens, visant à rendre les dispositifs plus « utilisables » par l'acteur humain, contribue inévitablement à l'émergence d'une « instrumentalisation » du sens, par l'utilisateur de ces mêmes dispositifs. Un sens qui correspond à la valeur attribuée par l'humain soit au dispositif-partenaire, soit aux pratiques inédites qu'il contribue à installer ou à rendre possibles. En quelque sorte c'est par leur plus grande « utilisabilité » que ces dispositifs concourent à « instrumentaliser » la signification des activités humaines qu'ils contribuent à réaliser.

Mais d'un autre côté, c'est l'« instrumentation du sens », induite par la conception des NTIC, qui alimente aussi de nouvelles synergies entre sciences humaines et sciences de l'ingénieur. Notamment, la possibilité d'informatiser l'accès ou l'indexation de corpora textuels invite à réfléchir sur les apports épistémologiques de cette nouvelle rationalité qui identifie le sens à une *logique du calcul* (Bachimont, 2000). Pour les études sur la communication langagière (écrite ou orale) entre humains, l'intérêt du travail d'analyse inauguré par l'« instrumentation » du sens est de focaliser l'attention du chercheur surtout sur comment la fonction communicationnelle d'une production langagière est ancrée à sa dimension *re/présentationnelle*.

8 Références bibliographiques

- Anis, J. (1999). *Internet, Communication et Langue française*. Hermès Science.
- Anis, J. (2001). Approche sémiolinguistique des représentations de l'ego dans la Communication Médinée par Ordinateur. *Langages*, n°144, 20-38.
- Bachimont, B. (2000). Connaissances et support d'inscription : entre raison graphique et raison computationnelle. *7ème Ecole d'Eté de L'Association pour la Recherche en Science Cognitive (ARCo)*, Vol. 1, 1-25.
- Berrendonner, A., Rouault, J. (1991). Sémantique des objets et calcul des noms. In *Actes de KMET'91*, Nice.
- Beust, P., Perlerin, V. (2003). Pour une instrumentation informatique du sens. In *Variation, construction et instrumentation du sens*, Siksou M. (Ed.), Hermès, Paris, 197-227.

- Calvary, G. (2002). Ingénierie de l'interaction homme-machine : retrospective et perspectives. In *Interaction Homme-Machine et recherche d'information*, Paganelli C. (Ed.), Hermès, 19-63.
- Chanet, C. (2002). Interpréter les buts de l'utilisateur dans le dialogue homme-machine écrit : vers un modèle linguistique. In *Interaction Homme-Machine et recherche d'information*, Paganelli C. (Ed.), Hermès, 19-63.
- Colombetti, M. (1985). *Le idee dell'Intelligenza Artificiale*. Arnoldo Mondadori Editore coll. Intelligenza Artificiale, Milano.
- Cormerais, F. (2007). « La lecture appropriative » - Un prototype de livre électronique universitaire. *Revue d'Interaction Homme-Machine (RIHM)*, volume 8, n. 2, pp. 37-49.
- Culioli, A. (1990). *Pour une linguistique de l'énonciation, Tome1*. Ophrys.
- Cusin-Berche, F. (1999). Courriel et genres discursifs. In *Internet et communication en langue française*, Anis J. (Ed.), Hermès, Paris, 31-54.
- Dinet, J., Rouet, J-F. (2002). La recherche d'information : processus cognitifs, facteurs de difficultés et dimension de l'expertise. In *Interaction Homme-Machine et recherche d'information*, Paganelli C. (Ed.), Hermès, 134-161.
- Dreyfus, M. (1979). *What computers can't do : The limits of artificial intelligence*. Harper & Row, Publishers, Inc. New York.
- Fondin, H. (2001). La science de l'information : posture épistémologique et spécificité disciplinaire. *Documentaliste – Sciences de l'information*, vol. 38, n°2, 112-122.
- Galatanu, O., Manes Gallo, M.C. (2007). Moi et « Autrui » : l'élaboration discursive des identités. In *Actes du Colloque International L'Autre, création et médiation*, Université de Nantes - 2006 (sous presse).
- Gallo, M.C., Rouault, J. (1992). Schéma d'action et types de procès. *Intellectica*, n° 1/2, 105-127.
- Ganascia, J.G. (1996). *Les sciences cognitives*. Flammarion, Paris.
- Guéneau, C. (2005). L'interactivité : une définition introuvable. *Communication & langages*, n°145, 117-129.
- Guichard, E. (dir.) (2001). *Comprendre les usages de l'Internet*. Editions Ens Rue D'Ulm.
- Johnson-Laird, P.N. (1983). *Mental Models*. Harvard University Press, Cambridge, Massachusetts.
- Le Bohec, O., Jamet, E., 2005. Les effets de redondance dans l'apprentissage à partir de documents multimédia. *Le Travail Humain*, tome 68, n°2, 97-124.
- Luzzati, D. (2007). Le dialogue oral spontané : quels objets pour quels corpora. *Revue d'Interaction Homme-Machine (RIHM)*, volume 8, n. 2, pp. 23-36.
- Manes Gallo, M.C. (2001). Communication humain/machine et recherche d'information : le « partage du sens » en représentation des connaissances. *Mémoire d'Habilitation à Diriger des Recherches*, Université Stendhal, Grenoble III.
- Manes Gallo, M.C. (2003). Prédication verbale et bimodalité écrit-oral. In *Variation, construction et instrumentation du sens*, Siksou M. (Ed.), Hermès, Paris, 123-158.

- Nicolle, A. (2003). Les interactions langagières personnes-machines. In *Variation, construction et instrumentation du sens*, Siksou M. (Ed), Hermès, Paris, 251-286.
- Panckhurst, R. (1999). Analyse linguistique assistée par ordinateur du courriel. In *Internet et communication en langue française*, Anis J. (Ed.), Hermès, Paris, 55-70.
- Ploux, S., Victorri, B. (1998). Construction d'espaces sémantiques à l'aide de dictionnaire de synonymes. *Traitement Automatique de la Langue*, vol. 39, n°1, 161-182.
- Quignard, M. (2007). Instrumenter l'argumentation - Conception de dispositifs de communication médiée par ordinateurs pour la modélisation cognitive de l'argumentation. *Revue d'Interaction Homme-Machine (RIHM)*, volume 8, n. 2, pp. 1-21.
- Rastier, F. (1991). *Sémantique et recherches cognitives*. PUF, Paris.
- Rouault, J., Manes Gallo, M.C. (2003). *Intelligence linguistique : le calcul des énoncés élémentaires*, Hermès, Paris.
- Rouault, J., Miège, B. (2003). Approches et fondements. In *Communication personnes systèmes informationnels*, Miège B. (Ed.), Hermès-Lavoisier, 21-40.
- Rouault, J. (1992). *Interfaces et systèmes d'information*. Cahiers du CRISS, n° 20.
- Rouet, J-F., Tricot, A. (1998). Chercher de l'information dans un hypertexte : vers un modèle des processus cognitifs. In *Les hypermédias : Approches cognitives et ergonomiques*, Tricot A., Rouet J-F., (Eds.), Hermès, 57-93.
- Searle, R.J. (1980). Minds, Brains and Programs. *Behavioral and brain sciences*, vol.3, 417-457.
- Searle, R.J. (1983). *Intentionality : an essay in the Philosophy of Mind*. Cambridge University Press, Cambridge.
- Tiberghien, G. (1999). La psychologie cognitive survivra-t-elle aux sciences cognitives ? *Psychologie Française*, tome 44, n°3, 265-285.
- Victorri, B. (1999). Le sens Grammatical. *Langages*, n° 136, p.85-105.
- Victorri, B. (2003). Langage et Géométrie : L'expression langagière des relations spatiales. *Revue de Synthèse*, Tome 124, 5° série, 119-139.
- Vivier, J. (2001). Introduction : La psycholinguistique au secours de l'informatique. In Vivier J. (Ed.), *numéro spécial « Psycholinguistique et intelligence artificielle »*, *Langages*, 144, 3-19.