

# De la compréhension de textes à l'intelligence artificielle

## Une interview avec Guy Denhière

Christian Vandendorpe

Numéro 57, mars 1985

URI : <https://id.erudit.org/iderudit/47260ac>

[Aller au sommaire du numéro](#)

Éditeur(s)

Les Publications Québec français

ISSN

0316-2052 (imprimé)

1923-5119 (numérique)

[Découvrir la revue](#)

Citer cet article

Vandendorpe, C. (1985). De la compréhension de textes à l'intelligence artificielle : une interview avec Guy Denhière. *Québec français*, (57), 48–49.

# De la compréhension de textes à l'intelligence artificielle

Une interview avec

**GUY DENHIÈRE**

• Depuis quand la compréhension de textes s'est-elle constituée en tant que champ d'étude ?

— C'est un domaine qui est né dans les années 70, et qui résulte d'un triple courant de recherches : psychologique, informatique et linguistique. La linguistique a ouvert la voie par ses travaux sur le texte, le récit en particulier, avec l'école structuraliste, qui a repris et développé les travaux de Propp. À partir de ce moment, l'unité d'analyse ultime n'était plus la phrase, mais le texte, le discours en général. La 2<sup>e</sup> partie de ce courant est formée par les recherches en intelligence artificielle. Après l'échec des tentatives de traduction automatique qui voulaient passer directement d'une langue dans une autre langue, on s'est rendu compte qu'il y avait nécessité de passer par une représentation interne. À la limite, cette dernière peut être ce que Schank, par exemple, a appelé une *base conceptuelle indépendante de la langue*, pour aboutir, après un certain nombre d'opérations, à une autre formulation qui peut être aussi bien une traduction qu'une paraphrase. Donc l'apport essentiel, ici, est dû à des gens du monde informatique qui ont voulu traiter le langage naturel et qui ont montré très rapidement que, d'une part, il fallait accorder une place importante à la sémantique par rapport à la syntaxe et que, d'autre part, il fallait tenir compte des connaissances des individus. Le troisième aspect, c'est l'apport de la psychologie cognitive qui étudie la représentation de la signification dans la mémoire humaine, l'organisation des connaissances et la manière dont ces connaissances interviennent dans la compréhension (lecture et audition), la mémorisation et l'acquisition d'informations nouvelles.

• Est-ce que la représentation qu'on se fait du rôle de la mémoire dans les activités cognitives s'est modifiée au cours des dernières années ?

— Effectivement, il y a eu une évolution très nette et c'est, je crois, un aspect essentiel de la recherche au cours des dix dernières années. Il faut partir d'un constat : sans mémoire nous ne pourrions ni parler, ni nous comprendre. À chaque fois qu'on ré-entend une configuration déterminée de sons ou qu'on voit à nouveau un pattern de signes, on le reconnaît. Reconnaître, c'est comparer ces sons ou ces signes à une image intérieure qu'on a dans la tête et décider si, oui ou non, c'est la même chose que ce qu'on a déjà vu antérieurement. Donc, toute reconnaissance est basée sur un acte de mémoire. Mais une autre manifestation évidente de la mémoire, c'est que lorsqu'on parle, au moment où je dis le mot « X » et que vous le traitez, que vous le recevez, j'ai déjà avancé dans le discours. Donc, pour me comprendre, vous êtes obligé de garder la trace de ce que je viens de dire. C'est-à-dire que lorsque nous conversons ou quand nous lisons, nous avons cette tâche importante de construire de la signification pas à pas, tout en la rattachant à ce qui a été dit ou lu antérieurement. C'est-à-dire que nous avons deux choses essentielles à faire : garder une trace à court terme de ce qui vient juste d'être dit et, en même temps, faire des liens, mettre en relation ; savoir s'il s'agit toujours du même personnage, de la même idée, par exemple. Donc on a des trous à combler dans le discours. Si je prends l'exemple classique : « À midi je suis allé au restaurant, j'ai mangé une lasagne », vous comprenez ces énoncés parfaitement. Je n'ai pas besoin de vous expliciter toute la séquence des actions qui se déroulent normalement quand on se rend au restaurant (entrer, s'asseoir, lire le menu, commander, etc). Votre connaissance de cette situation vous a permis de construire les liens manquants entre les deux énoncés : aller au restaurant, manger de la lasagne.

• Le récit semble avoir suscité plus de travaux sur les mécanismes de compréhension que d'autres types de texte. Pourquoi ?

— Avec le récit, on a peut-être affaire à quelque chose de particulier. À savoir que le récit fait partie de notre patrimoine culturel en tant que genre littéraire. À peu près toutes les sociétés que l'on connaît, à toutes les époques, ont raconté des récits, et on a pu mettre en évidence, pour un grand nombre de cultures, une forme canonique stéréotypée de récit. En général, un récit a pour caractéristique de commencer par poser un personnage ou des personnages et un cadre spatio-temporel. Ensuite d'indiquer, par exemple, l'état du personnage principal : soit un manque ou un désir. Et, pour que l'histoire ne soit pas trop triviale, il faut qu'il y ait une difficulté, ce qu'on appelle une complication. Donc le récit va nous mettre en scène ce personnage, les actions qu'il tente pour parvenir à ses fins, l'obstacle qui se dresse ou l'événement qui retarde et enfin la façon dont va se résoudre l'épisode. Tout cela pour dire que certains auteurs considèrent qu'il y a une structure propre au récit. Le problème qui se pose au point de vue de la compréhension du récit, c'est de savoir si le jeune enfant, par le fait qu'on lui raconte des histoires, apprend progressivement à interioriser cette forme prototypique, ce schéma canonique du récit et si l'acquisition de ce schéma joue un rôle à son tour dans la compréhension de nouvelles histoires. On retrouve là une similitude avec une piste de travail en intelligence artificielle : une structure avec des cases vides et des variables qui viennent remplir ces cases. Le processus de compréhension impliquerait l'utilisation de ce schéma de récit : au fur et à mesure de la lecture ou de l'audition, l'individu organiserait l'information sémantique en fonction de ce schéma hiérarchique et, ultérieurement, le même schéma servirait à récupérer l'information stockée en mémoire.



• **Comment, dans votre laboratoire, conçoit-on l'acte de lire ?**

— À peu près tous les auteurs sont d'accord pour dire que dans la lecture il y a intervention de deux familles de processus, de deux grands types d'activités qui s'exercent simultanément. D'une part, on a cette activité qu'on a appelée le déchiffrement, ou le décodage, qui consiste à percevoir des signes dans une succession de traits horizontaux, verticaux ou autres et à leur donner un sens. Ça peut commencer, par exemple, avec les lettres, aller à la syllabe, au mot, au groupe de mots ; c'est une activité qu'on appelle « de bas en haut », ou « dirigée par les données ». Tout aussi important, un autre type d'activité s'exerce simultanément ; c'est l'activité qui consiste à guider, à contrôler cette activité de déchiffrement à l'aide d'unités conceptuelles plus importantes : on l'appelle « dirigée par les concepts ». Par exemple si l'on traite d'un thème comme celui du Pôle Nord et qu'en lisant je trouve un mot qui commence par *ban*, je vais peut-être anticiper et dire *banquise*. L'acte de déchiffrement, le processus dirigé par les données, va me permettre

**L'ordinateur veillera à ce qu'il n'y ait pas trop d'informations nouvelles d'un coup...**

de confirmer ou d'infirmer cette hypothèse. Donc, vous avez, d'une part, ce qu'on pourrait appeler déchiffrement, identification et traitement des stimulus élémentaires et, en même temps, intervention de processus plus globaux qui permettent des anticipations. Ce qui est important, c'est que ces deux types d'activités doivent être coordonnées. Pour un lecteur expert ou très expert, il est très probable que le processus du type « dirigé par concepts » (ou du haut vers le bas) prend une part de plus en plus importante et que les activités élémentaires de déchiffrement sont automatisées. Il me semble que l'on arrive ainsi à une conception générale selon laquelle tout individu dispose d'une quantité déterminée de ressources cognitives, qu'il alloue une part de ces ressources à l'un ou l'autre de ces processus (du bas vers le haut ou du haut vers le bas) et que, suivant le degré d'expertise ou de familiarité avec ce texte lu, la répartition entre ces deux types d'activités peut être différente. En ce sens, du point de vue pédagogique, le fait d'annoncer le thème, ce dont il va être question, facilite la lecture et, par là, la compréhension.



• **Quelles sont les pistes de recherche les plus explorées actuellement ? Vous avez parlé de l'intelligence artificielle, de la psychologie cognitive. Quelles sont les percées qui pointent à l'horizon ?**

— Ce qu'il me semble, c'est qu'il y a émergence effective, même si on ne le perçoit pas encore nettement, des sciences cognitives. En France, il y a l'Association pour la recherche Cognitive qui regroupe des informaticiens, des psychologues, des linguistes et des philosophes du langage. Des organisations similaires existent dans d'autres pays européens. Les recherches actuelles sont le résultat d'une convergence de préoccupations. Le texte narratif, le récit, a jusqu'ici été le plus étudié. Maintenant, le virage qui se prend, me semble-t-il, consiste à se tourner vers les textes qui transmettent des connaissances (histoire, biologie, etc.). Ce qui est sûr, c'est qu'on peut poser le problème de l'individualisation de l'enseignement en des termes qu'on ne pouvait pas imaginer il y a quelques années. C'est-à-dire que, concrètement, on pourra de mieux en mieux tenir compte des connaissances antérieures des individus, connaissances qui jouent un grand rôle dans l'acquisition et la mémorisation d'informations nouvelles. On peut imaginer que, pour un

Guy Denhière est né en 1943. Il a donné des cours à Paris VIII et est maintenant chargé de recherche au CNRS, pour le Centre d'étude de psychologie (CEPCO) dirigé par Jean-François Le Ny.

Il travaille à un recueil sur la compréhension de textes, à paraître prochainement. Au cours de l'automne dernier, Guy Denhière a été professeur invité à l'Université Laval, dans le cadre d'un séminaire animé par Régine Pierre.

domaine thématique déterminé, par exemple la météo ou l'aquaculture, on stocke toutes les connaissances disponibles sur un gros ordinateur. Un individu intéressé par ces questions pourra très bien, à l'aide d'un terminal quelconque, interroger cette base de connaissances et cheminer en fonction de ses intérêts. On peut aussi envisager cela d'une manière plus systématique et, dans un cadre pédagogique, tenir compte de ce que l'on sait du point de vue cognitif et établir ce que Le Ny a appelé un « dialogue cognitif ». Il s'agit ici d'organiser les connaissances dans la mémoire de la machine en fonction de ce que l'on sait déjà du fonctionnement humain et de la manière dont les connaissances sont stockées dans la mémoire humaine. À partir de là, quand quelqu'un veut acquérir de l'information sur un domaine, l'ordinateur commence par l'interroger, de façon à se faire une idée de son état de connaissances initiales. À partir de ce diagnostic, l'ordinateur présentera — selon des modalités à étudier soigneusement — des informations nouvelles à l'apprenant. Il veillera à ce qu'il n'y ait pas trop d'informations nouvelles d'un coup, parce que si c'est le cas, il y a surcharge et l'individu n'arrive plus à relier ce qu'on lui présente à ses connaissances antérieures et à construire une représentation intégrée et cohérente... À ce moment-là, c'est l'échec, la rupture. Par contre, il ne faut pas non plus que ça soit trop facile, trop trivial. L'ennui arrive vite. Il faut qu'il y ait une quantité optimale d'informations nouvelles. Une telle réalisation, comme je l'indiquais au début, allie les apports de plusieurs disciplines scientifiques. Ainsi l'organisation des connaissances dans la mémoire de la machine suppose à la fois l'intervention d'informaticiens, d'experts du domaine enseigné et de psychologues ayant étudié la structuration des connaissances dans la mémoire humaine. Plus précisément, toutes les informations relatives à un thème déterminé n'ont pas la même importance et ne se situent pas au même niveau. Le questionnement initial doit porter sur les différents niveaux de la hiérarchie ; le texte présenté à l'apprenant doit tenir compte à la fois de ses réponses aux questions posées, de l'organisation du thème traité, du but visé, et, bien évidemment, de ce que l'on sait des lois de l'apprentissage et de la mémorisation. C'est peut-être là une des pistes d'avenir parmi les plus stimulantes pour la recherche et, tout aussi important, un domaine lié au développement des technologies nouvelles (les ordinateurs de la 5<sup>e</sup> génération) et à l'évolution sociale et politique de nos sociétés.

Propos recueillis par  
**Christian VANDENDORPE**