

**Une approche technologique  
pour l'entraînement à la vérification fiscale**

Louis-Claude Paquin

Centre d'analyse de textes par ordinateur

Université du Québec à Montréal

Claude Boivin

Direction générale de la vérification

Ministère du Revenu

Gouvernement du Québec

**Résumé**

Dans les organisations où la pratique est fondée sur des textes : lois, règlements, guides de procédures, etc., comme la Direction générale de la vérification, il est difficile d'assurer aux employés l'accès continu à toute l'information nécessaire à leur entraînement pour accomplir une tâche donnée ainsi que leur recyclage régulier. Nous proposons comme solution une application simple de l'informatique cognitive des organisations. Un système expert, qui reproduit des portions de la tâche, est jumelé à une fonction d'aide hypertextuelle constituée au moyen des techniques de l'analyse de textes par ordinateur. Ce système fournit à l'utilisateur, sur demande, mais avec parcimonie, toute l'information dont il a besoin pour exécuter sa tâche avec efficacité.

## **Le contexte**

Le travail de vérification fiscale se caractérise par des activités d'analyse et d'interprétation qui nécessitent la consultation de nombreux documents. Ces tâches font également appel à l'exercice d'un jugement qui s'affine avec l'expérience, avec l'accumulation et la diversité des dossiers traités. L'embauche de spécialistes des questions fiscales est difficile parce que de tels spécialistes sont rares; en effet, il faut plusieurs années pour former des experts en vérification fiscale québécoise. Après une courte formation magistrale, l'entraînement se fait en situation sur le mode maître-apprenti; un vérificateur expérimenté supervisera un novice pendant une période pouvant durer jusqu'à six mois. Par la suite, le superviseur ou le chef d'équipe dépannera les nouveaux vérificateurs. Il faut compter environ deux ans pour qu'un nouvel employé devienne pleinement autonome. Par ailleurs, dans le contexte de modifications fiscales fréquentes, les vérificateurs ont régulièrement besoin de se recycler.

La Direction générale de la vérification (DGV), pour remplir son mandat, fait face à des contraintes bien connues des organisations, lorsqu'il est question de formation de personnel :

### A) cours magistraux

- il est difficile de rendre les employés disponibles pour préparer ou suivre des cours de formation, leur temps devant être accordé en priorité aux activités de vérification;
- la formation ne répond pas suffisamment aux besoins reliés aux opérations;

### B) supervision-conseil

- souvent, l'entraînement se limite à la lecture de documents et à l'accompagnement d'un vérificateur d'expérience;
- l'implication des employés les plus expérimentés dans le processus de formation réduit la disponibilité de ce personnel pour les opérations. De plus, c'est une tâche impopulaire dans la mesure où ils ne sont pas préparés pour effectuer des tâches d'entraînement;

### C) information

- le temps de préparation et de diffusion des cours de mise à jour est jugé trop long. Le personnel n'a pas toujours à temps l'information nouvelle et se sent parfois démuné face à des dossiers complexes, mal à l'aise face à un mandataire bien informé;
- les cours de mise à jour sur les lois et règlements sont jugés insuffisants.

Dès 1985, un plan directeur des systèmes administratifs proposait, entre autres mesures, la recherche de nouvelles technologies, notamment pour que le personnel ait accès à de l'information qui soit à jour. En mai 1988, nous avons présenté un projet de système à base de connaissances pour assister l'entraînement des vérificateurs. En juin 1990, une maquette a été réalisée dans le domaine de la taxe sur les repas et hôtellerie et présentée aux superviseurs et gestionnaires. Ce développement fait partie des efforts investis pour améliorer les moyens de vérification du ministère du Revenu.

Après avoir exposé les caractéristiques de la formation en vérification fiscale, nous décrivons dans cet article le prototype d'un système à base de connaissances pour assister l'entraînement du personnel. Ce système a été développé avec l'appui de la Direction générale des systèmes d'information du ministère du Revenu et de la Direction générale des technologies de l'information du ministère des Communications dans le cadre d'un projet pilote. Nous énumérons d'abord les sources de la **connaissance** à acquérir. Nous cernons ensuite le domaine du prototype. Puis, nous présentons les réflexions que nous avons menées, d'une part, sur le choix de l'approche à la formation la plus adaptée et, d'autre part, sur le choix de la technologie pour implanter l'approche retenue. Par la suite, nous décrivons le prototype qui découle de ces choix. Nous concluons enfin sur les bénéfices anticipés d'un tel développement ainsi que sur l'élargissement que l'on pourrait envisager.

## **Les sources de la connaissance**

Un dossier de vérification de qualité doit contenir une argumentation cohérente et complète donnant toutes les références pertinentes. La documentation à connaître pour réaliser les vérifications est volumineuse (de 6 à 8 cahiers), très complexe, aride à lire et variée. Pour le domaine retenu, la taxe sur les repas et l'hôtellerie, le vérificateur doit souvent consulter plusieurs sources d'information, telles :

- la Loi concernant la taxe sur les repas et l'hôtellerie,
- la Loi concernant l'impôt sur la vente en détail,
- la Loi sur le ministère du Revenu,
- la Loi sur les Indiens,
- la Loi sur les licences,
- les cas de jurisprudence,
- les directives internes d'administration,
- les directives internes d'interprétation,
- les interprétations de Revenu Québec,
- les nouvelles fiscales de Revenu Québec,
- le manuel des techniques de vérification,
- les directives opérationnelles,
- les procédures opérationnelles,
- la classification des activités économiques du Québec.

Les recherches d'information entraînent la consommation d'un temps précieux. Sur la route, les vérificateurs utilisent un document de travail qu'il se sont confectionné, composé d'extraits des textes jugés les plus utiles. Lorsque les informations souhaitées ne s'y trouvent pas, ils appellent au bureau pour consulter un collègue ou le chef d'équipe, ce qui entraîne une double perte de temps et

une possibilité d'erreur. De plus, les mises à jour des documents sont fréquentes, mais parviennent aux vérificateurs avec un délai important.

Bien que tous ces textes constituent la documentation de base pour les vérificateurs, et malgré le caractère prescriptible des textes de loi et de règlement, ils ne fournissent que des principes en regard du travail de vérification à exécuter. Seule l'expérience acquise fournit des critères de jugement et des raisonnements adaptés aux situations particulières. Ce ne sont jamais des règles absolues, mais des connaissances intuitives (heuristiques) et donc incertaines. Quelle est la marche à suivre dans un cas précis? Tel point de la procédure suggérée peut-il être éliminé? Telle autre procédure de vérification doit-elle être entreprise? Ces connaissances devraient aussi, avec les précautions qui s'imposent, être mises à la disposition de l'ensemble des vérificateurs fiscaux d'un même groupe.

### **Le domaine de prototypage**

Le secteur responsable de la Loi concernant la taxe sur les repas et l'hôtellerie a été retenu comme domaine pour le prototypage d'un système informatique d'assistance à la vérification fiscale. Il s'agit d'un des secteurs où la période d'entraînement du nouveau personnel est la plus longue, étant donné la démarche d'apprentissage (cas par cas) et la diversité des établissements à vérifier. De plus, un projet précédent d'aide à la vérification, qui a fourni aux vérificateurs de ce secteur une application du type , a mis en place une infrastructure (micro-ordinateurs portatifs) dont la rentabilité peut être augmentée. Le sous-domaine d'activités retenu est la vérification de restaurants sans permis d'alcool (activité 9212 de la classification des activités économiques du Québec). Cette catégorie comprend : les restovolants, cafés, cafés-terrasse, cafétérias, casses-croûte et salles de réception.

À l'intérieur du processus de vérification, la phase qui a été retenue pour la maquette est la préparation d'un dossier. Cette tâche d'analyse, qui se déroule aux bureaux du ministère, a été découpée en trois blocs. Premièrement le vérificateur rassemble les documents suivants:

- le dossier général, loi repas et hôtellerie,
- le dossier général, loi de l'impôt sur la vente en détail ,
- le dossier de vérification ou le rapport du vérificateur,
- le rapport des remises de taxe, à jour,
- le rapport des remises antérieures si nécessaire,
- le dossier impôts-individus, ou le dossier impôts-corporations ,
- les autorisations de notes de repas ou de rubans de caisse,
- le relevé des quantités de notes expédiées par le ministère au mandataire.

Il s'agit d'une étape difficile pour le nouveau personnel et fastidieuse pour les employés expérimentés. Des changements sont apportés régulièrement aux procédures. Chaque dossier nécessite une attention particulière quant aux définitions et à la localisation des documents, en somme beaucoup de détails ennuyeux à retenir. Deuxièmement le vérificateur doit extraire des différents documents rassemblés les informations pertinentes, telles:

- la nature d'un établissement
- ses caractéristiques
- le genre de services fournis
- le code d'activité
- le numéro d'enregistrement
- etc.

Cette étape peut présenter certaines difficultés: les mêmes informations sont souvent à plusieurs endroits et peuvent être divergentes. De plus, il est souvent difficile d'identifier ce qui est pertinent au dossier. Troisièmement le vérificateur doit poser un diagnostic, incertain à cette étape, qui orientera la vérification sur place. À la fin de la phase de préparation, le vérificateur pourra avoir identifié des problèmes qui nécessiteront d'ajouter au plan de vérification standard des actions supplémentaires. Ces problèmes vont varier selon le genre de commerce et leur identification correcte dépend de l'expérience du vérificateur.

## **Choix de l'approche à la formation**

Les contraintes qui pèsent sur la formation du nouveau personnel de vérification sont nombreuses et contradictoires. D'une part, comme nous venons de le voir, les connaissances sont nombreuses, variées et complexes; et la tâche fait appel à un jugement qui ne s'acquiert qu'avec l'expérience. D'autre part, le nouvel employé doit être opérationnel le plus rapidement possible. Rappelons brièvement la formule de formation présentement utilisée dans le secteur de la taxe sur les repas et l'hôtellerie, en fonction de ces contraintes. Les cours de formation en groupe sur le mode magistral sont réduits au minimum. Le nouvel employé passe ensuite une courte période d'étude au bureau (1 à 2 semaines). Puis, pendant quelques semaines, sous supervision complète, il est confronté avec des cas réels. Enfin, il demeure plusieurs mois sous supervision partielle.

Cette situation indique qu'un support technologique à l'apprentissage dirigé, actuellement offert par des cours magistraux, ne constitue pas une cible privilégiée pour notre projet de développement. En effet, cette phase est marginale dans le cycle de formation et entraîne des coûts élevés de médiatisation qui, de plus, devraient être renouvelés pour tenir compte des changements fréquents dans les procédures. En raison des contraintes mentionnées précédemment, notamment de la nécessité d'être opérationnel le plus rapidement possible, la principale pédagogie demeure la supervision, qui consiste à apporter au nouvel employé, dans l'accomplissement de la tâche, les réponses aux questions qu'il soulève. On comprendra que, si une technologie ne peut remplacer complètement le superviseur, une aide est cependant possible pour répondre à des questions simples et régler de nombreux petits détails (définitions, acceptions, informations générales) entourant les différentes phases de la vérification. Ces connaissances, élémentaires pour un expert, dépassent toutefois la trivialité pour le nouvel employé. Selon le principe de Pareto (Association for systems Management 1974), nous estimons à titre indicatif que 80 % des questions posées par un nouvel employé ont trait à des connaissances élémentaires qui prennent 20 % du temps de réponse du



superviseur. Nous relierons ainsi l'efficacité de la supervision à la capacité d'être dégagé des questions accessoires pour accorder le temps aux questions les plus difficiles. Pour ce qui est de l'évaluation de l'apprentissage, peu de temps lui est accordé, en raison, d'une part, des contraintes de planification et, d'autre part, parce qu'un processus de révision en cascade des dossiers (plusieurs réviseurs, à différents paliers de décision) est en place. Ce qui compte finalement, c'est que l'employé ait l'information requise pour travailler efficacement.

L'objectif de l'entraînement en milieu de travail, c'est que l'employé acquière les habiletés nécessaires à la tâche, puis l'autonomie. Ce dernier objectif nous a inspiré le concept d'entraînement autonome ou autodidaxie, notion que l'on situe entre la formation dirigée et l'entraînement sous supervision. L'apprenant décide de ce qui est nécessaire à sa compréhension et demande le niveau d'explication qu'il juge suffisant lorsqu'il accomplit une tâche donnée. Une lecture volontaire et répétée d'explications pertinentes en contexte nous apparaît un facteur d'apprentissage non négligeable. Étant donné les changements fréquents, ce mode de formation va rencontrer aussi les besoins continus d'assistance à la tâche.

### Choix de la technologie

Le tableau suivant, tiré d'une étude (Hamilton, C. et Kleiner, B. 1987) compare différents moyens de formation du point de vue de la capacité de rétention du contenu :

Moyen pédagogique utilisé	Pourcentage de matière retenue		
	sur le fait	après 3 heures	après 3 jours
	%	%	%
Lecture	10		
Écoute	20	70	10
Visionnement	30	72	20
Visionnement et écoute	50	85	65
Recherche et exercice	80	> 95	>80

L'approche est présentée comme la plus performante. Les systèmes d'enseignement assisté par ordinateur (EAO) relèvent de cette dernière approche. Ils consistent habituellement en des enchaînements stricts de questions et de réponses selon un scénario pré-déterminé ce qui, par contre, ne favorise pas l'objectif d'autonomie de l'apprenant. De plus, les coûts associés à cette technologie sont élevés:

«...la nécessité de programmer les écrans et l'ensemble du scénario (directement ou au travers de langages ou de systèmes auteurs) conduit à des rapports de 150 à 1 entre le nombre d'heures de travail de médiatisation d'un personnel spécialisé et le nombre d'heures de dialogue "élève" produites"» (Meurrens, M. 1989)

À titre d'exemple, la grille de tarification d'une compagnie d'EAO établissait en mars 1989 à environ 15 000 \$ le coût de production d'une heure de cours et à environ 7 000 \$ le coût de traduction d'une heure de cours. En mars 1989, le prix offert au ministère du Revenu pour une version micro-ordinateur du logiciel d'EAO Phoenix était de 6 000 \$; le coût d'une version pour l'ordinateur central variait entre 45 000 \$ et 110 000 \$.

Les systèmes d'enseignement intelligemment assisté par ordinateur (EIAO) se situent à l'autre extrémité du spectre de ces technologies d'apprentissage. Ils se caractérisent par l'absence de déterminisme dans la progression des sessions de travail et par la prise en compte du profil psychopédagogique de l'apprenant. Le coeur des systèmes d'EIAO est en fait un système à base de connaissances (SBC) qui reproduit le domaine enseigné, constitue et tient à jour un modèle de l'apprenant et gère le dialogue d'apprentissage (Ozman, G. 1990). Il s'agit cependant pour l'instant de prototypes de recherche qui s'avèrent très intéressants mais beaucoup trop complexes pour la souplesse souhaitée par l'approche à la formation retenue.

Par ailleurs, la technologie des SBC en elle-même offre une approche intéressante pour l'assistance à la vérification fiscale. Il semble à la fois difficile et non souhaitable de confiner dans une

procédure aucun des trois blocs d'activités qui constituent la préparation d'un dossier de vérification. En effet, même si les informations peuvent être obtenues de bases de données ou autres systèmes du ministère, la mise au point d'un algorithme pour automatiser leur sélection et leur traitement (interprétation) est fort problématique. Comme nous avons pu le constater précédemment, le savoir-faire dans ce domaine est complexe, varié, étendu et pas toujours sûr; tous les cas sont différents (extension); de nombreuses manipulations logiques conditionnelles sont nécessaires et l'expérience accumulée est mise à contribution pour combler les silences de la procédure générale. Qui plus est, les procédures à suivre font l'objet de fréquentes mises à jour. Ainsi, le recours aux méthodes normatives s'avère trop coûteux, autant en temps de développement et de mise à jour qu'en temps machine. L'utilisation d'une méthodologie de prototypage de type adaptée à l'environnement du micro-ordinateur nous apparaît souhaitable. Ainsi, lorsqu'une méthode d'entretien aura été mise au point et validée, le système pourra régulièrement être mis à jour et amélioré, au fil de la stabilisation et de la croissance du savoir-faire par l'intermédiaire de ses experts et des cas traités.

Après avoir s'être interrogé sur la pertinence de la méthodologie des SBC pour offrir de l'assistance au vérificateur, voyons maintenant leur apport à la formation. D'abord, les SBC opèrent avec une certaine transparence quant aux étapes suivies (raisonnement). Les générateurs de SBC sont habituellement dotés d'un traceur qui est à même de fournir dans une fenêtre certaines explications en contexte (pourquoi une question est posée et comment le système est parvenu à une réponse donnée). Cette fonctionnalité des SBC permet déjà un minimum d'activité didactique qui, en gros, correspondrait à une approche (voir plus haut). De plus, les règles d'inférences offrent un mécanisme souple pour poser des questions, saisir la réponse et afficher des messages dans un ordre établi. Cependant tout le cadre d'apprentissage doit être défini à l'avance, tout comme dans le cas des systèmes d'EAO, mais le grand avantage des SBC sur ces derniers systèmes est la souplesse et la modularité d'une stratégie dirigée par les faits.

Dans les systèmes d'EIAO, l'initiative du dialogue est partagée entre le didacticiel et l'élève. Le cadre de la formation autonome que nous avons défini s'apparente plutôt à une situation extrême de dialogue où l'initiative repose entièrement sur l'apprenant. Lorsque (Meurens, M. 1989, p.39). C'est ainsi que nous sommes parvenus à la conception d'une fonction hypertextuelle qui serait adjointe au SBC modélisant la tâche de vérification. L'initiative serait alors totalement laissée à l'utilisateur qui décide du niveau d'information dont il a besoin. Le système ne lui fournirait que les informations qui correspondent à sa requête tout en lui permettant de naviguer à sa guise dans l'ensemble des informations par le biais d'un réseau de concepts associés.

Nous faisons l'hypothèse que les moyens que nous proposons, que l'on pourrait intituler résentent des avantages sur l'EAO en regard de la capacité de rétention des connaissances par l'utilisateur. En effet, d'une part, l'exécution d'une tâche réelle est une situation qui revêt une importance plus grande pour l'utilisateur que la solution d'exercices et lui impose une pression supplémentaire que nous estimons positive pour l'assimilation des connaissances. D'autre part, l'utilisateur est assisté par le système dans la recherche des informations nécessaires à l'exécution de la tâche; cette recherche est contrôlée par l'utilisateur qui reçoit une réponse immédiate et décide de l'ampleur à donner à sa requête. La rapidité des réponses fournies par le système, le besoin découlant d'une tâche réelle et un principe de parcimonie nous apparaissent des facteurs de motivation favorisant l'ancrage des connaissances chez l'employé. Ce genre d'aide rejoint la notion de tème aviseur» (Duchastel, P. J.et Braham, J., 1989).

### **Description du prototype**

Les objectifs du prototype développé consistaient à démontrer la faisabilité de l'entraînement autonome, à valider les choix technologiques, à identifier les caractéristiques désirées pour un

système final et enfin à déterminer l'envergure de l'effort requis pour son développement. Les tâches accomplies par le prototype sont les suivantes :

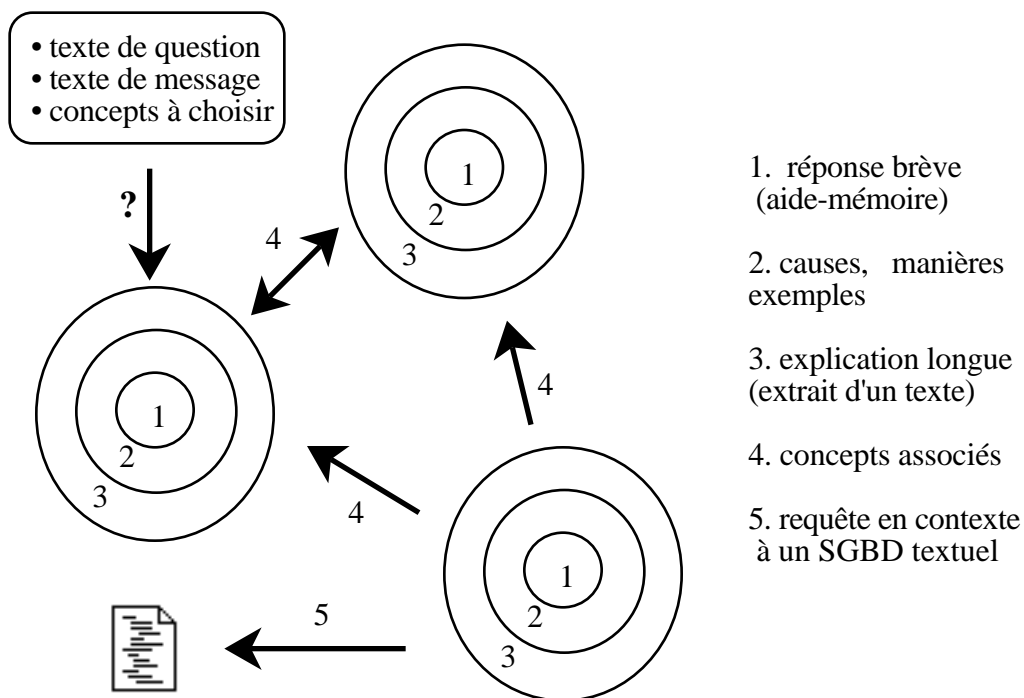
- guider un employé dans les documents à rassembler;
- demander l'information requise;
- fournir de l'assistance dans le diagnostic du cas;
- identifier les situations nécessitant une action supplémentaire au plan standard de vérification
- calculer un indice de la préoccupation
- produire un rapport d'analyse.

Lors de l'étude de besoins qui a précédé la mise en chantier du prototype et durant le prototypage, les caractéristiques suivantes ont été demandées:

- convivialité (menus et fenêtres de déroulement),
- éviter de demander des renseignements répétitifs (comme l'identification de l'employé),
- effectuer des retours en arrière lorsque nécessaire,
- faciliter la poursuite de plusieurs consultations en parallèle,
- pouvoir utiliser tout ce qui est entré dans le système pour produire automatiquement les rapports nécessaires,
- un mécanisme de traitement des connaissances incertaines,
- rendre l'information disponible directement sur les lieux du travail sur le matériel existant (PC-AT compatibles).

Le D\_expert, auparavant appelé Déredec-EXPERT, (Paquin L.-C. 1987 et 1990) développé au Centre d'analyse de textes par ordinateur de l'Université du Québec à Montréal est le générateur de systèmes experts retenu pour développer le prototype. Même s'il s'agit d'une première version en voie d'être recodée, ce générateur offre toutes les caractéristiques ci-haut mentionnées sauf la dernière qui est moins importante à ce stade-ci du développement. Ce choix nous offrait de plus la possibilité de faire réaliser dans le générateur des fonctions et spécifications répondant aux besoins spécifiques de notre projet.

C'est ainsi qu'a été développée la fonction, appelée , qui implante la formation autonome dans le contexte des SBC. Cette fonction se veut une solution générique au problème de la documentation des SBC; elle est indépendante de son contexte de développement. Disponible en tout temps lors de la consultation, elle consiste en l'affichage dans une fenêtre particulière d'une fiche comportant de l'information, texte ou image. Elle est activée par l'utilisateur qui sélectionne en contexte une chaîne et presse la clé du point d'interrogation ou sur une icône qui le représente. Les fiches sont structurées en trois niveaux selon un modèle de cercles concentriques. Les fiches peuvent de plus être réunies entre elles par des renvois. Un tableau de bord sous la fenêtre d'affichage du texte permet le passage d'un niveau à un autre de la fiche au moyen de trois et un quatrième intitulé permet l'accès à l'une des fiches qui lui est associée. On pourrait même imaginer un cinquième bouton qui permettrait de soumettre la chaîne sélectionnée en requête à une base de données textuelle afin d'obtenir tous les contextes d'occurrence de celle-ci. Le tableau suivant illustre cette structure:



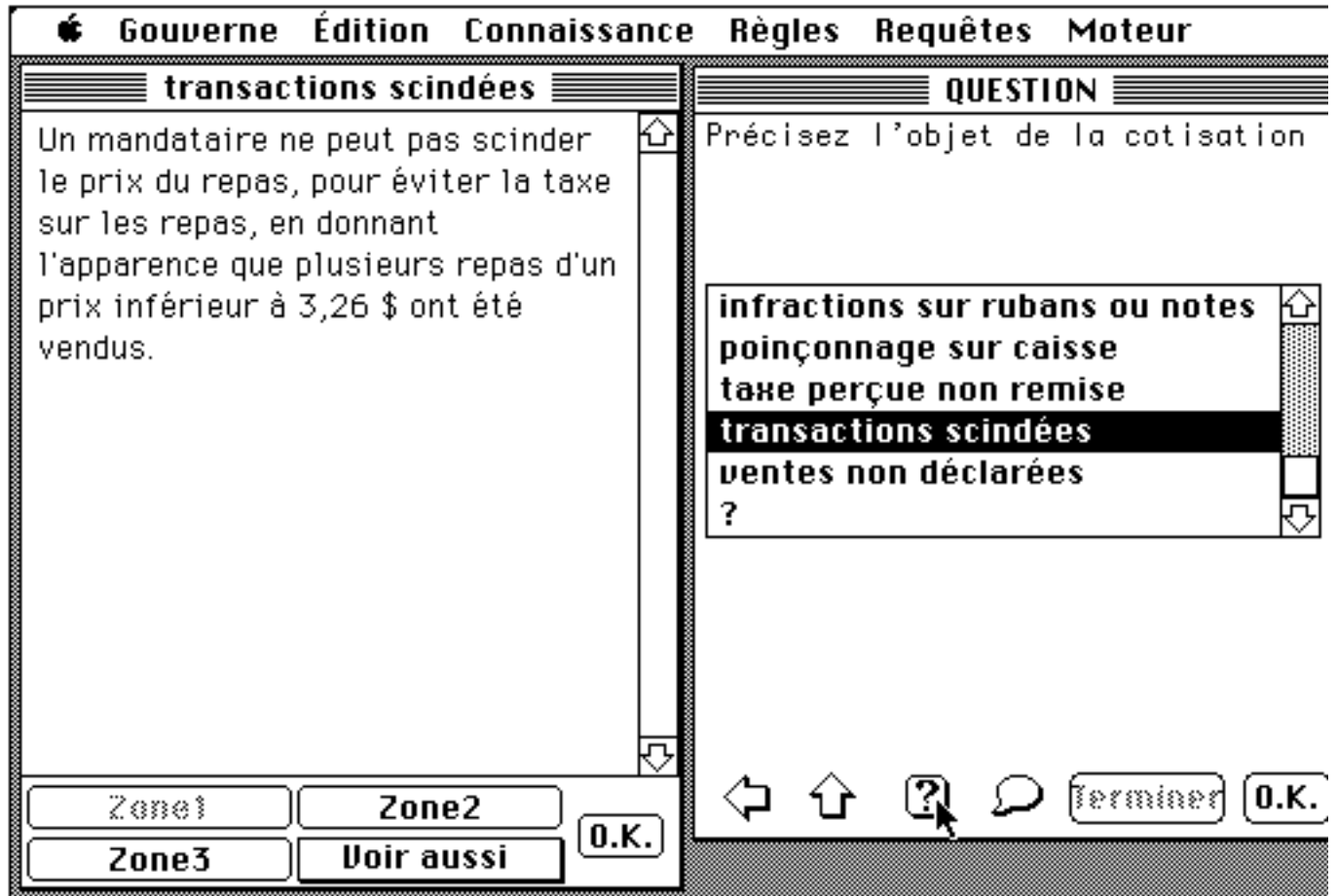
Dans le présent domaine, la fonction Hyper\_aide veut répondre aux questions suivantes :

- que veut dire ce terme ?
- comment faire cette tâche?
- que veut dire ce message?
- comment le système arrive à cette conclusion?
- pourquoi cette question?
- où est située l'information demandée sur ce formulaire?
- comment interpréter telle formulation?
- comment obtenir ce dossier?
- etc.

Par contre, dans Loge-expert, un système expert destiné à rendre accessible le droit du logement au grand public, cette même fonction donne accès à des définitions, des références et la jurisprudence pertinente (Thomasset, C.; Blanchard, F. et Paquin, L.-C. 1991).

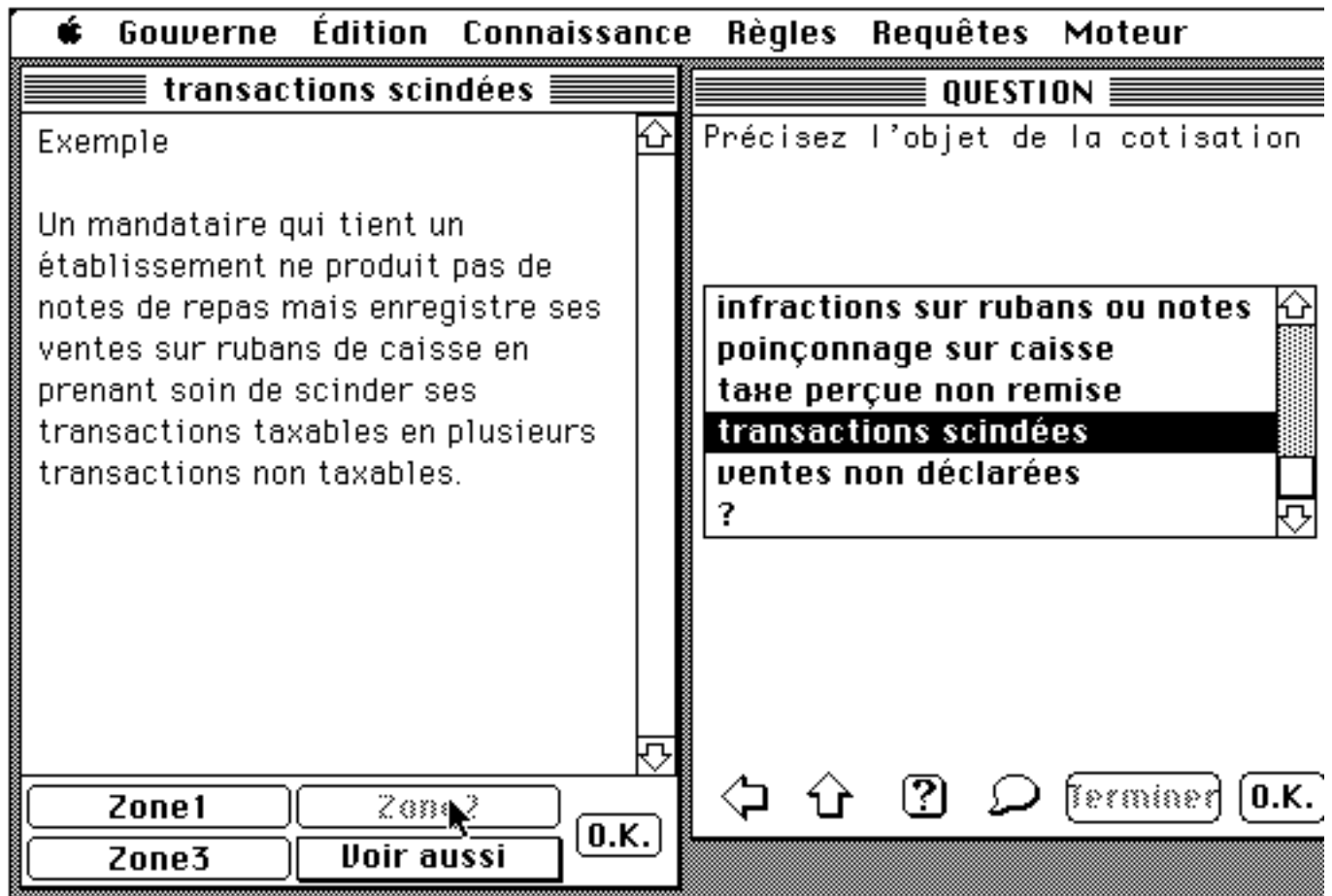
Voici une illustration commentée du fonctionnement de la fonction Hyper\_aide, dans le contexte de l'application de la Loi concernant la taxe sur les repas et l'hôtellerie pour la période d'avant 1991. Durant la consultation et à l'occasion d'une question, un choix parmi une série de termes est offert à l'utilisateur. Celui-ci sélectionne le terme désiré, qui apparaît alors noirci à l'écran; au lieu de valider la sélection en appuyant sur le bouton , l'utilisateur appuie sur le bouton , ce qui provoque l'affichage d'une seconde fenêtre, à gauche de la fenêtre de question.

La première zone d'information qui est affichée présente alors une brève explication ou une définition du terme sélectionné :

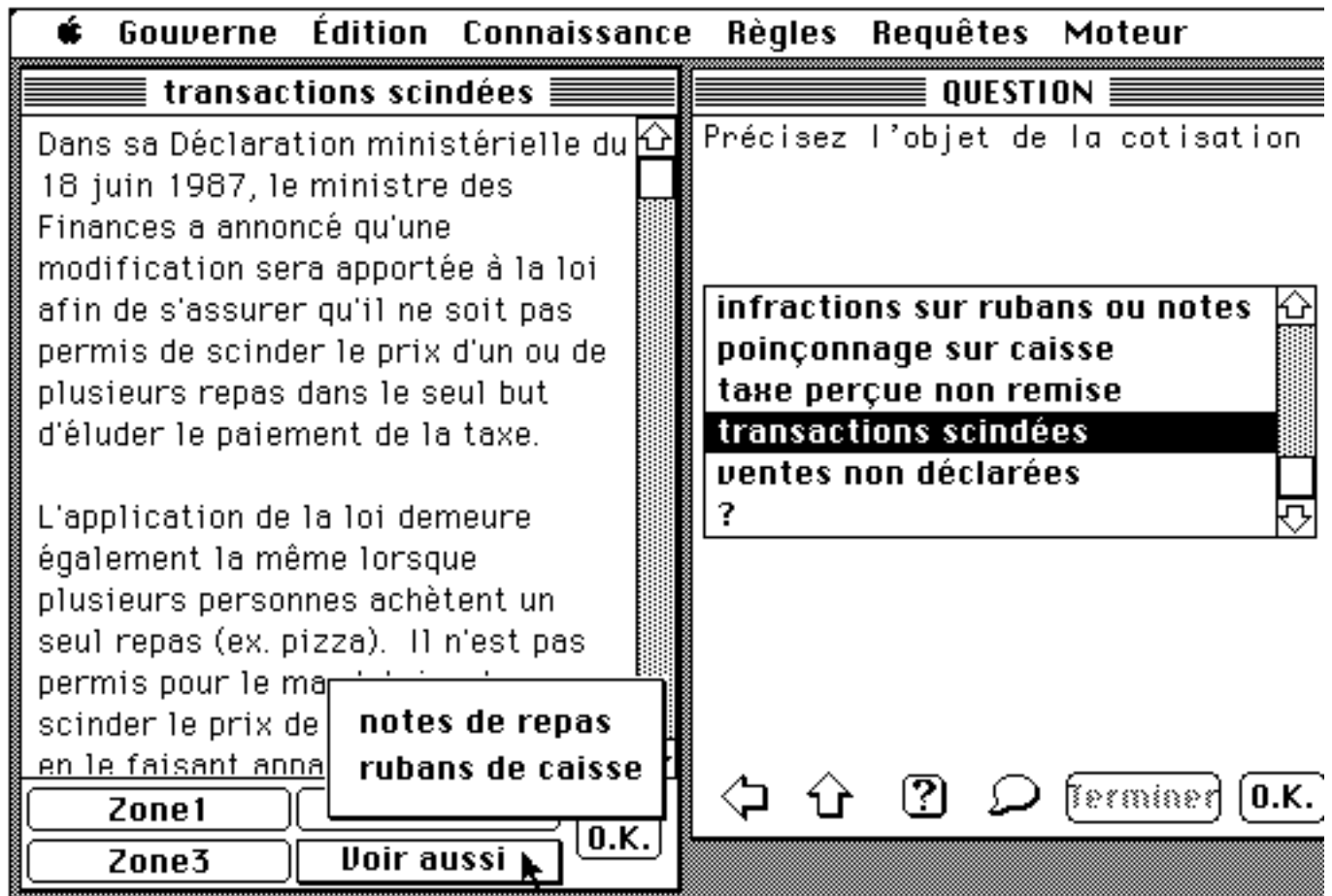




Si le besoin s'en fait sentir, un supplément d'information, un exemple dans le cas présent, peut être obtenu en appuyant sur le bouton :



Une explication détaillée peut aussi être demandée en appuyant cette fois sur le bouton . Lorsqu'on appuie le bouton , le cas échéant, un menu offrant des termes associés apparaît:



Si un terme est sélectionné, le texte de la fiche correspondante est affiché dans la fenêtre et ainsi de suite. L'exploration du réseau Hyper\_aide se termine en appuyant sur le bouton .

La fonction Hyper\_aide sera utile dans la mesure où les développeurs du SBC se doteront d'une fonction qui régira de façon constante, d'une part, la division des fiches en zones et, d'autre part, l'assignation des termes. La relation entre deux termes peut être explicative ou associative. Pour un principe d'économie et de lisibilité, le texte d'une fiche consacrée à une question du système renfermera des informations sur le sens précis de la question, sur la raison pour laquelle elle est posée ou sur la façon d'obtenir la réponse, mais ne donnera pas la signification de chacun des termes qui la

composent; des relations explicatives avec les fiches de ces termes seront plutôt établies. Il est plus difficile d'énoncer des principes clairs pour établir des relations associatives. «Elle couvre des relations entre paires de termes qui ne font pas partie d'un même ensemble d'équivalences et qui ne font pas partie de la même chaîne hiérarchique, mais qui sont mentalement associées de façon telle que le lien entre eux doit être rendu explicite» (ISO 2788, 1986, p. 17) pour une compréhension globale du domaine. La co-occurrence fréquente des deux termes dans les textes est un bon indice de relation associative. Les concepts associés peuvent aussi être identifiés par leur relation avec un ensemble de termes voisins. Par exemple la notion est associée à la notion par leur lien commun avec les notions voisines et . Afin d'éviter une explosion du réseau de relations entre les termes du domaine, nous avons limité les relations aux termes utilisés dans les deux premières zones des fiches.

La constitution des fiches passe par l'exploitation de l'ensemble des textes de l'organisation. À cet égard, la technologie des bases de données textuelles est très intéressante. Nous prôtons le recours à l'analyse de textes, tant pour le dépistage des concepts impliqués dans le calcul symbolique effectué par le SBC (Paquin, L.-C.; Dupuy, L. et Rochon, Y. 1990) que pour visionner leurs contextes d'occurrence pour constituer le texte des différentes zones de la fiche. Un recours minimum consisterait en la formulation de patrons de fouille à partir du terme que l'on souhaite documenter pour obtenir des concordances, c'est-à-dire, des extraits de textes où le patron de fouille est réalisé. L'exemple qui suit a été réalisé à l'aide du logiciel SATO (Système d'Analyse de Texte par Ordinateur, par F. Daoust du Centre d'ATO de l'Université du Québec à Montréal) et porte sur l'application de la Loi concernant la taxe sur les repas et l'hôtellerie pour la période d'avant 1991; l'exemple illustre une recherche pour valider l'association entre les concepts et , en les reliant aux concepts et :

Concordance logique scinder (prix,addition)

# 4 \*page=b-rh17/1/39 ... \*page=b-rh17/1/41/10

Un mandataire ne peut pas **scinder** le prix du repas, pour éviter la taxe sur les repas, en donnant l'apparence que plusieurs repas d'un prix inférieur à 3,26 \$ ont été vendus.

# 7 \*page=b-rh17/2/24 ... \*page=b-rh18/1/5/5

L'application de la loi demeure également la même lorsque plusieurs personnes achètent un seul repas (ex.. pizza). Il n'est pas permis pour le mandataire de **scinder** le prix de vente de ce repas en le faisant apparaître sur plusieurs additions. En effet, bien que ce repas sera éventuellement partagé entre plusieurs personnes, il constitue néanmoins, au moment de la vente, un seul repas lequel est taxable s'il est acheté pour un prix de plus de 3,25 \$. Taxes à la consommation. Texte pour bulletin d'Interprétation. Numéro: R..H.. 18 DATE: 30 novembre 1988 PAGE: 1 DE 3

commande: Abréviatiion Aide Commentaire Compter **Concordance** Contexte +

Concordance logique (diviser,divisé\$) (prix,addition)

# 2 \*page=l-rh1/8/5 ... \*page=l-rh1/8/14/8

Taxe sur les repas. -article- 4. Lorsqu'il n'y a qu'une seule addition pour plusieurs personnes ou pour plusieurs repas, cette addition ne doit pas être **divisée** suivant le nombre de personnes ou de repas et, malgré l'article 2, la taxe se calcule sur le total des prix des repas apparaissant sur cette addition. REFERENCES S.. R.. 1964, c.. 73, a.. 4; 1965 (1 sess..), c.. 30, a.. 3; 1988, c.. 4, a.. 192.

# 5 \*page=b-rh17/2/5 ... \*page=b-rh17/2/7/6

L'article 4 de la loi prévoit que lorsqu'il n'y a qu'une seule addition pour plusieurs personnes, cette addition ne doit pas être **divisée** suivant le nombre de personnes.

commande: Abréviatiion Aide Commentaire Compter **Concordance** Contexte +

Il est très rare qu'un extrait de texte convienne sans changement, comme c'est le cas pour la concordance #4 dans l'exemple de fiche plus haut. Habituellement un travail de révision des extraits de texte est requis pour les rendre compréhensibles en eux-mêmes. En effet, le processus d'extraction d'un fragment de texte cause un démembrement de la macro-structure textuelle qui entraîne une rupture dans la continuité. C'est ainsi qu'un fragment qui semblait significatif lors de la lecture continue, l'est beaucoup moins une fois extrait. La révision la plus usuelle consiste en la clôturation de la référence du fragment de texte sélectionné. Pour ce faire, les anaphores - les mots

qui sont mis à la place des référents premiers, comme par exemple les pronoms - et les synonymes - des termes qui sont contextuellement équivalents - doivent être remplacés par les termes utilisés dans le SBC. Pour obtenir une définition complète, notre intervention doit parfois être plus importante; elle peut aller jusqu'à synthétiser en une seule formulation des éléments qui apparaissent de façon isolée en raison d'un contexte donné. Cette opération est d'une importance capitale pour le succès de la formation autonome: pour soutenir l'intérêt de l'utilisateur, l'information qu'il recherche doit être accessible de façon claire et complète et concise.

La réalisation et la gestion des fiches de la fonction Hyper\_aide se fait indépendamment du D\_expert, à l'aide d'un traitement de textes et/ou d'une base de données en utilisant un champ dont le contenu non-indexé est un texte d'une longueur indéfinie - le format . Le lien avec le D\_expert se fait par une version en format ou de l'ensemble des fiches où chaque fiche est précédée de la clé - la chaîne de caractère, terme, question ou message, qui est documentée entre crochets et où chaque zone de la fiche est délimitée par le séparateur . La longueur du texte de chacune des zones est limitée à 32 000 caractères; les sont écrits à raison de un par ligne. Cette façon de faire permet un développement modulaire du système par plusieurs équipes de spécialités différentes: ingénierie de la connaissance pour la réalisation du SBC, la linguistique et la communication pour la fonction Hyper\_aide.

La fonction Hyper\_aide, telle que nous venons de la décrire, présente une grande similitude avec la technologie de l'hypertexte. Les principes de base sont les mêmes: rompre avec la dimension linéaire d'un document, favoriser le libre accès au document, etc. (voir Barret, E. 1988 et 1989). Toutefois plusieurs défauts qui ont été reprochés à des plateformes hypertextuelles ont été évités ici. L'accès à la fonction Hyper\_aide se fait par une clé et l'information pertinente est présentée à l'utilisateur sans aucune recherche de sa part. La navigation entre les noeuds ou segments de texte est simplifiée au maximum par l'usage d'un tableau de bord et d'un menu. Alors que l'hypertexte tend à rendre accessible le plus d'informations possible, la fonction Hyper\_aide veut raccourcir le plus possible le chemin entre une question de l'utilisateur et le type de réponse désirée. C'est

pourquoi les informations sont distribuées de façon uniforme en trois zones, de sorte que le délai pour trouver un type particulier d'information se trouve réduit. En effet, on est plus rarement à la poursuite d'un détail qui nécessite la consultation de toute l'information disponible, qu'à la recherche d'un aide-mémoire.

## **Conclusion**

Complété par la fonction Hyper\_aide, notre système à base de connaissances présente l'avantage principal de favoriser un entraînement continu du personnel sans pour cela nuire à l'efficacité du travail de vérification. Les questions simples sont confiées au système et les questions complexes sont laissées au superviseur. Nous prévoyons retirer des bénéfices tangibles de l'implantation de ce système. Premièrement, l'ingénierie cognitive, parce qu'elle nécessite l'explicitation de tous les détails et procédures — le savoir-faire — reliés au domaine d'expertise, aura pour effet une simplification du processus de révision, d'où une économie de temps. Deuxièmement, du temps de vérification sera récupéré par la simplification du travail de recherche d'information et une meilleure préparation des cas. Troisièmement, l'utilisation du SBC favorisera l'uniformité des méthodes de travail, ce qui tend vers une meilleure équité fiscale. Quant à la fonction Hyper\_aide, elle rendra en partie les connaissances générales de l'expert directement accessibles au nouvel employé, ce qui aura pour effet d'une part de libérer l'expert du traitement des cas de routine et de lui permettre de traiter plus de cas complexes et, d'autre part, de développer l'autonomie du nouveau personnel et sa motivation au travail. En dernière analyse, la solution que nous proposons constitue un moyen de communication de l'ensemble des connaissances accumulées dans le secteur et ainsi le champ de connaissances d'un nouveau vérificateur ne sera plus confiné à l'expérience de son superviseur.

Parmi les prolongements nécessaires de notre prototype, il y a la gestion de l'entraînement au moyen de données quantitatives : combien de fois chacune des fiches a été consultée, nombre de fiches

consultées par requête, combien de temps a été passé par requête. Une mesure d'utilité du système devra être mise sur pied, basée en partie sur les données quantitatives, en partie sur de l'observation directe des utilisateurs et en partie sur leurs commentaires, recueillis à l'aide d'un questionnaire, quant à leur compréhension des explications fournies. De plus, le système final devra être accompagné de la mise en place d'un centre d'entretien qui assurera la collecte des données d'utilisation et surtout garantira une information intègre et toujours à jour. Nous prévoyons la possibilité d'intégrer ultérieurement un système tuteur qui prendrait sur demande le relais du système conseiller. Nous prévoyons également une liaison avec le système de gestion de la base de données textuelle, actuellement en développement (Harvey, F.; Lamarre, C. et Garceau, S. 1990), en formulant une requête en contexte permettant de visualiser tous les textes qui traitent d'un sujet donné. Lorsqu'elle aura été validée sur le terrain, nous croyons opportun d'exporter aux autres secteurs de la vérification ces caractéristiques, autant celles du SBC lui-même que celles de la méthode de développement que nous avons suivie.

Remerciements à André Brisebois, Serge Garceau et Thomas Leblanc pour leur commentaires et suggestions ainsi qu'aux experts participants au projet.

## Références

- Association for systems Management (1974) "Effectiveness Depends On The Vital Few"  
*Organization Planning, Technical department report* (3), 4 January 1974
- Barret, E. éd. (1988) *Text, Context and Hypertext*, The MIT Press, Cambridge, Mass., 1988.
- Barret, E. éd. (1989) *The Society of Text, Hypertext, Hypermedia and the Social Construction of Information*, The MIT Press, Cambridge, Mass. 1989.
- Duchastel, P. et Braham, J. (1989) "Functionality considerations in a generic advisor system",  
*Colloque ICO'89, Informatique cognitive des organisations*, (B. Moulin et G. Simian, Éd.),  
Québec, 1989, pp. 317 - 326.
- Hamilton, C. ; Kleiner, B. (1987) *Personnel Journal*, Costa Mesa, California, february 1987  
article reproduit dans *Drake, revue d'affaires, "Soyez à l'écoute"*, Vol. 1 # 2, 1988, p. 42-43.
- Harvey, F.; Lamarre, C.; Garceau, S. (1990) "La rencontre de la gestion documentaire et de la gestion textuelle à Revenu Québec", *ICO Québec Gestion de l'information textuelle* 2 (3) , 1990, pp. 62-70.
- Meurrens, M. (1989) "La gestion du modèle de l'apprenant dans le système d'E.I.A.O. "G"" *ICO Québec* 1 (2), 1989, pp. 39 -45.
- Otman, G. (1990) "EIAO et enseignement des langues : le point de vue français", *ICO Québec* 2 (2), 1990 pp. 45 -.52.
- Paquin, L.-C. (1988) "Le Déredec-EXPERT, un atelier logiciel cognitif", *Actes du congrès 'Informatique cognitive des organisations'*, Québec, GIRICO, 1988, pp. 13-20.
- Paquin, L.-C. (1990), D-EXPERT, Manuel de l'utilisateur, (Version 2.0), Centre d'Analyse de Textes par Ordinateur, UQAM, 1990, 140 p.
- Paquin, L.-C.; Dupuy, L.; Rochon, Y. (1990) "Analyse de texte et acquisition des connaissances : aspects méthodologiques", *ICO Québec Gestion de l'information textuelle* 2 (3), 1990, pp. 95-113.
- Thomasset, C.; Blanchard, F. Paquin, L.-C. (1991) Loge-expert : an illustration of different phases of the development of an expert system in law", *Expert Systems with Applications*, R. Wagner éd. (à paraître)