

ÉTIENNE GIRARD

**USAGE DE LA COGNITION SPATIALE POUR LOCALISER LES LIEUX
D'ACTIVITÉ LORS D'UNE ENQUÊTE ORIGINE - DESTINATION**

Mémoire présenté
à la Faculté des études supérieures de l'Université Laval
dans le cadre du programme de maîtrise en sciences géographiques
pour l'obtention du grade de maître en sciences géographiques, (M.Sc.Géogr.)

FACULTÉ DE FORESTERIE ET DE GÉOMATIQUE
UNIVERSITÉ LAVAL
QUÉBEC

JUILLET 2004

RÉSUMÉ

Ce mémoire cerne la problématique de la description qualitative de la localisation d'un lieu décrit en langage naturel. C'est par une approche cognitive qu'est abordé successivement l'apprentissage de l'espace, le stockage de l'information et la restitution de l'information en langage naturel, par l'entremise des concepts de méronymie, de catégories hiérarchiques et de référents spatiaux. De ce cadre théorique, on propose de restructurer une base de données de lieux existants en y ajoutant des paramètres qui permettent de retrouver, d'une description en langage naturel précise ou floue, un lieu sans ambiguïté dans une base de données grâce à une interface usager offrant divers modes de repérage spatial.

RÉSUMÉ LONG

Ce mémoire s'intéresse aux descriptions qualitatives de la position des lieux d'activité dans l'espace géographique (toponymie et systèmes d'information géographique) avec comme objectif d'automatiser ou de faciliter la localisation des données d'enquêtes portant sur les comportements de mobilité des personnes en milieu urbain. Les bases de données contiennent normalement des informations normalisées et officielles ; par contre, les répondants utilisent des références spatiales nominales, souvent toponymiques, pour décrire les lieux d'activité qu'ils ont fréquentés. Or ces références contiennent des «erreurs» souvent prévisibles si l'on connaît les différents processus d'acquisition et de restitution de cette information.

Pour préparer les bases de données utilisées durant l'enquête origine-destination (OD) de l'automne 2001 dans la région de Québec, nous avons effectué un inventaire typologique des informations requises. Notre approche est fondée sur les recherches dans le domaine des sciences cognitives, de la linguistique, de l'intelligence artificielle et de la psychologie. Ces disciplines ont fourni un cadre théorique pour structurer et enrichir l'information géographique en fonction des principes reliés au langage (description d'un espace), à la mémoire (paronymie et catégorisation) et à la construction d'image mentale. Cette recherche repose sur une réflexion à propos des moyens à mettre en œuvre pour construire un outil de localisation qui effectue l'interface entre la perception de l'espace par les individus (géographie cognitive), la description verbale de la localisation des lieux d'activités fréquentés et sa représentation dans un SIG où les éléments doivent être localisés dans un système de coordonnées cartésiennes (x et y) ou géographiques (latitudes et longitudes).

Cette approche a permis de diminuer le temps de validation lors de l'enquête OD 2001 et d'augmenter l'exactitude des localisations. De nouvelles propositions de modification de la base de données et d'une interface usager permettent de diminuer encore le temps de recherche d'un nom de lieu qui est décrit de manière floue lors de l'enquête. Cette nouvelle structure de données est basée sur la recherche en cognition spatiale, doublée d'une interface

usager qui permet, lors d'une enquête, de valider la localisation des lieux décrite en langage naturel et au besoin d'apporter des éléments géographiques supplémentaires à l'enquête pour raffiner une description floue d'un lieu, ceci dans le but de s'assurer de l'exactitude et de l'unicité de sa localisation.

À mes trois fils, Damien, Maxence et Thierry

REMERCIEMENTS

Ce mémoire n'aurait vu le jour sans les gens qui m'entourent et qui m'ont encouragé à terminer. Le contexte dans lequel il a été produit n'était certes pas idéal, mais il est terminé.

Je voudrais remercier mon directeur de mémoire, Marius Thériault pour tout. Je le remercie d'avoir pris tant de temps, qui lui est si rare, pour réviser mon texte et m'encourager à poursuivre. Merci sincèrement Marius. Merci à ma codirectrice, Cécyle Trépanier et mon comité, Marie-Hélène Vandersmissen et Marc Saint-Hilaire.

Je ne pourrai passer sous silence l'opportunité que m'a procurée le département de géographie de faire ce mémoire. Ici je voudrais remercier tous mes collègues de travail, principalement le directeur du département de géographie Michel Allard et Yves Brousseau pour ses encouragements, les gens du laboratoire de cartographie, Sylvie Saint-Jacques, Serge Duchesneau et Karine Tessier pour m'endurer quotidiennement et du secrétariat pour leurs gentilleses et leurs multiples services qui nous rendent la tâche beaucoup plus facile, merci Marie-Josée, Nancy et Denise. J'en profiterai aussi pour remercier Louise Filion, en qui j'ai toujours eu beaucoup d'estime et à qui je dois beaucoup, je ne pourrai lui remettre, je le remettrai à mes étudiants.

Merci à Geoffrey Edwards du département de géomatique pour sa passion sur la cognition en géomatique et sa logique qui m'a permis de replacer certaines pièces du puzzle. Et merci à Sébastien Tremblay du département de psychologie pour m'avoir permis de suivre ses cours et m'initier à cette discipline.

Et finalement à mes proches qui ont du subir mes nombreuses heures de travail sur ce mémoire le soir et les fins de semaines. J'ai cru tellement de fois qu'il ne me restait que quelques jours de travail que j'ai créé cette illusion aux gens qui m'entourent, mille excuses Ka, ce fut très très long... Merci à mes trois fils, Damien, Maxence et Thierry d'être ce qu'ils sont

et de me prendre comme je suis, sincèrement je suis bien avec vous. Merci à mes parents qui m'ont appris à être curieux, critique et à respecter les gens qui m'entourent. J'espère ne pas vous avoir déçu trop souvent.

Je ne voudrais oublier tous les auteurs que j'ai cité qui m'ont torturé les méninges et qui m'ont fait découvrir tant de facettes de l'être humain et de la géographie, vraiment un beau domaine d'étude. Si on se fie à la qualité des articles publiés dans ce domaine, la recherche se porte bien et les jeunes chercheurs savent prendre la relève, il n'y a aucun doute.

La seule chose que je regrette c'est d'avoir sacrifié mes amis pour ce mémoire.

Table des matières

RÉSUMÉ.....	II
RÉSUMÉ LONG.....	III
REMERCIEMENTS.....	VI
TABLE DES MATIÈRES.....	VIII
LISTE DES TABLEAUX.....	X
LISTE DES FIGURES.....	XI
LISTE DES ANNEXES.....	XII
INTRODUCTION.....	2
HYPOTHÈSES ET OBJECTIFS.....	6
<i>Postulat général</i>	6
<i>Hypothèses</i>	6
<i>Objectif général</i>	7
<i>Objectifs spécifiques</i>	7
<i>But du mémoire</i>	8
<i>Structure du mémoire</i>	8
PREMIÈRE PARTIE.....	9
LA COGNITION DE L'ESPACE : ASPECTS CONCEPTUELS.....	9
CHAPITRE PREMIER.....	10
1 L'UTILISATION DES SIG LORS D'UNE ENQUÊTE DE MOBILITÉ.....	10
1.1 ÉNONCÉ DE LA PROBLÉMATIQUE ET DES BESOINS.....	10
1.1.1 <i>Les perspectives de recherche découlant du mémoire</i>	16
1.2 ÉTAT DE LA CONNAISSANCE EN MATIÈRE DE GÉORÉFÉRENCE.....	16
1.2.1 <i>Région floue</i>	17
1.2.2 <i>Positionnement dans l'espace géographique en langage naturel</i>	18
1.3 DESCRIPTION D'UN ESPACE DANS L'ENQUÊTE OD.....	19
CHAPITRE SECOND.....	22
2 LES SCIENCES COGNITIVES.....	22
2.1 INTRODUCTION AUX SCIENCES COGNITIVES.....	22
2.2 MODÈLE GÉNÉRAL.....	26
2.2.1 <i>Carte mentale et modèle cognitif de l'espace</i>	26
2.2.2 <i>Modèle d'acquisition des données géographiques</i>	27
2.3 QU'EST CE QUE L'ON MÉMORISE ?.....	31
2.3.1 <i>Point de repère</i>	33
2.3.2 <i>Réseau</i>	34
2.3.3 <i>Carte en survol</i>	36
2.3.4 <i>Portail</i>	37
2.4 COMMENT LE MÉMORISE-T-ON ?.....	37
2.4.1 <i>Carte mentale</i>	37
2.4.2 <i>Théorie de la mémoire et objet dans l'espace</i>	38
2.4.3 <i>Représentation externe</i>	43
2.4.4 <i>Autres relations</i>	44
2.5 QUAND ON LE MÉMORISE ?.....	44
2.5.1 <i>Apprentissage direct</i>	45
2.5.2 <i>Apprentissage indirect</i>	46
2.5.3 <i>Fréquence et durée</i>	47
2.5.4 <i>La restitution</i>	48
2.5.5 <i>Le rappel des métriques</i>	50
2.6 COMMENT L'EXPRIME-T-ON ?.....	51
2.6.1 <i>Description d'un espace</i>	51
2.6.2 <i>Capacités des systèmes d'information géographique actuels</i>	54

	ix
2.7	NOTES SUR L'ÉCHELLE 55
2.8	CONCLUSION DU CHAPITRE 56
DEUXIÈME PARTIE..... 58	
APPLICATION À UN SIG : ASPECT THÉORIQUE ET PRATIQUE..... 58	
CHAPITRE 3..... 59	
3	LES BASES DE DONNÉES SPATIALES 59
3.1	LES SYSTÈMES D'INFORMATION GÉOGRAPHIQUE..... 60
3.2	INTÉGRATION DES DONNÉES 61
3.3	MODÉLISATION DES DONNÉES 62
3.3.1	<i>Modèle Entités-Relations</i> 62
3.4	MODÈLE COGNITIF ET SIG..... 62
3.4.1	<i>Relation spatiale</i> 64
3.5	INTÉGRATION DES SCIENCES COGNITIVES..... 65
3.6	CONCLUSION DU CHAPITRE 67
CHAPITRE 4..... 69	
4	LA CONSTRUCTION DU SIG 69
4.1	L'INTÉGRATION DES BASES DE DONNÉES..... 70
4.1.1	<i>Les bases de données d'origines</i> 72
4.1.2	<i>Les modalités de l'intégration</i> 74
4.1.3	<i>Le résultat de l'intégration</i> 75
4.2	LE MODÈLE DE DONNÉES 76
4.2.1	<i>Tables de base</i> 77
4.2.2	<i>Ajustement cognitif</i> 79
4.2.3	<i>Ajustement géométrique et cognitif</i> 81
4.2.4	<i>Ajustement nominatif</i> 82
4.3	L'INTERFACE D'INTERROGATION..... 83
4.3.1	<i>Les langages</i> 84
4.3.2	<i>Le modèle d'interface</i> 86
CONCLUSION..... 89	
BIBLIOGRAPHIE..... 93	
ANNEXES..... 105	

Liste des tableaux

Tableau 1. Particularités des nomenclatures de lieux	20
Tableau 2. Pays d'origine et département de rattachement des auteurs des actes du colloque du COSIT 2001 à Morro Bay, Californie.....	25
Tableau 3. Les éléments de repères en milieu urbain selon Lynch (1960)	31
Tableau 4. Les éléments de la carte mentale selon Siegel et White (1975).....	32
Tableau 5. Le concept de portail selon Chown (2000)	32
Tableau 6. Types de représentation physique d'après Paivio (Sasse, 1997)	39
Tableau 7. Catégories et niveau hiérarchique	40
Tableau 8. Mode d'acquisition de l'espace	45
Tableau 9. Description d'un espace	52
Tableau 10. Taxonomie des conflits dans les systèmes d'information géographique	75
Tableau 11. Liste des grands générateurs fréquentés plus de 1000 fois lors de l'enquête OD 2001 dans la région de Québec tels que nommés dans la base de données de l'enquête (9 de 104 657 destinations).....	76
Tableau 12. Langages de requête d'après Aufaure et Trépied (2001)	85

Liste des figures

Figure 1. Territoire de l'enquête OD de la région de Québec en 2001.	11
Figure 2. Définition de « vagueness »	18
Figure 3. Les disciplines qui s'intègrent pour définir les sciences cognitives au sens large	23
Figure 4. L'espace chorotaxique aux espaces cognitifs	28
Figure 5. Facteurs influençant la construction de la cognition spatiale	30
Figure 6. Modèle de la mémoire selon Baddeley.....	38
Figure 7. Exemples de problèmes engendrés par la hiérarchisation spatiale	42
Figure 8. Hiérarchie spatiale.....	43
Figure 9. Cycle cognitif de l'acquisition et du raffinement de la connaissance	48
Figure 10. Résolution de problème dans la mémoire et restitution de l'information.....	49
Figure 11. Définition des types d'espaces.....	56
Figure 12. Modèle de Mennis et Peuquet (2000): composante connaissance et composante donnée.....	64
Figure 13. Système de représentation basé sur la représentation spatiale	65
Figure 14. Paramètres intégrés dans le modèle de données tel que représenté dans les sous-systèmes de la mémoire.	66
Figure 15. Schéma de base de données adapté pour faciliter la recherche de lieu lors d'une enquête en s'inspirant d'une approche cognitive.	78
Figure 16. Modèle d'interface usager proposé qui intègre des éléments permettant de retrouver un lieu de façon unique et non équivoque.....	87

Liste des annexes

Annexe 1 – Liste des bases de données intégrées pour l'enquête OD de la région de Québec en 2001 par le RTC et le MTQ	106
Annexe 2 – Modèle de données	107
Annexe 3 – Glossaire	108
Annexe 4 – WebKb.....	110

« La question de l'espace est présente dans toutes les disciplines qui traitent de la connaissance humaine. Nulle théorie cognitive ne saurait omettre de considérer la façon dont les individus « expérimentent » l'espace, dont ils mémorisent et s'en créent des représentations (mentales ou matérielles) et la manière dont ils utilisent ces représentations pour planifier leurs déplacements ou anticiper les déplacements d'autrui. » (Denis, 1997)

Introduction

Ce mémoire s'intéresse aux descriptions qualitatives de la position des lieux d'activité dans l'espace géographique et repose sur des notions de toponymie et de systèmes d'information géographique (SIG). L'objectif consiste à automatiser et à faciliter la localisation des données d'enquêtes portant sur les comportements de mobilité des personnes en milieu urbain. Il s'agit de développer une méthodologie de représentation informatique des lieux d'activité fréquentés par des individus dans un SIG en utilisant les références spatiales nominales, souvent toponymiques, utilisées par les répondants pour décrire les lieux visités. Il repose sur une réflexion à propos des moyens à mettre en œuvre pour construire un outil de localisation qui effectue l'interface entre la perception de l'espace par les individus (géographie cognitive), la description verbale de la localisation des lieux d'activités fréquentés et sa représentation dans un SIG où les éléments doivent être localisés dans un système de coordonnées cartésiennes (x et y) ou géographiques (latitudes et longitudes).

Afin de mesurer la demande en transport, plusieurs organismes, notamment le ministère des Transports du Québec (MTQ), réalisent des enquêtes durant lesquelles les personnes sont interrogées à propos de leurs déplacements et des activités réalisées durant une journée ou une semaine. De même, la réalisation d'enquêtes rétrospectives permettant de recueillir les lieux de résidence et de travail antérieurs des personnes améliorerait notre compréhension des mécanismes et critères de choix de localisations résidentielles. Cependant, pour être efficaces, ces enquêtes doivent fournir une gestion adéquate des références spatiales et temporelles telles que révélées par les répondants et des moyens appropriés pour manipuler les informations permettant de documenter les activités d'un grand nombre de personnes (souvent des milliers).

En offrant un lien direct avec la technologie des bases de données relationnelles et une manipulation efficace des localisations spatiales, les SIG apportent une solution partielle pour réaliser ce genre d'études. Cependant, bien que les métriques cartésiennes soient utiles pour dresser des cartes précises et déterminer des temps absolus, les personnes interrogées sont généralement incapables de livrer l'information spatiotemporelle sous cette forme et avec un tel degré de précision. En conséquence, il faut développer des mécanismes de transformation effectuant l'interface entre les localisations qualitatives exprimées par les individus et le besoin

de précision numérique (quantitative) requise pour les repérer dans un SIG. C'est cette nécessité de transformer les références spatiales nominales (adresses civiques, raisons sociales, établissements, coins de rue, toponymes, odonymes, lieudits, points de repère, etc.) en coordonnées géographiques utilisables dans un SIG qui constitue la motivation essentielle de ce mémoire.

C'est par les modèles cognitifs de la perception de l'espace géographique que nous aborderons l'identification par les individus des entités géographiques remarquées, leur façon de concevoir et de hiérarchiser leur importance et leur manière d'exprimer les interrelations entre les lieux. Ces modèles cognitifs permettent d'étudier les similitudes et les dissimilitudes des représentations des phénomènes géographiques en relation avec des points de repère spécifiques et à travers le sentiment d'appartenance des personnes (par exemple, identification à un quartier, à une municipalité ou à une paroisse). Idéalement, une procédure d'enquête de mobilité devrait être capable de s'accommoder de toutes les formes de références spatiales usuelles utilisées par les répondants et les traduire en coordonnées cartésiennes utilisables dans un SIG avec le maximum de rapidité et un minimum d'intervention humaine. Les SIG actuels n'offrent pas ce type de service, ce qui limite fortement leur utilisation pour réaliser des enquêtes verbales de mobilité des personnes sans recours à des technologies de positionnement global ou de télémétrie. Par ce mémoire, nous répondons à ce besoin de traduction des références spatiales pour les enquêtes par questionnaire.

Une innovation émergente dans la recherche en géographie sociale

L'évolution rapide de la société amène la recherche en études urbaines à se pencher de plus en plus sur la dimension spatiotemporelle des comportements quotidiens de mobilité (migrations pendulaires et utilitaires) et les choix de localisation à plus long terme (choix résidentiels ou positionnement des établissements). Les principaux changements observés dans la société concernent un accroissement des contraintes familiales (divorce, garde partagée, familles reconstituées), une diversification du marché de l'emploi (emplois contractuels, instabilité professionnelle, emplois multiples, travail à temps partiel), une plus grande segmentation du marché du logement, la variation des coûts de transport selon le mode et l'émergence de nouvelles approches et technologies (télétravail, commerce électronique, livraison juste à

temps) qui influencent le comportement des individus et des ménages (Thériault *et al.*, 2001). De plus, la diversification des formes structurelles des ménages (deux parents en emploi, familles éclatées ou reconstituées) a vraisemblablement entraîné certaines modifications dans la mobilité des individus (allongement des distances parcourues) et un ajustement des habitudes de consommation (Vandersmissen, 2000). Enfin, avec l'avènement des nouvelles technologies, l'espace est en pleine mutation et les géographes sont amenés à redéfinir certains concepts notamment ceux de lieu (réel ou virtuel), de distance, de connexité et d'accessibilité.

En géographie urbaine et économique, les chercheurs ont longtemps travaillé avec des données agrégées publiées à l'échelle des secteurs de dénombrement ou de recensement. Cependant, la technologie informatique actuelle permet de constituer des bases de données désagrégées décrivant les comportements individuels (personnes, ménages et entreprises) et de les exploiter pour étudier les processus urbains à l'échelle désagrégée où ils surviennent, permettant ainsi de bonifier le caractère expérimental de leur démarche scientifique. Bien que cette évolution pose quelques problèmes d'éthique et de confidentialité des données, cette échelle d'étude permet de documenter les comportements et d'identifier les facteurs décisionnels qui influencent la localisation et la mobilité des personnes, des ménages et des entreprises. Même si elles permettent de réduire l'erreur écologique, qui survient quand on construit des inférences à propos des processus individuels à partir de mesures agrégées (notamment causées par le problème de l'aire spatiale modifiable), ces études désagrégées introduisent une possibilité d'erreur atomistique si elles ignorent le contexte de l'environnement local et social en postulant que les seules caractéristiques de l'individu suffisent pour expliquer son comportement (Courgeau, 2000).

De plus, certaines études en géographie humaine et en démographie ont recours aux enquêtes longitudinales et rétrospectives pour collecter des informations sur les biographies individuelles afin de retracer la séquence temporelle « d'évènements » ou « d'états » qui ont marqué l'existence des personnes interrogées (Courgeau, 2000; Séguin *et al.*, 2000) en considérant trois séries d'évènements interreliés : la trajectoire familiale, la trajectoire professionnelle et la trajectoire résidentielle. La comparaison des évènements de ces trois trajectoires permet une validation des données à référence spatiale, mais doit être appuyée par des outils qui permettent de tenir compte des changements d'état de l'espace à travers le temps.

À partir des années 1970, le développement des SIG a favorisé l'émergence d'une métascience qui ne traite pas du monde géographique, l'objet de la géographie, mais de l'information relative au monde géographique (Couclelis, 1999). Les outils se sont développés ainsi que la science de l'information géographique (Goodchild, 1990), pour initier, dans les années 1980, une révolution dans la représentation spatiale. Bien que des progrès soient encore souhaitables, il est ainsi devenu plus facile de produire des cartes et des modèles spatiaux, d'interroger et d'analyser la distribution des phénomènes dans l'espace et de créer de nouvelles formes de représentation cartographique.

Cependant, ces nouveaux outils ont été l'objet de plusieurs critiques malgré les progrès technologiques qu'ils ont permis. Les objections ont été nombreuses au début des années 1990 (Sheppard *et al.*, 1999; Schuurman, 2000), en commençant par un débat, devenu historique, entre Taylor (1990) et Openshaw (1991) qui opposaient une critique du positivisme inhérent à cet outil et de la nature réductionniste des SIG face à une perspective visionnaire de cette technologie pouvant entraîner une libération et une unification de la géographie. Depuis, le débat a progressé, l'outil progresse, son utilisation surtout...

En tant que technologie d'information, les SIG sont utiles pour aider à structurer les données spatiotemporelles afin de résoudre des problèmes concrets liés à la distribution des phénomènes sur le territoire et aux changements qui y surviennent. Les acteurs, tant utilisateurs que concepteurs, de même que les méthodes employées doivent répondre à certaines exigences :

- Le concepteur de la base de données doit avoir une compréhension empirique des phénomènes géographiques représentés dans le SIG afin de bien représenter les aspects sémantiques (entités, relations, processus, faits) pertinents du domaine d'application.
- Le modèle de données sur lequel repose la base de données doit être conforme à une des structures spatiotemporelles imposées par la technologie, ce qui limite la gamme des ontologies utilisables.
- L'utilisateur du SIG doit sélectionner et extraire les données nécessaires pour accomplir une tâche particulière grâce à une compréhension adéquate de la question, du domaine d'application et de l'information requise pour résoudre le problème.
- Le contexte social et culturel (incluant la discipline) dans lequel s'effectue la recherche détermine, entre autres, la manière dont la question sera posée et quelles formes de réponse sont jugées acceptables (Couclelis, 1999).

Les systèmes d'information géographique permettent de gérer efficacement des informations géographiques s'ils ont comme base une structure de données qui correspond à la problématique de recherche et non à une structure purement informatique. Une approche cognitive permet de mettre en perspective le problème dans un contexte de compréhension de l'apprentissage de l'espace, de sa mémorisation, de ses représentations et de sa description en langage naturel par les individus.

Hypothèses et objectifs

Postulat général

L'utilisation d'une indexation cognitive des localisations exprimées en langage naturel lors d'une enquête, doublée d'une interface de requête appropriée, permettent de rehausser la validité (exactitude et précision) des références spatiales obtenues afin de réaliser des études de géographie urbaine utilisant les SIG.

Hypothèses

Notre première hypothèse est que l'ajout d'informations corollaires dans la base de données afin de répliquer la logique des références spatiales relatées par les personnes permet de pallier certaines lacunes inhérentes aux bases de données institutionnelles et aux langages informatiques, parfois inadaptés pour refléter tous les attributs phénoménologiques et relationnels de l'espace géographique.

Les bases de données institutionnelles contiennent des informations et des nomenclatures officielles et actuelles. La référence à l'espace se limite à l'adresse civique, aux codes postaux et à la municipalité. Et rarement nous retrouvons les anciennes appellations des noms de rues, de villes ou encore le nom des édifices. C'est par une approche cognitive que l'on peut comprendre les modifications nécessaires à une base de données pour qu'elle soit plus compatible avec la description des lieux en langage naturel. Cette approche fait référence à la façon dont les gens saisissent et expriment les lieux dans l'espace, avec des vocables périmés, des localisations floues en faisant référence à des points de repères sans adresse civique mais situés à proximité de ces derniers.

Les langages informatiques, tels que le « structured query language » (SQL) n'offrent pas, à priori, la souplesse requise pour retrouver des références spatiales utilisant des concepts relatifs comme « près de », « sur », « devant », ni pour fournir toute l'information utile pour valider une référence spatiale. Il faut combler ces lacunes par une interface qui permettra de rechercher un lieu sous différents vocables et d'obtenir toute l'information pour s'assurer de la validité du lieu géoréférencé.

Notre seconde hypothèse est qu'une structure entité-relation est suffisante pour opérer, entretenir et exploiter une telle base de données intégrée dans un SIG vectoriel opérant en mode géorelationnel.

Objectif général

L'objectif général consiste à élaborer un SIG de la région de Québec qui soit représentatif de la distribution des lieux d'activité en 2001. Fondé sur un schéma logique de base de données de type entités-relations, ce dernier est conçu à l'échelle des édifices et des repères physiques. Il vise à accélérer la reconnaissance, par un opérateur, des endroits désignés par les répondants d'une enquête. Ceux-ci s'expriment en langage naturel et il faut déterminer rapidement et de manière adéquate les références aux objets uniques qui fournissent les coordonnées requises pour positionner les lieux d'activité sur une carte et dans le SIG. Le but ultime du système consiste à accroître la qualité des localisations en termes de précision et d'exactitude.

Objectifs spécifiques

Les objectifs spécifiques sont :

- Amorcer une réflexion pour ajouter des faits géographiques corolaires afin que le SIG soit opérationnel dans le cadre d'une enquête OD où les localisations sont, d'une part, exprimées de façon nominale par des personnes ayant une conception de l'espace qui diffère de la géométrie euclidienne et, d'autre part, colligées par des enquêteurs non initiés aux SIG et aux langages de requête informatique (par exemple, SQL). En fait, il s'agit d'adapter le système informatique pour favoriser un meilleur dialogue entre les répondants et les enquêteurs afin d'identifier la localisation la plus précise et exacte possible des lieux d'activité fréquentés.
- Élaborer une typologie des références spatiales naturelles utilisées par les gens et effectuer une première tentative de schématisation de ces dernières dans une base de données relationnelle exploitée dans un SIG.

- Structurer, intégrer et transférer des tableaux de données disparates, d'un modèle tabulaire vers un modèle relationnel, afin de les gérer et de constituer un outil informatisé de localisation assez efficace pour être utilisé en temps réel lors d'une enquête.

Le SIG construit dans le cadre de ce mémoire sera ensuite étendu afin de constituer un répertoire spatiotemporel de l'évolution de lieux d'activité sur le territoire. Ce dernier sera développé au CRAD (Centre de recherche en aménagement et développement) durant les prochaines années.

But du mémoire

Le but de ce mémoire consiste à améliorer le processus d'identification des lieux visités par les répondants des enquêtes de déplacement, ces derniers utilisant le langage naturel pour les désigner, le système permettant de réaliser un positionnement fiable et unifié (un lieu, plusieurs vocables) selon des principes compatibles avec la technologie des SIG. L'approche par les sciences cognitives pour comprendre le traitement de l'information spatiale présente une solution intéressante pour la structuration de la base de données afin de créer une interface adéquate pour la saisie des références positionnelles et pour l'analyse des résultats par complémentarité cognitive.

Structure du mémoire

Le mémoire s'articule autour de deux grandes parties. La première partie est théorique et présente les sciences cognitives et les théories sur l'acquisition de l'information géographique, le stockage de l'information dans la mémoire et la restitution de l'information en langage naturel. La seconde partie traite en premier lieu des systèmes d'information géographique et de certaines réflexions liées à cette technologie. Puis on aborde la problématique de la construction de la base de données en s'appuyant sur les théories énoncées précédemment. Nous proposons dans un dernier chapitre une interface usager afin de permettre une utilisation optimale de la base de données pour les entretiens téléphoniques lors des enquêtes origine-destination.

Première partie

La cognition de l'espace : Aspects conceptuels

Chapitre premier

1 L'utilisation des SIG lors d'une enquête de mobilité

1.1 *Énoncé de la problématique et des besoins*

Ce mémoire veut contribuer au développement des SIG grâce à un apport méthodologique visant à accroître leur potentiel d'application en géographie humaine. Plus spécifiquement, il s'agit, au plan opérationnel, de réduire l'écart entre les géoréférences de type nominal et les coordonnées géographiques essentielles pour les SIG en construisant des outils capables de manipuler divers types de références toponymiques et de détecter de manière semi-automatique et efficiente (assistée par ordinateur) les lieux concernés afin de produire des coordonnées semblables pour des lieux identiques. Bien que la multiplicité des nomenclatures utilisées par les répondants doive être conservée, il faut que le système de géolocalisation soit capable de déduire l'unicité des lieux référencés lorsque ces derniers sont identiques. Par exemple, si on effectue une localisation à la granularité de l'édifice, les personnes qui affirmeront avoir effectué des achats chez Sears à Sainte-Foy, à Place Laurier ou au 2700 boulevard Laurier devront toutes obtenir la même référence spatiale puisqu'il s'agit du même immeuble.

Le terrain d'étude est la région métropolitaine de Québec telle que définie par l'enquête origine-destination (OD) du MTQ et du RTC (réseau de transport de la Capitale) en 2001 (figure 1). Le produit visé est un SIG géorelationnel en mode vectoriel qui servira, dans un

au système afin de compléter la représentation géométrique de l'espace par une dimension nominale laissant place à un certain degré de flou (précision en fonction de l'échelle de référence; par exemple : quartiers versus coins de rue).

Il s'agit de doter le SIG d'une interface entre l'espace perçu (traduit sous une forme nominale) et l'espace cartographique afin d'ajouter à la carte une dimension sémantique compatible avec la perception de l'espace et sa traduction en langage naturel. Par exemple, à partir de la phrase « j'ai fait des achats dans la pharmacie située au nord-ouest de l'intersection de la rue Cartier et du boulevard René-Lévesque à Québec », on doit pouvoir identifier l'édifice où est situé ce commerce, son adresse civique, la bannière commerciale, en plus de fournir une position sur le territoire. Pour chaque phénomène, il faut donc prévoir plusieurs modalités d'identification et les gérer dans un modèle unique. Ces modalités permettront de géoréférencer des descriptions complémentaires d'une même localisation (édifices, noms des commerces qui y sont localisés, types de produits, adresses civiques, quartiers où ils sont situés). Dans le cadre de cette recherche, nous nous limiterons à l'identification des lieux d'activité de l'enquête OD, ce qui simplifie la problématique puisque nous n'avons pas à rendre compte de la description du trajet effectué pour s'y rendre. Utilisant le mode relationnel, le modèle de données permettra d'établir un lien entre la sémantique des localisations et leur positionnement géographique.

Des études antérieures ont identifié les difficultés inhérentes à la collecte de données spatiotemporelles. Il s'agit de problèmes liés à la mémoire qui limitent la fiabilité des données recueillies ou encore de variabilité des indications utilisées pour désigner les lieux. Séguin *et al.* (2000, p. 8) proposent une procédure de validation *in situ* : « L'utilisation interactive d'un système géomatique pourrait constituer un moyen efficace pour assister le processus de remémoration de localisation grâce à une visite virtuelle des lieux, en plus de fournir un outil supplémentaire d'interaction entre l'enquêteur et le répondant ». Thériault *et al.* (2000) proposent de regrouper, dans une interface-usager conviviale, différents types de géoréférences et des procédures d'aide à la localisation. D'autres auteurs proposent d'intégrer dans les bases de données une série d'informations corolaires pour faciliter cette recherche (photos, photographies aériennes, ...). Une synthèse de ces propositions d'extension du système de saisie des localisations permettra d'évaluer leur utilité pour notre application.

Le deuxième volet du mémoire est méthodologique. Il s'agit de structurer des attributs géographiques (objets et identifiants) dans une base de données d'un SIG en mode géorelationnel et de concevoir une interface permettant de formuler des requêtes (recherche de localisation) en temps réel. Cette application utilise les données requises pour réaliser l'enquête OD de la région de Québec durant l'automne 2001.

Depuis 1977, la STCUQ (Société de transport de la Communauté Urbaine de Québec, maintenant RTC) réalise des enquêtes OD dans la région de Québec afin de connaître les comportements récurrents de déplacement durant une journée de semaine, d'évaluer la demande en transport ainsi que d'étudier sa distribution dans l'espace et sa répartition dans le temps. Depuis, cette enquête est répétée avec une fréquence quinquennale (1981, 1986, 1991, 1996 et 2001). L'enquête de 2001 comporte certaines modifications, dont deux qui nous intéressent plus particulièrement par rapport au géocodage et à la saisie de l'information.

La première nouveauté est une augmentation de la précision du géocodage. L'information géographique est saisie et codée en temps réel lors de l'entrevue téléphonique de 15 minutes. En 1977, 1981 et 1986, la plupart des géoréférences ont été spécifiées en utilisant un découpage spécifique (modifié à chaque enquête) du territoire en quelques centaines de zones. Depuis 1986, celui-ci se faisait en utilisant les codes postaux à six positions (résolution de la façade d'îlot en milieu urbain, au Canada). Dorénavant, l'information sera géoréférencée à l'échelle de l'édifice. Ce changement d'échelle est essentiel pour alimenter des procédures de microsimulation des comportements de mobilité et pour affiner l'étude des décisions individuelles. Il permet d'affranchir l'étude des dynamiques urbaines de la «tyrannie des zones» (Spiekermann et Wegener, 2000) en autorisant l'établissement de relations topologiques avec les réseaux de transport. De plus, cette désagrégation des localisations permet de réduire les difficultés créées par les changements fréquents des découpages zonaux. Le découpage des zones doit être modifié à chaque enquête pour tenir compte des développements survenus depuis l'enquête précédente, élargir le territoire transformé par l'étalement urbain et les changements notables aux réseaux de transport (routes, circuits de transport en commun, etc.). On se bute donc invariablement sur le problème de l'aire spatiale modifiable (Openshaw, 1984) dès qu'on désire comparer les situations à deux dates ou étudier des transects historiques.

La deuxième nouveauté concerne la saisie et le codage de l'information géographique en temps réel lors de l'entrevue téléphonique de 15 minutes. Auparavant, la transcription des informations nominales était effectuée en temps réel, mais le codage spatial était réalisé en différé. Cette dernière opération durait plusieurs mois et le déphasage des deux opérations ne permettait pas d'effectuer un éventuel rappel téléphonique pour obtenir des précisions afin de résoudre les ambiguïtés, le temps écoulé étant trop long pour que la personne enquêtée puisse se souvenir des localisations mal décrites. Il s'en suivait une perte brute d'information et une diminution générale de la qualité des références géographiques qui devaient souvent être ajustées à posteriori, sans information corolaire. En effet, les adresses ou localisations imprécises ne peuvent être validées de la même façon (par rapport à la collecte des dates) comme suite à la non-disponibilité d'information corolaire (Séguin *et al.*, 2000).

En 2001, l'enquête téléphonique a été exécutée par une trentaine d'enquêteurs qui disposaient d'une base de données incluant plusieurs listes de références spatiales potentielles dans le but de diminuer au maximum les risques d'erreur. Ces listes ont été utilisées à des fins indicatives, sans restreindre les choix possibles, mais comme un outil permettant une validation immédiate de la connaissance des lieux fréquentés dans la base de données du SIG, une identification ad hoc des ambiguïtés et une référence de base pour détecter les dénominations trop vagues pour être localisées (par exemple, plusieurs écoles dans le même quartier) ou simplement erronées (par exemple, intersection de deux rues parallèles). La procédure d'enquête n'a pas limité les possibilités de réponses pour ce qui est des localisations de lieux fréquentés par le répondant durant la journée de l'enquête. Comme suite aux expériences réalisées par le MTQ, au printemps 2000 dans la région de Trois-Rivières, cette première étape permet de générer plus de 95% de localisations exactes durant l'entrevue. Les autres lieux sont traités lors d'une seconde étape, réalisée le soir même, par une équipe de professionnels qui utilise le SIG afin de localiser les références géographiques résiduelles et de les ajouter dans la banque de données. L'opération d'un tel outil requiert un SIG pour organiser et représenter les données et d'une interface-usager d'interrogation pour effectuer le pairage entre les références nominales et les lieux déjà répertoriés dans le système. La saisie *in situ* a permis de valider les données de localisation durant l'enquête OD de la région de Québec qui a été réalisée de septembre à décembre 2001.

Pour construire le SIG nous disposons de plusieurs bases de données de sources diverses, dans tous les sens du terme. Les problèmes, reliés aux échelles discordantes, aux dates de validité diverses, aux types de géoréférences hétérogènes (points, lignes, zones avec ou sans limites fixes) et à la validation des données, sont complexes à résoudre. Ils constituent un des obstacles importants pour la mise en œuvre d'un répertoire de géolocalisation à l'échelle d'une région métropolitaine de la taille de Québec. Qui plus est, beaucoup de références spatiales varient avec le temps (apparition, disparition d'édifices, changements de noms des rues, des commerces, ...). Il faut donc prévoir des mécanismes permettant d'identifier la période de validité de chaque référence et d'effectuer des liens entre les variantes successives de nomenclature.

Les solutions apportées pour la localisation des données spatiales devront être automatisables afin de permettre des mises à jour fréquentes des données durant les prochaines décennies. Le système doit donc être conçu afin de pouvoir évoluer avec le territoire, au gré des ouvertures et des fermetures d'institutions et de commerces, afin de refléter l'évolution des lieux d'activités, tout en gardant une trace historique des états précédents. Cette information est structurée dans un modèle entité-relation (ER), ce qui permet de formuler des requêtes complexes afin d'identifier les variantes toponymiques utilisées pour désigner un même lieu. Ce type d'approche facilitera la mise à jour des données et le suivi des modifications à l'utilisation du sol.

Cette structure de données de type ER n'est pas nouvelle, les premières références en informatique datent du début des années 1970 (développée par Codd chez IBM, (1970)). Cependant, à ce jour, ce type d'approche a été très peu utilisé pour réaliser des études en géographie urbaine et économique. Notre cadre de référence porte essentiellement sur l'étude des comportements de mobilité des individus pour vaquer à leurs activités quotidiennes. Toutefois, nous n'abordons pas cette problématique directement dans le mémoire en réalisant une étude de cas spécifique. Il nous faut néanmoins situer de façon précise le contexte d'utilisation du SIG que nous allons élaborer afin de répondre aux besoins de l'enquête OD, de manière immédiate, mais aussi, dans un contexte plus large pour appuyer les recherches effectuées au CRAD axées sur l'analyse des comportements de mobilité des personnes et des

ménages dans le cadre de recherches financées par les grands travaux du CRSH et le RCE (Réseau de Centres d'Excellence) Géoïde (La géomatique pour des décisions éclairées).

Il est également possible d'ajouter de l'information dans la base de données afin de mieux refléter l'organisation résiliaire (en réseau) des objets géographiques (Gerevini et Renz, 1998). Ces auteurs proposent une nouvelle technique pour ajouter aux relations topologiques (RCC-8) une relation de dimension afin de qualifier la relation entre les objets. En rehaussant la dimension sémantique de l'information, ce type de proposition ajoute des faits dérivés dans la base de données permettant d'assouplir la recherche des références spatiales, en plus de produire des résultats plus faciles à valider grâce à l'obtention de faits complémentaires. Par exemple, une requête visant l'ensemble des pharmacies sur un territoire donné permettra une validation ad hoc plus robuste si on ajoute l'identification du coin de rue le plus rapproché au nom de l'entreprise et à l'adresse civique qui sont déjà consignées dans la base de données. Certains prétraitements visant à établir des relations de concordance (par proximité ou par connexité) sont réalisables dans le SIG et permettent de produire des données corolaires compatibles avec la perception de l'espace par les répondants, facilitant ainsi l'identification des lieux d'activité.

1.1.1 Les perspectives de recherche découlant du mémoire

Cette base de données constitue la première phase de mise en œuvre d'un vaste SIG pour la région de Québec. Ce SIG est alimenté par plusieurs projets de recherche du CRAD. Dans un horizon de cinq ans, une fois restructurée dans un modèle orienté objet (OO), cette base de données servira à alimenter un simulateur multi agents qui «permettra d'étudier les interrelations entre les choix résidentiels des ménages, les décisions de localisation des entreprises et les comportements de mobilité des individus» (Thériault *et al.*, 2001).

1.2 État de la connaissance en matière de géoréférence

La géoréférence est définie comme la représentation de la localisation des entités du monde réel à l'intérieur d'un cadre spatial et dans un système de coordonnées particulier (Lo et Yeung, 2002). L'objectif de la géoréférence est de donner un cadre spatial formel et univoque dans lequel les entités du monde réel sont mesurées, décrites, enregistrées et analysées.

Nous allons nous attarder ici à deux problèmes liés à la géoréférence. Beaucoup d'entités géographiques ont une frontière floue (Campari, 1996) et la description dans un langage naturel des espaces géographiques est rarement précise. Si on fait abstraction des problèmes de projection cartographique et d'erreurs de positionnement, la question du décalage sémantique entre les références de positionnement dans les SIG et les références spatiales utilisées par les individus est complexe, mais peut être simplifiée si on s'appuie sur les connaissances en cognition spatiale.

1.2.1 Région floue

Le concept de région floue a été largement traité en géographie et en raisonnement spatial qualitatif (Erwig et Schneider, 1997; Worboys, 1998; Stell, 1999; Winter, 2000; Cohn et Hazarika, 2001; Duckham *et al.*, 2001; Hazarika et Cohn, 2001; Kulik, 2001; Roy et Stell, 2001; Zhan, 2001; Bittner et Stell, 2002). Ce n'est pas tant l'objet qui est flou que sa frontière. Hazarika (2001) définit très bien le terme « vagueness » dans lequel le flou peut être dû à une imprécision, à l'ignorance de localisation de cette frontière, ou à une variation temporelle ou thématique (figure 2). Il est possible aussi que cette imprécision soit due à l'objet même, on parle alors de flou intrinsèque comme par exemple le « sud du Québec ».

Un objet avec une frontière vague est aussi appelé un « objet vague ». La localisation d'une frontière peut-être inconnue, ce qui a comme résultat une information incomplète. Même si certaines divisions administratives sont bien délimitées dans l'espace, les gens ne connaissent pas toujours leurs frontières. Les gens ont une perception floue des limites spatiales des zones, par « ignorance » (municipalités, quartiers, paroisses, arrondissements, etc.), même quand ces dernières sont découpées en unités spatiales discrètes de type administratif, ce qui peut amener un biais lors de la description de la localisation d'une entité géographique en référence à ces découpages. Il en est de même pour les points de repère lors de la navigation, « dès que l'on s'approche d'un point de repère il s'opère un changement dans la granularité et nous pouvons facilement identifier ce qui entoure le point de repère » (Hazarika et Cohn, 2001, p. 99).

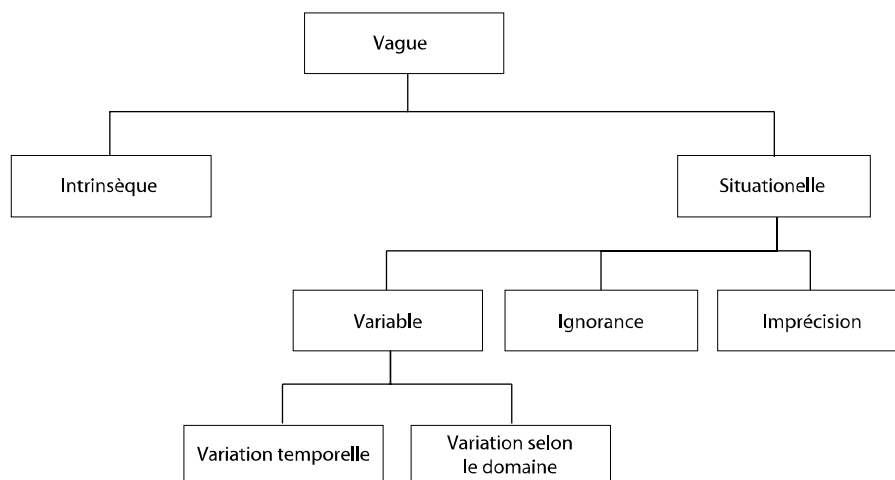


Figure 2. Définition de « vagueness » d'après Hazarika & Cohn (2001).

1.2.2 Positionnement dans l'espace géographique en langage naturel

Le positionnement dans l'espace géographique utilise deux systèmes totalement différents. Le premier mode d'expression de la localisation utilise une structure de langage naturel et des prépositions verbales pour désigner des lieux ou exprimer leur position relative. L'autre, plus récent, est fondé sur des principes de géométrie et repose sur des bases mathématiques de nature cartésienne. Pour spatialiser les enquêtes dans un SIG, la difficulté consiste à transformer des données recueillies dans le mode verbal afin de produire des références géométriques. Les expressions « en avant », « en arrière », « à gauche » et « à droite » sont couramment utilisées pour désigner des positions relatives des objets dans l'espace. L'interprétation de ces relations dépend du point de vue du locuteur (sa propre position par rapport aux deux objets) au moment où elles sont formulées, ce qui rend leur interprétation dépendante de cette information contextuelle. Une partie du problème de conversion provient de la grande diversité des formes d'expression et du caractère souvent conjonctuel des expressions verbales. Ainsi, l'expression « près de » ou « à deux minutes de » n'aura pas la même signification pour un automobiliste et un piéton.

Dans l'état actuel de la technologie, les SIG supportent très peu ce type d'expression. Plusieurs recherches, principalement en linguistique (Levinson, 1996; Ligozat, 1998; Eschenbach, 1999; Grabowski, 1999; Talmy, 2001), en psychologie (Tversky et Lee, 1998; Werner et Habel, 1999; Werner et Schmidt, 1999) et en géographie (Egenhofer, 1992; Mark et Egenhofer, 1994;

Egenhofer, 1997) tentent de formaliser le langage pour permettre une meilleure interface entre les exigences d'un SIG et la «souplesse» du langage courant. C'est cette approche que nous retenons pour construire notre interface. Il faut établir un compromis opérationnel entre la versatilité du langage courant et la rigidité formelle des systèmes mathématiques afin d'établir une certaine diversité de relations de correspondance entre les formes les plus précises du langage naturel et les systèmes de référence numériques.

Eschenbach (1999) propose d'utiliser une approche cognitive afin de distinguer le *système de référence spatiale* du *cadre de référence spatiale* pour analyser la représentation sémantique de l'espace dans le langage courant, les modes de formation de la connaissance de l'espace, de même que sa perception. La structure de l'espace ne peut être perçue ou exprimée que par l'intermédiaire d'entités concrètes. Il existe deux ordres de structures spatiales : un ordre concret qui repose sur des relations de position entre des entités réelles et un ordre abstrait qui utilise une structure géométrique continue pour positionner les objets dans un cadre de référence. Ainsi, l'expression «système de référence spatiale» est utilisée pour désigner une collection d'entités concrètes distribuées dans l'espace, tandis que le terme «cadre de référence spatiale» désigne une collection d'objets informatiques qui modèle la structure géométrique du système de référence spatiale.

1.3 Description d'un espace dans l'enquête OD

Durant une enquête OD, les répondants indiquent des localisations à travers un système de référence spatiale (tableau 1) basé sur des toponymes, en utilisant des adresses civiques d'édifices fréquentés, des odonymes, des intersections désignées par les rues qui se croisent, des noms d'édifices, des noms d'établissements (éventuellement localisables dans des édifices), ou encore des entités zonales dont la délimitation est plus vague (limites floues): des quartiers, des paroisses, d'anciennes municipalités, etc.

Tableau 1. Particularités des nomenclatures de lieux

Nomenclature des lieux en langage naturel		Exemples	Difficultés
Adresses civiques		975 rue Belvédère	Mauvais générique, absence de particule de liaison, utilisation d'ancien nom de rue
Raisons sociales		Saint- Hubert	Occurrences multiples dans la région et absence du générique Rôtisserie
Établissements		Place de la Cité	Chapeaute plusieurs toponymes
Toponymes		Complexe G	Parallélisme (Dorion et Poirier, 1975) : toponymes multiples pour un même édifice (Ex. Marie-Guyart)
Odonymes		Blvd Saint-Cyrille	Ancien nom de rue, absence de direction et mauvais générique
Description qualitative en référence à d'autres objets	Ponctuel	En face de la statue du bonhomme Carnaval	Utilisation de point de repère et de la position relative
	Linéaire	Sur la rue Saint-Jean	Ambiguïté concernant le segment de rue visé
	Zonal	Dans Saint-Jean Baptiste	Quartier ou paroisse avec frontière floue

À la base, le système visé doit donc établir une correspondance univoque entre ces dénominations désignant des entités du monde réel et le cadre de référence spatiale définissant leur positionnement géométrique. Idéalement, on doit identifier les incohérences durant l'entrevue afin de corriger les erreurs (tableau 1). Dans une seconde étape, le système devra également être capable de distinguer les positions qui correspondent directement à l'entité visée (« dans », « sur », « à », etc.) de celles qui l'utilisent comme référence proximale (« près de », « à coté de », « connecté à », etc.) ou positionnelle (« devant », « derrière », « à gauche », « à droite », « au nord », « à l'est », etc.).

Une enquête sur la localisation d'un lieu a ses particularités. Dans l'enquête OD du MTQ, le répondant doit identifier un lieu d'activité avec une brève description et non un trajet, ce qui simplifie beaucoup la description en langage naturel de la localisation du lieu d'origine ou de la destination. De plus cette enquête est téléphonique. Le répondant doit donc décrire l'espace

sans utiliser le langage gestuel et il ne peut pas visualiser le niveau de compréhension de la description de la localisation du lieu par l'enquêteur. L'effort cognitif lui incombe donc en très grande partie. Il n'y a pas non plus de possibilité de valider à l'aide d'une carte. C'est donc le langage qui est le seul lien de communication entre l'enquêteur et le répondant.

Nous abordons dans le prochain chapitre les sciences cognitives et leur apport à la compréhension des différentes étapes de l'acquisition de l'information géographique jusqu'à la restitution en langage naturel de la description d'un lieu.

Chapitre second

2 Les sciences cognitives

Dans ce chapitre, nous abordons quelques principes de cognition dans un contexte historique et interdisciplinaire en nous concentrant sur la théorie des sciences cognitives. Ceci nous permet d'appréhender la problématique de la mémorisation et de la restitution de l'information géographique afin de décrire l'identification, la localisation et la description d'un lieu d'activité.

2.1 Introduction aux sciences cognitives

Les sciences cognitives ont pour objet de décrire, d'expliquer et, le cas échéant, de simuler les principales dispositions et capacités de l'esprit humain, soit le langage, le raisonnement, la perception, la coordination motrice et la planification par exemple (Ladrière, 2002).

On peut remonter très loin dans l'histoire pour connaître ses origines, mais on attribue le plus souvent à Alan Turing la « paternité immédiate » de ces sciences en citant deux de ses articles (Turing, 1937, 1950). Le but de ces derniers consistait à donner une définition mathématique précise d'un algorithme ou d'une procédure mécanique. Turing y présentait les conditions d'une expérience consistant à établir une conversation entre un homme et une machine.

C'est aussi à la première cybernétique que l'on doit les bases des sciences cognitives, notamment grâce aux travaux de Wiener (1948). Il effectue une « synthèse des mathématiques pures (théorie de la prédiction statistique), dans le domaine de la technologie (machines à calculer, télécommunications), dans ceux de la biologie et de la psychologie, et jette les bases d'une science nouvelle, à support mathématique, destinée à couvrir tous les phénomènes qui, d'une manière ou d'une autre, mettent en jeu des mécanismes de traitement de l'information » (Ladrière, 2002).

À la même époque eurent lieu les conférences de Macy (entre 1946 et 1953) où, sous l'égide du philanthrope Josiah Macy Jr., les « grands esprits » de l'époque se réunissaient pour définir une science générale du fonctionnement de l'esprit. Déjà moult disciplines étaient présentes (mathématiciens, logiciens, ingénieurs, physiologistes et neurophysiologistes, psychologues, anthropologues et économistes) par l'entremise de gens comme Norbert Wiener, Warren McCulloch, John Von Neumann, Karl Popper et Margaret Mead.

Tolman (1948), lui, publiait sur la cognition spatiale un article, devenu classique, intitulé « Cognitive Maps In Rats And Men » qui est identifié comme l'article qui amena la psychologie à s'intéresser à la relation isomorphique entre la manière de se déplacer dans le monde et la flexibilité mentale produite par les représentations internes ou les cartes mentales selon Devlin (2001).

Aujourd'hui les sciences cognitives sont intégrées dans plusieurs disciplines reliées à la cognition naturelle (philosophie, psychologie, linguistique et neurosciences) et à la cognition artificielle (intelligence artificielle et informatique) (figure 3).

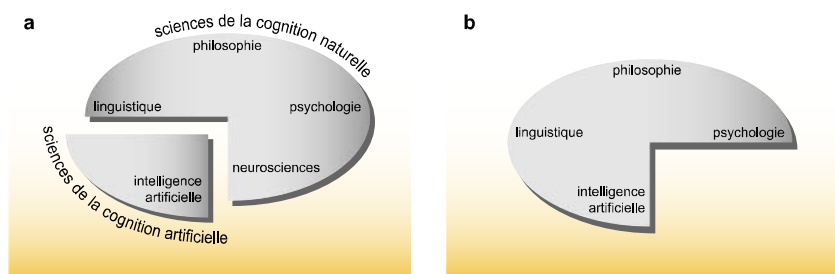


Figure 3. Les disciplines qui s'intègrent pour définir les sciences cognitives au sens large (a) et plus particulièrement les sciences des systèmes de traitement de l'information (b) (source : Encyclopedia Universalis, 2001).

La géographie est aussi présente par l'entremise de chercheurs et de groupes de recherche. Il en est ainsi des chercheurs du National Center for Geographic Information and Analysis (NCGIA) dont le projet Varenus (1997-2001) porte sur les modèles cognitifs de l'espace géographique (Michael F. Goodchild, David M. Mark, Max J. Egenhofer et Karen K. Kemp). Le groupe européen ChoroChronos (1996-2000) orientait ses recherches plus particulièrement sur le temps et l'espace avec une approche cognitive (Andrew Frank, Timos Sellis et Christian S. Jensen). Le groupe allemand « Spatial Cognition Priority Program » (1998-2002) est aussi très présent dans la littérature. Parmi les autres acteurs importants, mentionnons les Autrichiens (Andrew Frank et son équipe), le groupe de l'Université de Leeds dirigé par Anthony Cohn, les Français (Michel Denis et Gérard Ligozat), le projet européen SpaceNet (1994-1998) et, plus récemment, la chaire de recherche du Canada de Geoffrey Edwards à l'Université Laval, intitulée « Cognitive Geomatics - Spatial Referencing at a Crossroads » (2000-2007). Ces projets demeurent fondamentalement multidisciplinaires, en conformité avec la nature même des sciences cognitives¹. La conférence bisannuelle du COSIT (Conference on Spatial Information Theory) est présentement la rencontre internationale qui permet le plus d'échanges dans le domaine des sciences cognitives appliquées à l'information spatiale. Une synthèse de l'origine géographique et scientifique des participants de ce colloque nous permet de saisir qui s'intéresse aujourd'hui à cette science encore en pleine élaboration. Le recensement par départements des auteurs ayant publié dans les actes du colloque de 2001, qui a eu lieu à Morro Bay en Californie aux États-Unis (tableau 2), nous indique que 46% des chercheurs provenaient d'un département d'informatique, 27% de psychologie, et que 16% étaient issus de départements de géographie ou de géoinformatique. Seulement 3% (soit deux auteurs par discipline) proviennent de linguistique, de science cognitive et d'intelligence artificielle. En ce qui concerne la provenance géographique des auteurs, on note une nette dominance des pays occidentaux, avec une forte concentration en Amérique du Nord et en Europe.

¹ Pour une lecture plus approfondie sur le sujet, on peut lire l'article de Mark (1999).

Tableau 2. Pays d'origine et département de rattachement des auteurs des actes du colloque du COSIT 2001 à Morro Bay, Californie

Département		Pays d'origine	
Informatique	46%	USA	34%
Psychologie	27%	Allemagne	24%
Géoinformatique	10%	UK	14%
Géographie	6%	Canada	7%
Linguistique	3%	France	6%
Science cognitive	3%	Autriche	6%
Intelligence artificielle	3%	Nouvelle-Zélande	3%
Géomatique	1%	Danemark	3%
Philosophie	1%	Brésil	3%
		Suisse	1%
Total	100%	Total	100%

Les nombreuses publications scientifiques des dernières années, principalement depuis 1998, démontrent bien l'importance de la recherche dans le domaine. On retrouve des articles sur la cognition spatiale dans les revues scientifiques de diverses disciplines, par l'entremise des revues *IJGIS*, *Computers Environment and Urban systems*, *Spatial cognition and Computation*, *Environment and behavior*, *Geoinformatica*, *Transaction in GIS*, *Journal of visual language and computing*, et *Cognitive psychology* pour ne nommer que celles-ci. Bien entendu en plus des actes de colloques du COSIT, on peut trouver des livres rédigés par plusieurs auteurs, dont ceux édités par David Mark et Andrew Frank (1991), Juval Portugali (1996), Emile van der Zee et Jon Slack (2003) et Michel Denis (1997), un des rares livres en français sur le langage et la cognition spatiale. Il faut citer aussi les deux livres de Golledge « *Spatial Behavior* » (Golledge et Stimson, 1997) dont il est coéditeur et la monographie « *Wayfinding Behavior* » (Golledge, 1999). Et pour terminer deux livres récents de géographie, un de Robert Kitchin (Kitchin et Blades, 2002) coécrit avec un psychologue et la monographie de Donna Peuquet (2002), qui mettent bien en évidence la pluridisciplinarité et la profondeur de la recherche en cognition spatiale.

Les ouvrages fort abondants et très riches en concepts, mais aussi en contradiction nous ont amenés à synthétiser cette information dans les pages qui suivent. La diversité de méthodes et des sciences qui abordent cette problématique amène une complexité certaine pour en faire la synthèse. Certains choix ont été faits qui permettent, nous l'espérons, de saisir la grande variété

des paramètres qui influencent l'apprentissage et le rappel verbal de la localisation d'un lieu en langage courant. Ces choix nous ont amenés à traiter plus particulièrement le lieu à grande échelle, mais non sans effleurer à l'occasion d'autres échelles (de la description des localisations spatiales d'objets sur un écran d'ordinateur jusqu'à l'échelle d'un pays) et la problématique, si riche en littérature, de la description d'un trajet. Des sciences cognitives, nous retenons principalement les théories élaborées en géographie, en linguistique, en psychologie et en neurosciences en délaissant la robotique et l'intelligence artificielle. Pour ce faire, nous abordons cette partie du mémoire en commençant par les modèles et les processus mentaux reliés à l'acquisition des connaissances géographiques.

2.2 *Modèle général*

2.2.1 **Carte mentale et modèle cognitif de l'espace**

Différentes théories et différents modèles ont été proposés dans la littérature selon les disciplines. Les théories sont basées pour la plupart sur les recherches de Piaget (1948) sur le développement de l'apprentissage de l'espace chez les enfants, les travaux de Lynch (1960) sur les cartes mentales, et le chapitre de Siegel et White (1975) sur le développement progressif de la carte cognitive du point de repère puis du réseau et dans un dernier temps d'une vue en survol. De nos jours, les deux grands courants en imagerie, analogique et sémantique, influencés par les recherches en neuropsychologie amènent un ajustement de ces théories, ce que nous verrons plus loin dans le texte.

Les modèles, eux, réfèrent à la connaissance individuelle de l'espace et aux relations avec le milieu, au processus cognitif associé à l'encodage et la récupération de l'information spatiale, lesquels composent la carte cognitive. Les modèles mentaux représentent le monde réel et c'est la théorie cognitive qui les décrit.

En psychologie, la recherche a été principalement orientée vers le dysfonctionnement des cartes mentales; « de 1978 à 1985 plusieurs ouvrages en psychologie ont cherché comment l'environnement était représenté mentalement, en se concentrant principalement sur les distorsions de jugements à propos de l'environnement » (Mark *et al.*, 1999).

Le terme « carte mentale » a souvent été utilisé dans la littérature pour désigner des modèles cognitifs de représentation de l'espace. « L'utilisation du mot « carte » est seulement un terme convenable pour résumer l'information encodée dans les représentations cognitives de l'espace par les individus » (Kitchin et Blades, 2002). Mais d'autres termes aussi ont été proposés dans la littérature : Lynch avec ses images environnementales (environmental image, (1960)), Kuipers avec ses modèles; Tour (Tour model (1978; 1982)) et la hiérarchie sémantique spatiale (Kuipers *et al.*, 2003), les cartes mentales (mental maps (Gould et White, 1986)), de Tversky avec les collages mentaux (cognitive collage (1993)) et les modèles mentaux de l'espace (Spatial mental models (1994))².

La définition des cartes cognitives proposée par Sholl (1996) introduit bien la suite de ce chapitre. Il les définit comme étant le codage graphique des relations euclidiennes (distances linéaires et directions) entre des points de repère importants dans un système de référence de coordonnées centré sur l'environnement. En fait, cette définition nous amène à nous questionner sur chacune des composantes de cette définition. Le codage de l'espace est-il graphique, ou peut-il être codé autrement ? Est-ce vraiment des relations euclidiennes ? Dans l'utilisation du terme relation, fait-il référence à la topologie ? Qu'est-ce qu'un point de repère ? Comment le définir ? Quel système de référence utilise-t-on ?

2.2.2 Modèle d'acquisition des données géographiques

« La connaissance de l'espace est basée sur des expériences sensorimotrices; c'est pourquoi la définition des symboles dans une représentation symbolique est contrainte par l'expérience de perception et d'action » (Harnad, 1990). Ces expériences se déroulent dans un espace chorotaxique, « c'est-à-dire l'étendue qui contient l'homme et ce qui l'entoure, déterminée par les arrangements entre les lieux, avec des attributs spécifiques, mesurables universellement. Cet espace existe nécessairement ; mais il n'est pris en compte que parce qu'il a été *re-connu* en fonction des motivations particulières d'une ou plusieurs personnes. » (Cauvin, 1999).

Le modèle proposé par Cauvin (1999) permet de comprendre la différence entre l'espace réel et l'espace vécu qui mène à la création d'une carte mentale (figure 4). Elle décrit le passage d'un espace chorotaxique (espace réel) aux espaces cognitifs qui sont « des espaces où peuvent

² Kitchin en dresse une liste encore plus exhaustive (Kitchin et Blades, 2002, p. 1).

s'effectuer les mouvements, les déplacements, en fonction d'un but déterminé ». Les espaces fonctionnels, intermédiaires aux deux autres espaces, sont en fait « ces espaces qu'enregistrent et utilisent les habitants d'une ville bien plus que l'espace de référence. Ce sont ces espaces qui vont être filtrés, remaniés par chacun de nous et qui vont donner naissance aux espaces cognitifs, comme le montre le schéma portant sur les différents espaces ».

Ce modèle met en évidence l'espace cognitif qui englobe seulement l'espace d'activité et demeure sujet certains filtres liés à l'expérience du sujet percevant l'espace. Cet apprentissage se fait dans un contexte particulier, pour différentes motivations de déplacement. Il introduit aussi l'importance de l'organisation spatiale des éléments dans l'espace et les attributs subjectifs des éléments constituant cet espace. Ce modèle d'acquisition de l'information géographique reste incomplet, il est orienté vers l'expérimentation du territoire et fait abstraction des espaces acquis par d'autres intermédiaires. Ce modèle fait référence à la carte comme produit, mais pas comme un outil d'apprentissage du territoire à une échelle géographique. Nous verrons plus loin dans ce chapitre l'importance de ce média comme porteur d'information qui permet de comprendre le territoire.

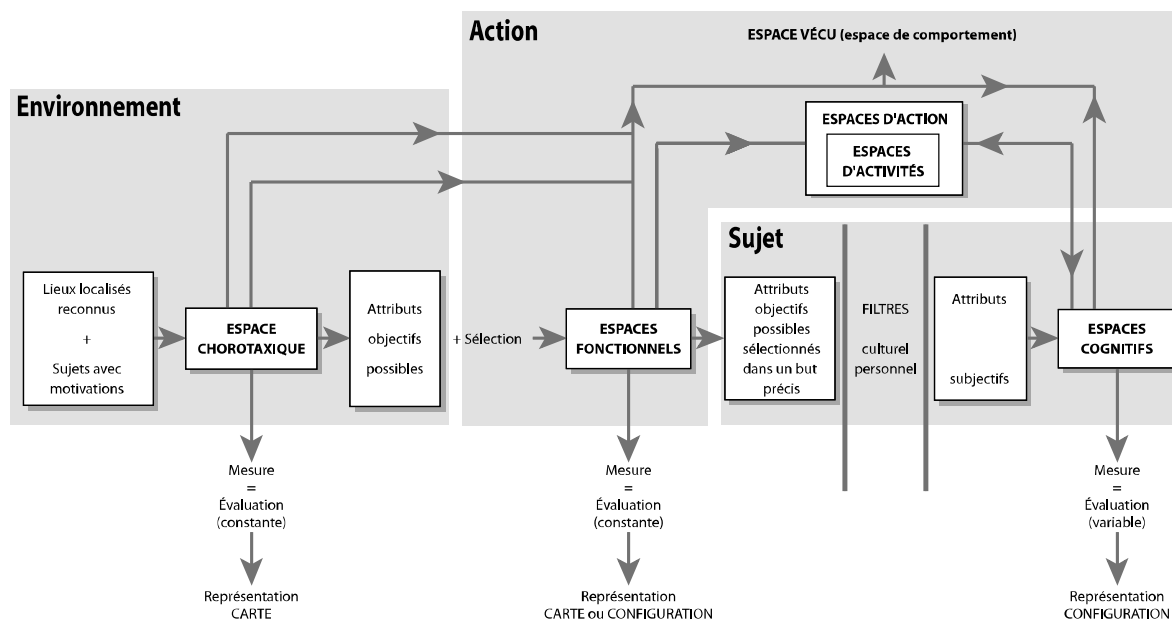


Figure 4. L'espace chorotaxique aux espaces cognitifs d'après Cauvin (1999)

2.2.2.1 Des représentations cognitives liées à divers facteurs

Toujours selon Cauvin (1999), les facteurs qui vont influencer le développement de la cognition spatiale sont les facteurs liés au sujet, à l'action et à l'environnement (figure 5).

Premièrement, les facteurs reliés au sujet sont définis par le statut et les caractéristiques personnelles. Plusieurs études nous amènent vers cette caractérisation de l'espace différent pour chaque individu, mais « il existe un accord substantiel entre les individus qui ont des expériences et des caractéristiques similaires » (Golledge, 1976). La question reste ouverte encore aujourd'hui sur l'existence ou non des caractéristiques de groupe. Dans « Bridging psychology and biology », Kosslyn (2002) suggère d'étudier tant les résultats individuels que les résultats de groupe. Il arrive à la conclusion que dans le cas des images mentales et de leur « géographie cervicale », il semble y avoir une grande similitude des résultats entre les individus.

D'autres études sur les caractéristiques personnelles, portant sur l'influence du genre dans l'acquisition et la capacité de s'orienter dans l'espace, ont à certaines occasions proposé une différence (Cutmore *et al.*, 2000; Fiore et Schooler, 2002) et d'autres des différences non significatives (Davies, 2002). Montello *et al.* (1999), par l'entremise de plusieurs tests, arrivent à la conclusion que les femmes sont de beaucoup supérieures aux hommes pour les tâches de mémoire de localisation d'objets statiques ; par contre, les hommes sont beaucoup plus performants pour l'apprentissage de nouvelle localisation à partir d'expérience directe³. Cette controverse dans les résultats démontre bien la complexité de faire des recherches qui nous permettent de bien comprendre les processus qui entrent dans le traitement de l'information spatiale et surtout de reproduire ces tests dans un environnement naturel différent.

Pour réussir à diminuer les paramètres qui pourraient influencer les résultats, beaucoup de travaux ont porté sur des individus avec des lésions cervicales ou des handicaps. La vision étant considérée comme le mode d'acquisition de l'espace, les résultats sur les individus ayant un handicap visuel sont les plus étonnants : « Les recherches sur la cognition des aveugles suggèrent qu'ils possèdent une représentation spatiale qui est en plusieurs points similaire aux voyants » (Devlin, 2001).

³ Pour une lecture plus exhaustive sur le sujet, il faut consulter Kitchin (2002, p. 103-109) et Devlin (2001 p. 41-95).

Les phases de développement de l'être humain et son influence sur la cognition spatiale sont aussi fort intéressantes, mais moins pertinentes pour ce mémoire (Piaget et Inhelder, 1948; Siegel et White, 1975; Gouteux et Spelke, 2001; Kitchin et Blades, 2002).

Deuxièmement, l'action est la motivation qui nous incite à faire usage de l'espace et qui induit certaines représentations de ces espaces. En fait, la création d'une image, ou carte mentale, résulte d'un problème à résoudre. L'image est construite de manière différente pour chaque problème et il n'est créé en général qu'une seule solution. On ne recherche pas le tracé optimal, mais plutôt celui qui répond au problème à résoudre dans son espace cognitif. Ici Cauvin suppose que sans problème il n'existe pas de carte mentale, hypothèse contestée par d'autres.

Troisièmement, l'environnement est constitué des lieux et de leurs propriétés physiques, de leurs voisinages, de l'organisation des objets et des liens entre ces objets. Les objets ont comme propriétés non pas leur localisation en tant que telle, mais bien leur relation et leur différence avec les autres objets.

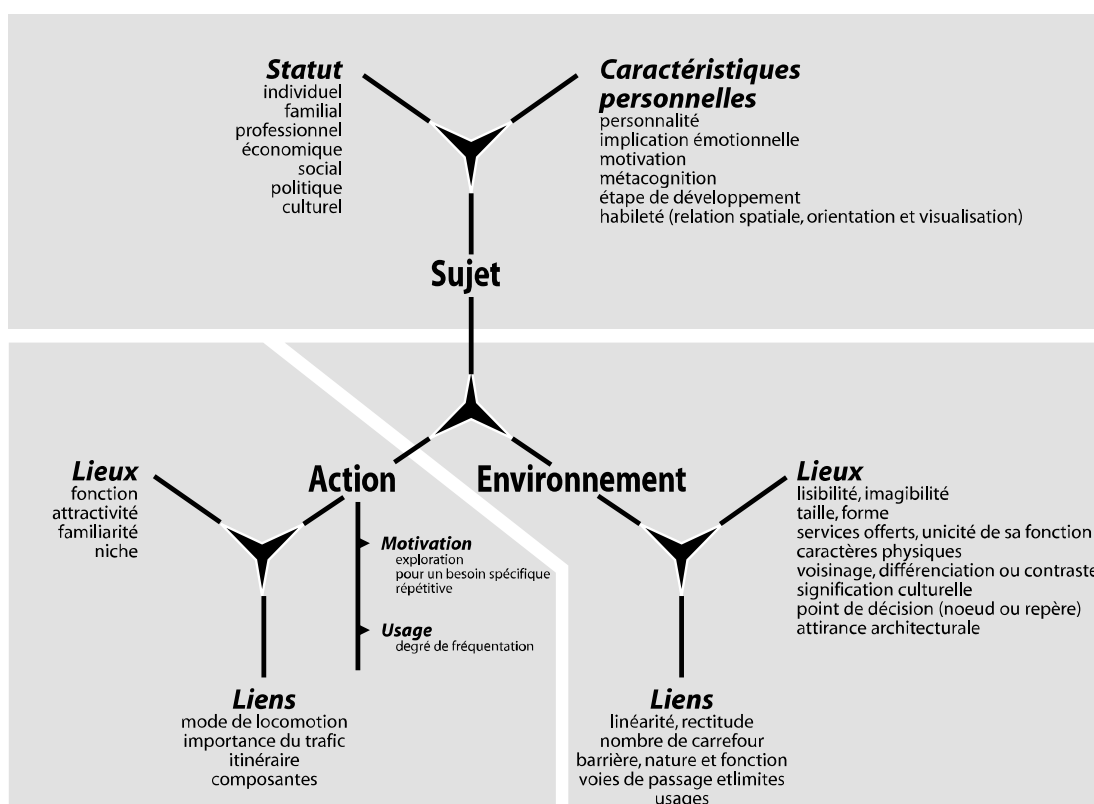


Figure 5. Facteurs influençant la construction de la cognition spatiale d'après Cauvin (1999).

De ce modèle général proposé par Cauvin, voyons maintenant ce que l'on mémorise exactement de l'ensemble des objets. Il n'est pas possible de retenir la totalité de l'environnement de l'espace vécu, ni l'ensemble d'une carte.

2.3 Qu'est ce que l'on mémorise ?

Une théorie de la microgenèse de la cognition spatiale a été proposée dès 1960 par Lynch dans « The image of the city ». L'auteur précise que les cartes cognitives devaient comprendre différents types d'éléments (tableau 3) permettant de construire une image du milieu urbain. Lynch critique dès le départ cette idée de carte; les représentations mentales de l'environnement spatial ne sont pas précises et ne sont pas des modèles miniatures de la réalité, elles sont déformées. Malgré ces distorsions, cette représentation préserve un aspect important de l'espace : les relations topologiques entre les entités du monde. Ces éléments de repère, décrits par Lynch, sont presque toujours cités, mais avec des articulations différentes.

Tableau 3. Les éléments de repères en milieu urbain selon Lynch (1960)

Types	Fonctions	Usage	Dimension	Exemples
Réseau	Réseau pour se déplacer	Naviguer	1	Rue
Frontière	Indication de limite de district	Structurer	1	Barrière, rivière
District	Point de référence	Structurer	2	Voisinage
Point de repère	Point de référence dans lequel on ne peut entrer	Structurer	0	Statue
Nœuds	Point de communication	Naviguer	0	Intersection de rues et édifice public

Ces éléments sont aussi présents dans un premier modèle embryonnaire proposé par Shemyakin (1962) et sont repris en partie par Siegel et White (1975). « Depuis plus de deux décennies, l'acquisition de connaissance de l'espace a été axée sur la connaissance des points de repère, des trajets puis des configurations (en survol) » (Mark *et al.*, 1999). Dans la littérature on traite de l'acquisition de ces trois composantes (tableau 4) comme une suite séquentielle d'acquisitions. D'abord, les points de repère sont obtenus, ils sont des objets qui se démarquent des autres par leurs unicités dans un espace restreint. Suit l'apprentissage des connexions entre ces points de repère, soit la création d'un réseau topologique initial. En dernier lieu, lorsque la connaissance de l'espace atteint un certain niveau, l'individu agence le

tout en une organisation plus générale du territoire en retenant la configuration du réseau et des points de repère pour produire une information plus euclidienne, « une image en survol ». Cette suite séquentielle d'apprentissages de ces trois éléments ne fait pas l'unanimité ; certains auteurs (Montello, 1998) croient plutôt que les trois composantes s'acquièrent au même moment et se raffinent avec le temps.

Tableau 4. Les éléments de la carte mentale selon Siegel et White (1975)

Élément	Description	Perception	Spécificité	Exemple
Point de repère	Objet - Objet proéminent ou une marque distinctive d'un élément dans un paysage	Évènement perceptuel	Propriété relative	Grand Théâtre
Route	Séquence - Lien entre les points de repères	Sensorimoteur Séquence précise Élément de topologie	Distance et direction	Boul. René Lévesque entre l'université Laval et le Parlement
Connaissance globale	Configuration générale de l'environnement	Apprentissage à long terme		Centre-ville de Québec ou de la région de Québec

Chown *et al.* (1995) ajoutent le concept de « gateway » ou portail qui permet de diviser l'espace en espaces plus petits (tableau 5). Le portail marque la transition entre différentes parties d'un espace.

« Dans un environnement interne, les portes sont les portails les plus marquants, dans un environnement naturel les portails incluent les passages en montagnes, les ponts ou les clairières. Les portails sont spécifiquement des places où les gens arrêtent, prennent une pause et regardent autour. Cela peut être une intersection où de nouvelles informations sont disponibles... Ils peuvent être facilement détectables en cherchant des endroits où soudainement la distance perçue de l'arrière-plan change. » (Chown, 2000, p. 2)

Tableau 5. Le concept de portail selon Chown (2000)

Type	Description	Perception	Fonction	Exemple
Portail « Gateway »	Structure visuelle - Marqueur environnemental, Transition naturelle entre deux régions et localisation idéale pour un ancrage visuel.	Visuel essentiellement	Divise l'espace en plus petits espaces	Coin de l'avenue Cartier et le boulevard René Lévesque ouest

2.3.1 Point de repère

On décrit le point de repère, « landmarks » dans la littérature anglaise, comme étant un objet ou une structure qui marque une localisation et qui est utilisé comme point de référence. Pour qu'il soit un point de repère, il doit être proéminent ou constituer une marque distinctive d'un élément dans un paysage. Cette proéminence ne dépend pas des attributs spécifiques à cet objet, mais de la distinction de ces attributs par rapport aux objets voisins. Être un point de repère est donc une propriété relative. De plus, ce point de repère est plus important s'il cumule plus d'un élément distinctif. Certains fournisseurs de données les nomment points d'intérêts (POI ; *points of interest*) géocodés dans les bases de données spatiales sans vraiment les définir.

Lynch (1960) définissait les objets qui étaient des points de repère comme étant des objets singuliers, proéminents ou qui étaient des amers lointains ou des points de repères locaux. Les auteurs, contemporains, traitent des points de repères locaux du moins dans la littérature sur le milieu urbain. Sorrows (1999), en relevant des éléments dans la littérature traitant de l'espace électronique, va plus loin dans la définition des points de repère en considérant leur accessibilité, leur contenu et leur prototypicalité. On entend par prototypicalité, que le point de repère est le meilleur exemple pour une catégorie et typicalité marque l'appréciation du degré d'appartenance à une catégorie. Il s'agit d'un concept majeur en psychologie cognitive sur lequel nous reviendrons.

Raubal et Winter (2002) utilisent les trois grands critères de Sorrows pour définir quel objet peut être un point de repère. C'est en fait dans la description d'une procédure qu'ils nous proposent de qualifier les objets dans une base de données. Les auteurs parlent d'abord des critères d'attraction visuelle (surface de la façade, forme, couleur, visibilité), deuxièmement des critères d'attraction sémantique (importance culturelle et historique, marque sémantique), puis des critères d'attraction structurale (intersection et frontière) pour identifier les objets potentiellement utilisés par les gens pour décrire des trajets ou des lieux à proximité de ces points de repère.

Certains problèmes restent à résoudre dans cette classification d'objet en points de repère dans un SIG notent ces mêmes auteurs. Les points de repères peuvent être classés différemment

selon le mode de transport (piéton, vélo, voiture), selon les usagers (touriste, homme d'affaires, personne handicapée) et selon le moment (nuit-jour, été-hiver). Un bon exemple local est probablement le collège Saint-Charles Garnier sur le boulevard René-Lévesque, un grand édifice qui se distingue des autres par son attraction sémantique et visuelle, sauf qu'en été il est complètement caché par les feuilles et en hiver, difficilement visible en autobus.

Du point de repère, l'on mémorise :

- sa localisation en relation avec le réseau routier et avec les autres points de repère non pas en coordonnées cartésiennes, mais en relation avec les autres points de repère; cette localisation est spécifiée par des paramètres de direction, de distance et voisinage;
- son identification sommaire et générale est plus souvent reliée à sa prototypicalité qu'à son contenu réel;
- sa description, sans comprendre tous les détails, mais certes les détails qui le distinguent des autres, qui font son unicité.

2.3.2 Réseau

« Dans plusieurs modèles de cartes mentales (Piaget et Inhelder, 1948), le réseau est conceptualisé comme une séquence de points de repère. Pour suivre un chemin, il faut suivre une séquence de points de repère. À chaque point de repère sur le réseau, il est nécessaire de sélectionner le prochain point de repère dans la séquence » (Chown *et al.*, 1995).

De ces premiers modèles, l'étude sur les liens entre les points a marqué beaucoup la recherche sur les cartes cognitives, mais aussi les systèmes d'information géographique. « La représentation et les processus des relations spatiales ont reçu beaucoup d'attention dans les applications géographiques parce que les relations dans l'espace géographique entre les entités spatiales sont fréquemment considérées aussi importantes que les entités elles-mêmes. La compréhension et la formalisation des relations spatiales sont fort importantes dans les requêtes spatiales et le traitement de l'information géographique » (Papadias et Kavouras, 1994)

Ce qui est mémorisé des réseaux sont des éléments :

- de métrique ; direction ou orientation et distance
- de topologie, principalement la relation spatiale entre les points de repère.

2.3.2.1 Métrique

Les directions ou orientations font partie des premiers paramètres que l'on mémorise, mais il existe une différence entre les métriques d'une scène et celle d'un espace beaucoup plus grand non perceptible d'un seul coup d'œil.

L'identification des directions des éléments d'une scène requiert plusieurs étapes : l'observateur doit interpréter la scène, distinguer la cible et les éléments de référence en général dans un cadre de référence égocentrique, par rapport à l'observateur. L'apprentissage d'une scène d'un seul angle ne permet pas de construire une vue indépendante de l'orientation de cette scène (Sholl et Nolin, 1997). Pour affranchir cette scène d'une orientation spécifique, il faut qu'il y ait apprentissage dans différents angles, ce qui peut facilement se faire par simple déplacement dans cet espace. Sholl suggère d'ailleurs que si plusieurs études arrivent à la conclusion que l'espace est saisi sans orientation préférentielle, c'est probablement occasionné par un déplacement préalable vers la scène lors des expérimentations, ce qui occasionne déjà un premier apprentissage.

Dans le cadre de l'apprentissage avec une carte ou dans un espace plus grand, donc avec navigation, Werner et Schmidt (1999) démontrent l'importance de l'effet de contexte. L'orientation et les directions sont encodées facilement. C'est le déplacement et la fréquence de déplacement dans cet espace qui permettent de rendre indépendant le cadre de référence lors du rappel d'un lieu ou d'un trajet. La difficulté semble résider plus dans le rappel que dans l'encodage.

La distance métrique entre deux lieux est en général un des paramètres les moins bien emmagasinés. En fait la distance est souvent plus liée à la distance temps ou distance vécue qu'à la distance réelle au sens métrique. De plus, la fréquence de déplacement dans un espace permet de mieux préciser les distances.

« D'un point de vue scientifique (positiviste), la mesure est considérée comme étant la seule façon de « voir » l'espace d'une manière objective. Tout de même, il est possible de la définir à un certain niveau de précision. Les gens normalement expérimentent l'espace non par des mesures, mais plutôt par observation des résultats des processus qui

sont reliés au déplacement dans l'espace, le plus souvent par un déplacement dans l'espace qui requiert du temps. Le temps de déplacement et l'effort sont normalement proportionnels à la distance entre deux points, quoique cette relation soit rarement linéaire. » (Mark et Frank, 1996, p. 11).

2.3.2.2 *Topologie*

La topologie exprime les liens de connexité et d'adjacence entre les entités géographiques. C'est une branche des mathématiques qui étudie les relations d'adjacence, de connectivité et d'inclusion entre les objets.

La topologie est un des éléments le mieux encodés dans la mémoire des individus. Ce que l'on encode c'est la relation entre les points de repère en omettant souvent les distances métriques et en alignant les éléments. Si le réseau est non rectiligne, nous aurons tendance à l'enligner pour en faciliter l'emmagasinage. C'est la méthode la plus économique d'emmagasiner cette information. « Smolenski a qualifié un système de ce type de sémantiquement transparent (Smolensky, 1988) car il ne nécessite pas d'autre traitement que d'emmagasiner ou d'interpréter. » (Chown, 1999).

La recherche sur la topologie dans les SIG a été fort importante depuis plus de trente ans. L'implantation est de plus en plus sophistiquée (Hoel *et al.*, 2003) dans certaines applications et très faible dans d'autres (Theobald, 2001). Quoi qu'il en soit, l'approche cognitive renforce le besoin de poursuivre ce développement, et probablement de poursuivre les recherches, car selon l'usage que l'on fait du territoire, la topologie peut-être différente, un réseau routier peut servir de lien, mais aussi de barrière, une rue à sens unique doit être bien modélisée pour simuler le déplacement en voiture, mais modélisée différemment pour les piétons.

2.3.3 Carte en survol

La carte en survol telle que décrite par Kuipers (2000) en résume l'essentiel : « Une multitude de petites cartes ne sont combinées dans un seul cadre de référence à un niveau métrique que lorsqu'il y a suffisamment d'information et de ressources traitables disponibles ». Cette vue d'ensemble constitue la carte mentale, elle est construite au moment du rappel et particulière au contexte de rappel.

2.3.4 Portail

En milieu urbain, le concept de portail fait référence à des intersections ou des endroits où le paysage change abruptement, ce qui nous permet de saisir un espace d'un seul coup d'œil et nous permettra de construire une « petite carte » de plusieurs éléments du paysage. Ce concept nous permet d'aller plus loin que les points de repère, qui ne font référence qu'à un seul objet et permet probablement de rejoindre le concept d'intersection de Lynch où la visibilité non pas de l'objet, mais du paysage (plusieurs objets) prend toute sa signification.

Dans une étude de Wender (2003) sur la connaissance du réseau, l'auteur note la faiblesse du point de repère et propose plutôt d'utiliser le terme « place » pour tenir compte de l'effet de contexte dans lequel se trouve le point de repère. Même si le concept est différent du portail, il démontre bien que l'approche des points de repère comme le définit Siegel et White est insuffisante.

2.4 Comment le mémorise-t-on ?

2.4.1 Carte mentale

Depuis les premières études de Tolman (1948) portant sur les cartes cognitives, en passant par les travaux de Piaget et Inhelder (1948), sans oublier les contributions de Lynch (1960) et Gould et White (1986), les études portant sur la perception de l'espace ont largement influencé le développement de certains concepts de géographie cognitive. Plus récemment, Montello (1997) définissait la perception de l'espace comme la connaissance intuitive chez les individus des propriétés spatiales incluant la localisation, la dimension, la distance, la direction, la forme, le patron, le mouvement et les relations entre objets.

L'intégration des perceptions spatiales s'effectue à travers des formes de représentation qui sont hétérogènes parce que:

- le langage courant utilisé pour décrire l'espace est particulier et complexe,
- les cartes mentales ne sont pas dotées d'une structure cartésienne,
- la navigation utilise des repères qui ne sont pas nécessairement répertoriés dans les bases de données spatiales,
- et les échelles de perception varient selon les individus et les circonstances.

« Le concept qui tend à s'imposer aujourd'hui est celui de « modèle mental spatial », structure cognitive affranchie de toute perspective, mais utilisable dans le raisonnement spatial pour construire des représentations selon différents points de vue en fonction des nécessités de la tâche » (Denis, 1997).

2.4.2 Théorie de la mémoire et objet dans l'espace

Le cadre théorique auquel nous référons est un modèle souvent utilisé en psychologie, le modèle de Baddeley sur la structure de la mémoire. On peut le résumer ainsi : « Les études en psychologie cognitive et en neuropsychologie ont mis en évidence l'aspect multi-unitaire de la mémoire (Baddeley, 1990). Elle apparaît, en effet, être dissociée en processus discrets, distingués selon un certain nombre d'axes incluant la capacité de stockage (limitée en temps et en quantité d'informations versus illimitée), l'accès à la conscience (dissociation de processus conscients versus automatiques), le type de matériel mémorisé (verbal versus visuospatial), et les systèmes cérébraux sous-jacents. » (Croize, 2001). Ce modèle est représenté graphiquement à la figure 6.

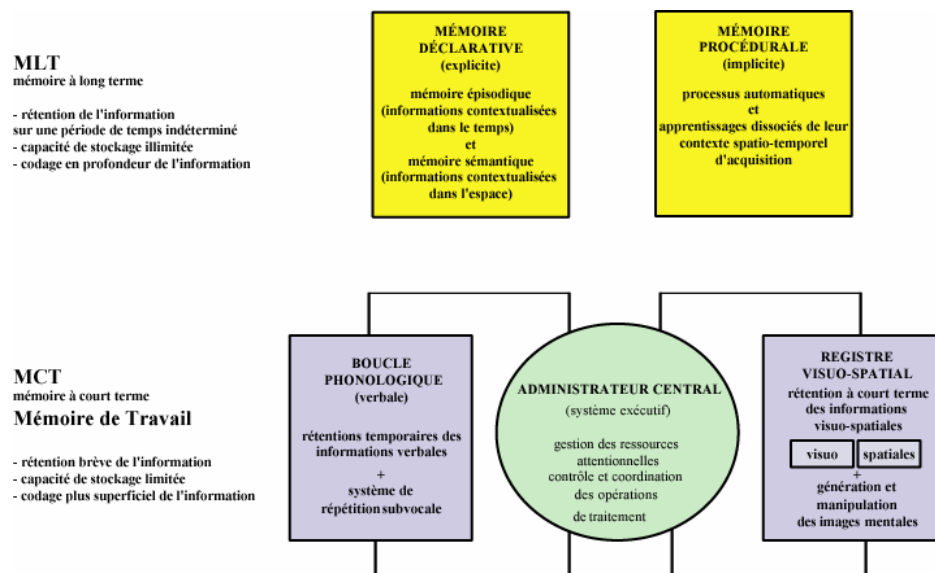


Figure 6. Modèle de la mémoire selon Baddeley (d'après Croize, 2001).

Ce qui ressort de ce modèle, pour les besoins de notre étude, concerne l'existence d'un registre visuo-spatial dans la mémoire de travail et le codage en profondeur dans la mémoire à long terme.

Le registre visuo-spatial est une fonction spécialisée dans le traitement de l'information spatiale et d'une mémoire sémantique qui sert à emmagasiner des informations reliées à leur contexte spatial en utilisant la mémoire à long terme. En fait, Kosslynn (1980; 1994) affirme que l'information est entreposée simultanément sous forme analogique (image) et sous forme verbale (langage). Les travaux de Paivio (1986) sur le codage double (*dual coding*) mettent en relief les avantages de cette forme de mémorisation (tableau 6).

Ces deux modèles complémentaires, celui de Kosslynn et celui de Baddeley, sont très répandus, mais contestés sous deux aspects principalement, soit à propos du principe de division en mémoire de travail et en mémoire à long terme, de même que sur le type de matériel mémorisé. Phyllyshin (2003) prétend encore que seule la forme verbale est emmagasinée. Bien que le débat reste ouvert, les derniers travaux en neuropsychologie ont tendance à favoriser le modèle de Baddeley et à accepter la thèse du modèle multi-unitaire.

Le codage en profondeur est la forme que prend l'information lorsqu'elle est emmagasinée dans la mémoire à long terme. Ses formes de codage sont la catégorisation, la méronymie et la hiérarchisation.

Tableau 6. Types de représentation physique d'après Paivio (Sasse, 1997)

Représentations physiques	Analogique (image)	Verbale (langage)
Exemples	photographies dessins cartes diagrammes	langage humain système formel: maths, logique symbolique logiciel
Propriétés	analogique iconique continue	non -analogique non iconique numérique/discret
Cartes	référentiel isomorphisme	référentiel arbitraire

2.4.2.1 *Catégorisation nominale*

La catégorisation est le processus où des entités distinctes sont traitées comme équivalentes. C'est un des principes fondamentaux et omniprésents des activités cognitives. Il est fondamental parce que la catégorisation nous permet de comprendre et de faire des prédictions à propos des événements et des objets dans notre environnement (Medin et Aguilar, 1999).

Selon Rosch (1978), la catégorisation s'applique à un ensemble d'objets qui sont considérés comme équivalents et généralement désignés par un nom générique (exemple, restaurant asiatique versus chinois, japonais, etc.) ou une catégorie formelle (centre commercial, hôpital, etc.). Le système de catégorisation cognitif permet la rétention et l'utilisation d'une quantité maximale d'information avec un minimum d'effort, c'est-à-dire une économie cognitive maximale. On retrouve deux dimensions au système de catégorisation : une dimension horizontale soit la structure interne des catégories et une dimension verticale soit le niveau d'abstraction. Le tableau 7 illustre la dimension verticale des catégories.

Ce niveau moyen de base correspond au premier niveau où l'on peut concrétiser des objets. L'utilisation d'un niveau de base moyen permet de minimiser l'ambiguïté et de maximiser la compréhension, d'augmenter la similarité et la simplicité pour une interaction entre usagers et de résoudre des conflits durant l'intégration de catégories complexes.

Tableau 7. Catégories et niveau hiérarchique

Niveau	Description cognitive	Niveau d'abstraction	Exemple	Exemple 2
Sur-ordonnées	Très inclusives Peu homogènes	Élément abstrait	Milieu Urbain	Monument religieux
Niveau de base	Information maximum avec le moindre effort cognitif	Dernier niveau où l'on peut décrire un élément de manière concrète	Ville	Église
Subordonnées	Peu d'éléments Très peu inclusives Éléments très homogènes, partageant un grand nombre de propriétés	Description d'élément particulier. Plus l'objet est particulier, moins il y a d'élément dans cette catégorie.	Ville de Québec	Église Saint-Jean Baptiste

Frutos et Berger (2001) arrivent à la conclusion que le niveau de base n'est pas privilégié lors de la généralisation des propriétés, « le statut privilégié des catégories de niveau de base, serait largement déterminé par le contexte de la tâche ».

Un des problèmes rencontrés dans les bases de données est l'utilisation d'un niveau de catégorisation subordonnée trop spécialisé. En fait, la catégorisation par des spécialistes dans un domaine amène souvent une trop grande spécification. Les non spécialistes n'utilisent pas

ce niveau de précision, ce qui nuit à l'opération de la base de données. Il est à noter aussi que les non-spécialistes ne se limitent pas non plus à l'utilisation du niveau de base, mais que le niveau diffère selon leur connaissance (Waxman *et al.*, 1997). L'implantation de différents niveaux de hiérarchie permet de retrouver le niveau employé par les répondants d'enquêtes. Par exemple, pour la description d'un restaurant szechwannais, il faut probablement prévoir plusieurs nomenclatures et hiérarchies différentes pour le retrouver, par exemple un restaurant chinois, oriental, asiatique.

2.4.2.2 *Hiérarchisation spatiale*

À partir de ce principe de catégorisation dite nominale, plusieurs auteurs ont proposé l'existence de catégorisations géographiques (Spencer et Hund, 2002).

Raisonnement spatial hiérarchique

Stevens et Coupe (1978) furent les premiers à trouver une évidence empirique d'un raisonnement spatial hiérarchique et à démontrer comment ce pouvoir heuristique pouvait déformer le jugement et la mémoire des relations spatiales. Cette hiérarchie est formée par la relation spatiale d'un objet qui est identifié à une zone où les propriétés de localisation de la zone sont reprises par l'objet. Cette substitution amène les individus à penser, par exemple, que Montréal est au nord de Seattle, car on attribue la localisation du Canada et des États-Unis aux villes selon leurs pays respectifs.

L'exemple de San Diego et Reno, illustré dans la figure 7, démontre bien le problème qu'occasionne cette catégorisation hiérarchique des entités spatiales. Steven et Coupe (1978) dans leur étude empirique sur les représentations mentales demandaient à des individus d'identifier l'orientation spatiale relative entre une paire de lieux bien connus. Ces participants avaient indiqué en général que San Diego (Californie) était à l'ouest de Reno (Nevada). Barkowski (2001) a noté le même problème pour les villes de Genève (Suisse) et Nice (France).

L'hypothèse retenue par Steven et Coupe ne fait pas l'unanimité. Portugali et Omer (2003) critiquent l'hypothèse que le problème soit relié uniquement à la hiérarchisation, mais

proposent plutôt qu'il y ait combinaison avec l'effet de barrière tel que défini par Lynch et l'effet de rotation tel que défini par Tversky (1992).

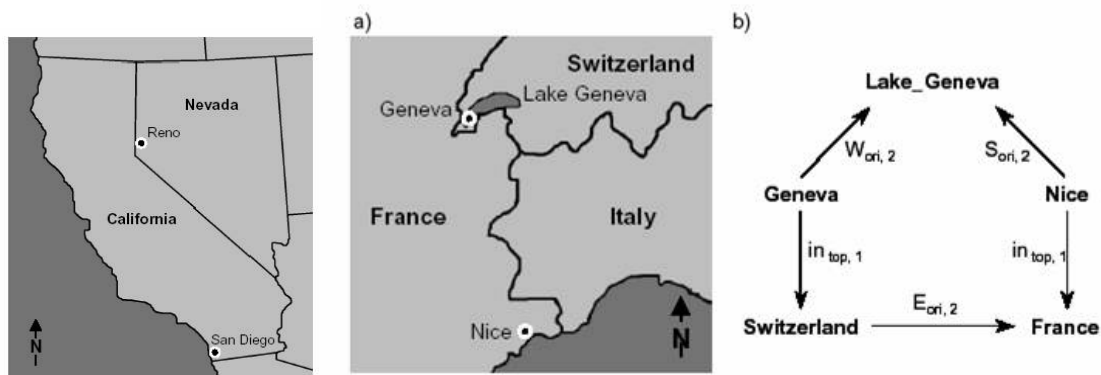


Figure 7. Exemples de problèmes engendrés par la hiérarchisation spatiale (Barkowsky, 2002).

Raisonnement spatial qualitatif

Plus récemment, ces relations spatiales ont été étudiées dans des recherches en intelligence artificielle et en robotique. Le raisonnement spatial qualitatif développe des algorithmes pour décrire et raisonner à propos de l'espace d'une manière robuste et efficace. Ses calculs sont basés sur des concepts mathématiques de topologie et d'orientation.

Plusieurs publications traitent du sujet sous le vocable « qualitative spatial reasoning » (QSR) (Frank, 1992; Gerevini et Renz, 1998; Cohn et Hazarika, 2000; Moratz *et al.*, 2000)⁴. Ces relations peuvent constituer des réseaux d'entités géographiques destinés à faciliter la recherche intelligente des lieux d'activités⁵.

2.4.2.3 Hiérarchie

Les hommes sont capables d'organiser et de structurer l'espace de différentes manières pour simplifier le raisonnement spatial. « Il y a plusieurs tâches où le raisonnement spatial humain performe admirablement et où l'ordinateur est inefficace ou performe peu. La recherche de tracé dans un réseau très vaste est un de ceux-là. Il y a évidence que l'homme utilise dans ce cas

⁴ Les lecteurs qui voudraient avoir un bon aperçu de QSR peuvent consulter l'article de Cohn et Hazarika (2001).

⁵ Les relations de direction ont été revues par Ligozat (1998).

ou dans des cas similaires la structure hiérarchique qu'il impose à l'espace pour réduire l'effort cognitif et améliorer sa performance. » (Car et Frank, 1994).

La cause de cet encodage hiérarchique des relations spatiales peut s'expliquer par la limite du traitement de l'information du système visuel et de la mémoire. Cette limitation nécessite l'organisation d'événements ou d'objets en groupe. Quand le nombre de groupes dépasse la limite, ceux-ci doivent être à nouveau regroupés dans un ordre de plus haut niveau. Ce regroupement séquentiel constitue une hiérarchie. Dans le contexte de la mémoire spatiale, il est facile d'apprécier cette forme d'encodage (figure 8).

Selon Miller (1956), la limite du nombre d'éléments d'un groupe est plus ou moins égal à sept. Le « nombre magique » mesure la capacité de la mémoire à court terme à traiter une information unidimensionnelle. Au-delà de cette limite il y a réarrangement de l'information en groupe, selon une structure hiérarchique ou en ordre séquentiel. Ce nombre revient souvent dans la littérature et semble constituer un seuil « naturel ».

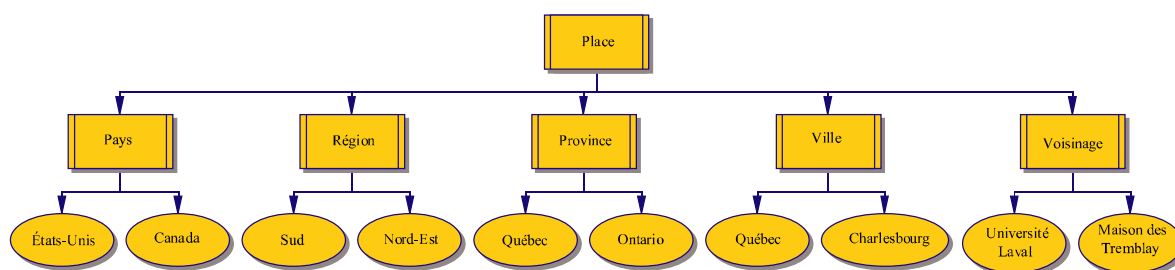


Figure 8. Hiérarchie spatiale.

2.4.3 Représentation externe

Nous avons abordé jusqu'à maintenant les représentations internes de l'information dans la mémoire, mais récemment certains auteurs ont proposé d'intégrer son externalisation et son interreprésentation (Turvey *et al.*, 1990; Varela *et al.*, 1991; Kelso, 1995). Portugali (2002) argumente que le système cognitif en général, et celui associé aux cartes cognitives en particulier, s'étend au-delà de la mémoire des individus, dans l'environnement externe. Le système cognitif est un réseau composé de représentations internes et externes. Les représentations internes étant les entités construites par le cerveau afin de synthétiser

L'information de l'environnement externe, les représentations externes réfèrent à l'environnement ou l'action que représente cette représentation interne.

C'est cette théorie qui permet de comprendre l'importance de la lisibilité d'une ville, cette information inhérente à la forme de la ville que l'on n'emmagasine pas, mais que l'on emploie en relation avec notre représentation interne.

2.4.4 Autres relations

La localisation d'un objet est un des attributs fondamentaux qui peut-être utilisée pour l'identifier. Un autre des attributs fondamentaux est la relation de cet objet avec les autres objets : « normalement la position d'un objet est exprimée en terme de relation avec un autre objet » (Vorweg et Rickheit, 1998). Ces relations peuvent être exprimées de différentes manières comme en témoigne d'ailleurs la diversité des approches théoriques sur cette question. Dans la littérature, selon Grabowski (1999), on aborde les relations selon :

- les propriétés physiques des objets dans le monde (Lang, 1993; Theobald, 2001)
- les propriétés linguistiques des éléments (Talmy, 1983; Levinson *et al.*, 2002)
- les préférences du système cognitif (Levelt, 1982; Landau et Jackendoff, 1993)
- ou le mélange du cognitif et des catégories sémantiques (Herskovits, 1986; Mark et Frank, 1996; Egenhofer et Shariff, 1998)

Les propriétés physiques des objets concernent principalement l'utilisation des SIG pour permettre de définir celles-ci et est essentielle dans la définition même des objets dans les SIG par l'entremise de la topologie. « Un élément essentiel des SIG est la capacité de créer et de manipuler la structure topologique des données en mode vectoriel (Theobald, 2001). »

Les principales relations spatiales sont la distance⁶ et la dimension (Jones, 1997; Cohn et Hazarika, 2001), l'orientation, la direction (Vorweg, 2003), la forme, les relations floues et, bien entendu, la topologie au sens strict, soit l'adjacence et la connectivité.

2.5 Quand on le mémorise ?

Plusieurs auteurs ont proposé que l'information géographique est acquise d'abord avec des caractéristiques qualitatives et que le quantitatif est acquis par une connaissance plus

⁶ Voir la thèse de Berendt sur la distance (Berendt, 1998).

approfondie du territoire, par répétition (Siegel et White, 1975). D'autres par contre, proposent que, à un lieu ou à la première utilisation d'un trajet, une information quantitative est perçue dès le premier passage et que la précision augmente avec la répétition (Montello, 1998).

En fait, l'apprentissage de l'espace se fait de multiples façons. On peut regrouper en deux grandes classes les modes d'acquisition de l'information ; soit l'apprentissage direct, dans la vie quotidienne en parcourant l'espace fonctionnel tel que décrit précédemment et l'apprentissage indirect utilisant des cartes, des systèmes d'information géographique ou des descriptions verbales ou écrites (tableau 8).

Tableau 8. Mode d'acquisition de l'espace

Acquisition	Directe	Indirecte	
Mode	Espace vécu	Carte	Description verbale ⁷ Description écrite
Type d'expérience	Expérience progressive de l'espace	Expérience isochrone de l'espace	Expérience isochrone de l'espace
Point de vue	Variable	Fixe	Variable
Cadre de référence	Égocentrique	Géocentrique	Variable

2.5.1 Apprentissage direct

L'apprentissage direct est le mode d'acquisition le plus fréquent. Quoiqu'en général on traite du mode visuel comme principal mode d'acquisition de l'espace, il ne faut pas oublier les autres modes d'acquisition : auditif, olfactif, haptique (toucher) et proprioception, soit le mouvement du corps dans l'espace. L'importance de ce dernier est bien démontrée par exemple par l'effet d'illusion de déplacement qu'entraîne le déplacement d'un véhicule à proximité lorsque nous sommes arrêtés à un feu de circulation ou la sensation musculaire lorsqu'on circule à vélo sur un faux plat dont la pente est non perceptible pour un automobiliste.

2.5.1.1 Organisation de l'espace

Cet apprentissage direct est modulé par l'agencement des objets dans l'espace et par leurs interrelations. Plusieurs auteurs se sont intéressés à l'importance de l'organisation de l'espace

⁷ Voir tableau 9 pour une explication plus complète des descriptions verbales.

dans la facilité à lire ou à comprendre une ville. Les effets de barrière de certains éléments physiques, tels que les rivières, les forêts ou les montagnes induisent en général les individus vers une surestimation des distances entre les objets qui sont séparés par ces barrières (Canter, 1977; Cohen *et al.*, 1978; Kitchin et Blades, 2002).

Des descripteurs de haut niveau ont été proposés par différents auteurs pour prédire le niveau de précision inhérent à la représentation mentale de l'espace par les gens. Lynch (1960) propose le concept de lisibilité, Weisman (1981) celui de complexité, et plus récemment, Hillier (2001) celui d'intelligibilité. Ce dernier utilise la topologie des axes routiers telle que définie dans la théorie de la syntaxe spatiale (space syntax theory), qui traite des rapports spatiaux entre sous-ensembles locaux ou en tant qu'éléments des systèmes globaux. Dans son article intitulé « A Theory of the City as Object. », il justifie l'utilisation de la corrélation entre les indicateurs de connectivité (topologie et angle) et l'indice d'intégration global comme une simplification analogue à celles des cartes mentales. Pocock (1978) souligne, par exemple, la faible connaissance des paysages dans les villes très uniformes.

2.5.2 Apprentissage indirect

L'apprentissage de l'espace par l'entremise d'une carte permet d'améliorer la représentation mentale en survol d'une région et de réduire le temps d'apprentissage (Thorndyke et Hayes-Roth, 1982). « Pour l'acquisition de l'information de configuration de l'espace, les cartes sont préférables à l'exploration... parce qu'elles font ressortir les éléments de configuration » (Benelli *et al.*, 2002).

Barkowski (2002) fait remarquer dans sa thèse qu'« en observant l'abondance des relations spatiales existant dans une carte géographique il est difficile de la mémoriser en conservant toutes les relations disponibles. De même, il n'est pas possible d'emmagasiner en mémoire la carte entière, comme une photo. »

Selon Thorndyke & Hayes-Roth (1982), les individus avec une expérience directe sur le terrain s'orientent de manière plus précise et sont plus précis pour estimer des distances de parcours que les individus ayant étudié la carte de ce même territoire. Les deux groupes sont aussi précis pour localiser les objets et pour estimer des distances sur une ligne droite. La difficulté pour les

lecteurs de la carte serait en partie due à la vue en perspective du territoire et à l'orientation du Nord au sommet de la carte, ce qui n'est pas représentatif de la réalité.

2.5.3 Fréquence et durée

La fréquence à laquelle un lieu est visité et la durée de la visite jouent un rôle important. Le premier apprentissage d'un lieu peut s'effectuer en général d'un point de vue fixe, dans un seul angle. Le rappel utilise habituellement un cadre de référence égocentrique. « Le cadre de référence dominant est l'expérience égocentrique. Le système de référence sélectionné pour cette première vue est typiquement non mis à jour avec d'autres angles ou mouvement de l'observateur. Toutefois si cette première vue est désalignée, mais qu'une vue subséquente est alignée dans un axe naturel et saillant de l'environnement, le nouveau système de référence est sélectionné et l'assemblage est réinterprété dans ce nouveau système de référence » (McNamara, 2003).

Mennis et Peuquet proposent un modèle d'acquisition et de raffinement de la connaissance des objets géographiques (figure 9). Le problème de ce modèle est qu'il semble perpétuel, pourtant certains auteurs s'entendent pour dire qu'à un certain niveau, lorsque le repérage d'un trajet ou d'un lieu est suffisant, les individus cessent d'apprendre. Étant donné que le paysage urbain, en ce qui nous concerne, continue d'évoluer, il existe un décalage temporel entre ce que l'individu a encodé et l'espace chorotaxique. C'est un aspect fort important qu'il faut retenir lors de la construction et de la mise à jour des bases de données. L'information descriptive des éléments périmés doit être gérée dans le temps et non éliminée des bases de données sous le prétexte qu'elle n'existe plus. Cette information reste encodée dans les cartes mentales des individus et sert lors du rappel. L'exemple du poste d'essence Pétro-Canada située au coin de l'avenue Belvédère et du boulevard René-Lévesque Ouest est révélateur. Le vocable Frigo, nom du dépanneur de l'époque, et l'odonyme Saint-Cyrille, pour désigner l'axe routier, sont encore utilisés par les résidents du quartier, ces références toponymiques ne sont pourtant plus visibles dans le paysage.

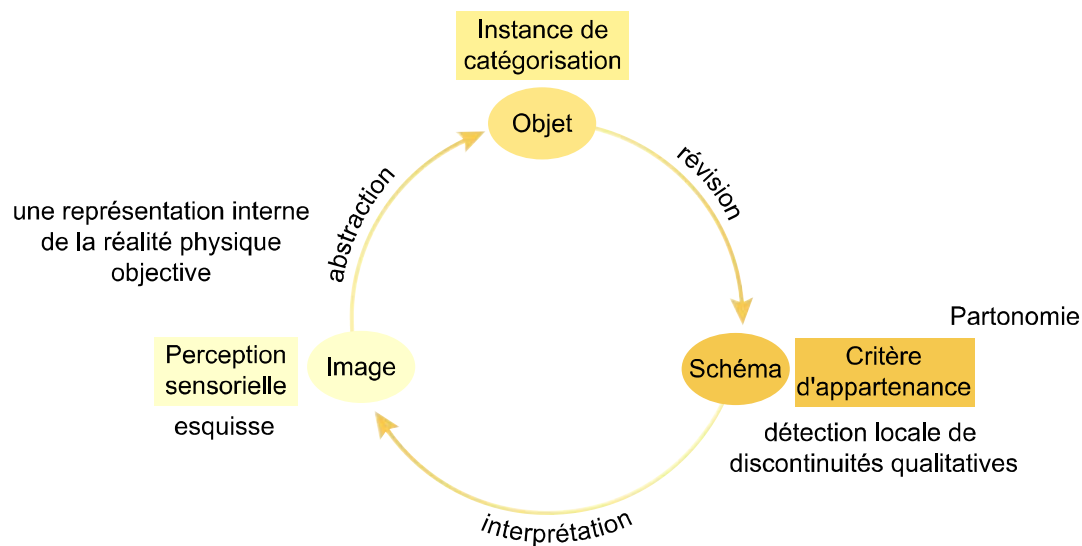


Figure 9. Cycle cognitif de l'acquisition et du raffinement de la connaissance (Mennis et Peuquet, 2000).

2.5.4 La restitution

« La restitution consiste à récupérer à partir des différentes composantes de la mémoire, les éléments qui y sont stockés, en vue de leur réutilisation dans un contexte actuel » (Cadet, 1998).

- En accord avec le modèle d'échelle contextuelle, des asymétries émergent parce que:
 - le stimulus diffère selon le contexte de rappel;
 - l'amplification estimée est mesurée selon le contexte où l'information a été acquise.
- Dans le domaine spatial, la distance entre deux lieux est retrouvée dans la mémoire à long terme et mise à l'échelle dans la mémoire de travail en fonction du contexte courant.
 - Les points de repère permettent une meilleure estimation; sans point de repère, les distances sont souvent sous-estimées. (McNamara et Diwadkar, 1997)

L'information restituée lors de la résolution d'un problème est variable selon le contexte et l'information disponible. Lors de la description d'un lieu, il peut y avoir un décalage entre la réalité et ce qui est rappelé, et même, entre ce qui est emmagasiné et ce qui est restitué. Lors de

la résolution d'un problème « simple », l'information est restituée telle quelle. Lorsqu'on interroge un individu sur un lieu qu'il fréquente souvent, comme sa maison ou son lieu de travail, l'information est restituée sans distorsion. Lorsque la connaissance de la localisation est inexacte, les gens reconstruisent l'information à partir de schémas et des données catégorielles. Si cette information est simplement oubliée, une valeur par défaut est restituée (figure 10). « Quand la mémoire est inexacte et que les gens doivent produire une valeur exacte, ils utilisent un processus d'estimation qui combine une donnée (*remembered stimulus value*) avec une information catégorielle, estimée entre les frontières d'une catégorie et la valeur centrale de cette catégorie (prototype) » (Huttenlocher *et al.*, 1991).

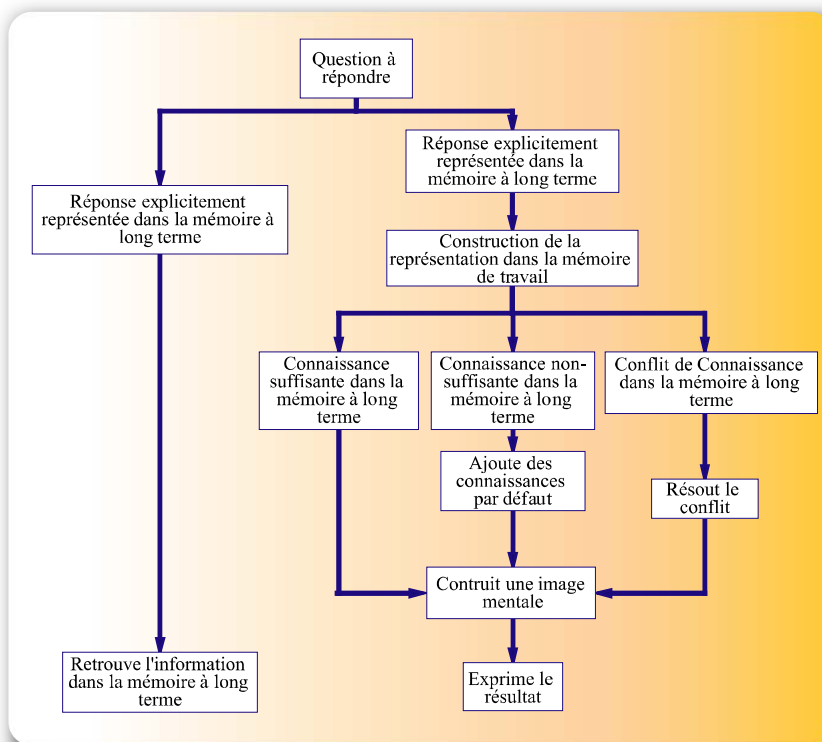


Figure 10. Résolution de problème dans la mémoire et restitution de l'information.

Le sujet des images mentales est abordé dans la thèse de Barkowski (2002). Celles-ci sont construites dans la mémoire de travail, utilisant au besoin des parties d'information récupérées de la mémoire à long terme. Selon lui, la perspective d'un processus de cognition spatiale, comme la procédure de construction d'une image mentale, est très efficace pour décrire

l'information géographique. Il identifie différentes modalités mises en œuvre dans la construction d'une image mentale:

- La représentation mentale d'une configuration spatiale peut être spécifique au problème à résoudre (les entités requises, l'échelle, la résolution et les caractéristiques nécessaires pour la représentation de l'espace).
- Plus ou moins de connaissances spatiales peuvent être emmagasinées efficacement et peuvent être utilisées de manière flexible.
- Les pièces d'information requises n'étant pas toujours disponibles dans la mémoire, la construction sur demande peut pallier des informations manquantes en utilisant des informations plus générales, potentiellement erronées. Ces connaissances par défaut peuvent compléter l'information en lui assignant des détails qui semblent véridiques.
- Les connaissances acquises de sources et de modalités différentes peuvent être combinées dans une représentation unique. Plus spécifiquement, les deux types généraux de connaissances distingués en intelligence artificielle et en psychologie cognitive, le verbal et l'image, sont combinés dans une représentation commune appelée image mentale. D'ailleurs, les avantages respectifs des deux types de représentation peuvent être exploités conjointement.

2.5.5 Le rappel des métriques

Identifier les directions des éléments d'une scène requiert plusieurs étapes : l'observateur doit interpréter la scène, distinguer la cible et les éléments de référence, prendre un point de vue et déterminer l'origine de la description pour décrire la relation spatiale. Chacune de ces composantes est appelée perspective, soit séparément, soit dans son ensemble. La description de cette relation spatiale est purement verbale.

Les directions font partie des choses que l'on apprend dans la phase de vue d'ensemble (survey). On considère qu'il y a une erreur d'environ 20 degrés lors de la restitution de l'information, mais en fait cette dernière est très variable d'un individu à l'autre (Hunt et Waller, 1999). Le problème ne réside pas seulement dans le stockage de l'information de la direction, mais beaucoup dans son rappel. Dans une étude de Haber *et al.* (1993), le rappel en pointant la direction avec le bras est plus précis que l'expression verbale, s'orienter, découper l'espace en segment et décrire selon un cadre de référence qui peut être différent que celui dans lequel on a « appris » le lieu ou le chemin à décrire est beaucoup plus complexe. C'est principalement à ce moment que les erreurs apparaissent et que l'imprécision augmente. L'article de Haber est critiqué du point de vue méthodologique par un article de Montello *et al.*

(1999). Ces derniers dénotent la difficulté de faire la différence entre l'erreur absolue et l'erreur constante.

Premièrement, il y a le problème du cadre de référence. Si l'apprentissage d'une scène s'effectue dans un seul angle, le rappel dans un autre angle est complexe.

Deuxièmement, le problème d'échelle. Quand on parle d'objet proche, on a tendance à utiliser les termes gauche, droite, avant et arrière (Hunt et Waller, 1999) qui sont intrinsèques ou relatifs. À une plus petite échelle, on utilise des descriptions extrinsèques, géocentriques nord-sud, est-ouest.

Troisièmement, lors de l'utilisation des termes gauche-droite, avant-arrière, il y a le problème de différencier la gauche de la droite, problème en partie physiologique, le corps étant symétrique. Par contre, l'asymétrie existe pour devant-derrrière tout comme pour l'axe tête-pied. Ce dernier axe est le plus accessible, car il est renforcé par l'effet de la gravité.

L'autre particularité est que les gens ne divisent pas les 360 degrés les entourant en parties égales. Ils ont tendance à considérer comme devant eux tout objet qui entre dans leurs champs de vision, soit environ 180 degrés.

2.6 Comment l'exprime-t-on ?

2.6.1 Description d'un espace.

Les gens décrivent l'espace en utilisant des « points de repère » *bien définis* dont la localisation est connue (bâtiments identifiés - avec ou sans adresse civique -, raisons sociales, odonymes, intersections de rues) ou *moins bien définis*, soit avec une localisation floue (lieudits, quartiers, secteurs, paroisses), des *vocables qui n'existent plus* (odonymes désuets, bâtiments démolis) ou des *descriptions erronées* (odonymes inexistantes, intersections de rues parallèles). Selon Siegel et White (1975), ces points de repère sont structurés dans une procédure mentale en deux étapes qui consiste d'abord à les relier par une « route » qui trace un lien sémantique, pour ensuite effectuer une coordination de cette perception de l'espace, grâce à la construction d'une carte mentale. Certains auteurs parlent d'une construction en parallèle plutôt que par étapes (Raper, 2000). Ce que l'on peut retenir cependant, c'est que les cartes mentales répondent à une

logique différente des cartes usuelles, elles sont souvent constituées d'éléments discrets, hiérarchisés ou non géométriques. Elles sont aussi empreintes d'émotions et différentes d'un individu à l'autre. Enfin, leur contenu et leur structure varient selon l'échelle, ce qui ne facilite en rien leur représentation dans un système informatique.

Cette information spatiale est verbalisée dans des expressions du langage courant le plus fréquemment non-quantitatives (tableau 9) (près de, à côté de) ou utilisant des quantités floues (environ cinq minutes). La notion de distance est souvent exprimée en temps parcouru plutôt qu'en distance métrique et son évaluation est très souvent erronée. Les temps de déplacement rapportés par les individus sont très difficiles à traduire en distances réelles (Montello, 1997) et leur appréciation varie selon les individus, les technologies utilisées et les circonstances du déplacement.

Il faut bien distinguer ici le système de référence spatiale qui est une collection d'objets spatiaux concrets et le cadre de référence qui est utilisé pour la collection d'entités géométriques qui modèle la structure géométrique du système de référence spatiale.

Tableau 9. Description d'un espace d'après Taylor & Tversky (1996)

	Types de perspectives		
Propriétés	Parcours du regard Scénique	En trajet cinématique	En survol cartographique
Point de vue	Fixe	Variable	Fixe
Verbe	D'état	D'action	D'état
Référent	Objet ou personne	Personne	Objet
Terme de référence	Gauche Droite Avant Arrière	Gauche Droite Avant Arrière	Nord Sud Est Ouest
Cadre de référence	Relatif égocentrique	Intrinsèque allocentrique	Extrinsèque géocentrique
Spécification		Ordre de succession temporelle	
Analogie	Visualisation d'une scène entière à partir d'un point fixe, perspective horizontale	Visualisation en déplacement, perspective cinématique	Visualisation d'une scène entière à partir d'un point fixe, perspective verticale

Le concept de localisation étant intrinsèquement relatif, on ne peut spécifier la localisation d'un objet sans fournir, du moins implicitement un cadre de référence (McNamara, 2003). Par exemple, pour localiser une statue sur une rue, on peut la spécifier sur la rue elle-même (au centre de la rue) par rapport à un autre objet (à côté de la pharmacie Jean Coutu) ou par rapport à un observateur (en face de moi).

À cette dimension spatiale, il faut ajouter la dimension temporelle qui influence notamment le degré de cohérence de la représentation et sa concordance avec le monde réel (Raper, 2000). Plusieurs phénomènes géographiques doivent être situés dans le temps pour préciser leur identité. Comme nous l'avons vu précédemment, les individus font référence à des objets ou des espaces souvent périmés (anciennes limites municipales, paroisses fermées ou dénominations commerciales abandonnées). Cette dimension spatiotemporelle et cette rémanence des repères spatiaux désuets sont des composantes importantes, mais que nous ne traitons pas dans ce mémoire.

Le cadre de référence est déterminant lors de la description d'un espace et il est très difficile à coder dans un système informatique. La lecture des articles de Levinson (2002) et de Vorweg (1998) permet de bien comprendre l'importance du cadre de référence.

Dans le cadre de ce mémoire, ce qui nous intéresse le plus est la description spontanée d'un élément⁸. Bien entendu, il faut bien comprendre ici que l'on fait référence à un élément dont la localisation reste floue et que le locuteur doit spécifier (localisation de l'édifice) en référence à d'autres éléments. Dans ce cas, selon Tversky (1997, p. 45), la sélection d'un référent dans les descriptions brèves est affectée par un ensemble de biais cognitifs, modulés par les caractéristiques spécifiques des tâches. Le biais le plus notable est un biais de communication en faveur de la perspective du participant qui a la charge cognitive la plus élevée, comme moyen de minimiser cette charge ou d'égaliser la difficulté cognitive conjointe. Le second biais est un biais spatial qui conduit les locuteurs à localiser préférentiellement les objets qui sont

⁸ La description de tracé et la recherche de tracé sont très peu abordées dans ce mémoire sinon pour évoquer quelques notions de base. Tout au long du mémoire, nous avons apporté une attention particulière à ne pas s'étendre sur le sujet de la recherche de trace, un sujet très bien couvert dans la littérature (Weisman, 1981; Car et Frank, 1994; Vishon et Cutting, 1995; Allen, 1999; Amarin, 1999; Chown, 1999; Golledge, 1999, 1999; Hunt et Waller, 1999; Kettani et Moulin, 1999; Nadel, 1999; Davies et Pederson, 2001; Raubal, 2001; Hochmair et Frank, 2002; Raubal et Winter, 2002; Timpf et Kuhn, 2003).

« près » d'autres objets ou personnes. La description du référent est relativement directe, mais la relation spatiale est beaucoup plus complexe et plus difficilement modélisable, quoique ici l'on spécifie bien que le locuteur a tendance à utiliser les objets « près »; ces objets sont fort probablement des objets bien distinctifs, donc potentiellement des points de repère connus.

2.6.2 Capacités des systèmes d'information géographique actuels

La représentation et le traitement des relations spatiales ont beaucoup retenu l'attention en géographie appliquée du fait que, fréquemment, les interactions spatiales sont aussi importantes que les phénomènes géographiques eux-mêmes (Papadias et Kavouras, 1994). Comprendre et formaliser les relations spatiales constituent un ajout significatif pour accroître la pertinence des requêtes spatiales et bonifier le traitement de l'information géographique.

Pullar et Egenhofer (1988) distinguent les relations spatiales suivantes :

- relations topologiques : décrivent des relations de voisinage et de connexité;
- relations de direction : décrivent une orientation cardinale (nord, nord-est, etc.);
- relations de distance : décrivent un concept de distance métrique;
- relations ordinales ou comparatives : qui décrivent les inclusions et l'ordre de perception des phénomènes (« dans », « sur », « avant », « après », « plus proche voisin », etc.);
- relations floues de proximité ou de position relative: « près de », « à côté de », etc.

Le contexte actuel des fusions municipales actualise grandement ce problème puisque, dans la région de Québec, de nouvelles villes ont été établies à la fin de 2000 et qu'il faut en conséquence renommer des dizaines de rues dans les nouvelles villes de Québec et Lévis durant les années subséquentes. Ces changements majeurs et officiels prennent un certain temps à entrer dans les habitudes des gens lorsqu'ils parlent de lieux connus. Les bases de données, plus que jamais, doivent tenir compte du temps et ce changement est un bon moment pour implanter des systèmes plus performants afin de gérer la variation temporelle des toponymes. L'approche usuelle se limite souvent à la nomenclature officielle à un moment donné, ce qui limite fortement leur potentiel pour constituer des systèmes de géolocalisation utilisable auprès de la population.

2.7 Notes sur l'échelle

Les théories et modèles proposés proviennent de diverses disciplines qui traitent de l'espace au sens large. Mais il faut noter une certaine mise en garde dans l'utilisation du terme espace. Peut-être faudrait-il le plus souvent parler d'espace géographique tout comme le fait Goodchild (2001).

Freundschuh et Egenhofer (1997) proposent une classification de la perception de l'espace selon l'échelle spatiale et selon la nature des objets qui l'occupent. L'espace peut requérir un déplacement pour être perceptible et il peut être petit ou grand en rapport avec l'individu concerné. Le déplacement réfère à la nécessité de se déplacer à travers un espace pour le percevoir ou l'expérimenter, et subséquemment, apprendre l'espace. Les objets, eux, sont manipulables ou non par l'individu. La manipulation réfère à la capacité de préhension, de tourner ou de déplacer des objets dans l'espace. Dans ce sens, l'espace est formé par la configuration d'objets avec lesquels les gens peuvent éventuellement interagir ou, si ces derniers sont suffisamment grands, entre lesquels on peut se déplacer. C'est à partir de cette catégorisation, qui constitue une assise importante de la géographie naïve, qu'ils distinguent divers types d'espaces (figure 11).

Si l'on prend la définition de ces auteurs, notre étude est à une échelle d'espace environnemental incluant des objets non manipulables (parc, édifices, voisinages et villes), dans de larges espaces qui impliquent un déplacement pour expérimenter, et une acquisition directe qui intègre ces connaissances pendant une période de temps significative (Freundschuh et Egenhofer, 1997; Montello, 1997).

Bien que, dans ce mémoire, nous faisons référence à l'espace environnemental (espace acquis par l'expérience directe), la description floue d'un lieu fait également référence à d'autres échelles géographiques qui incluent cet espace. Ainsi, l'espace géographique est souvent utilisé par les gens afin de situer un lieu qu'ils fréquentent moins souvent. Cet espace géographique correspond à ligne 11 de la figure 11.

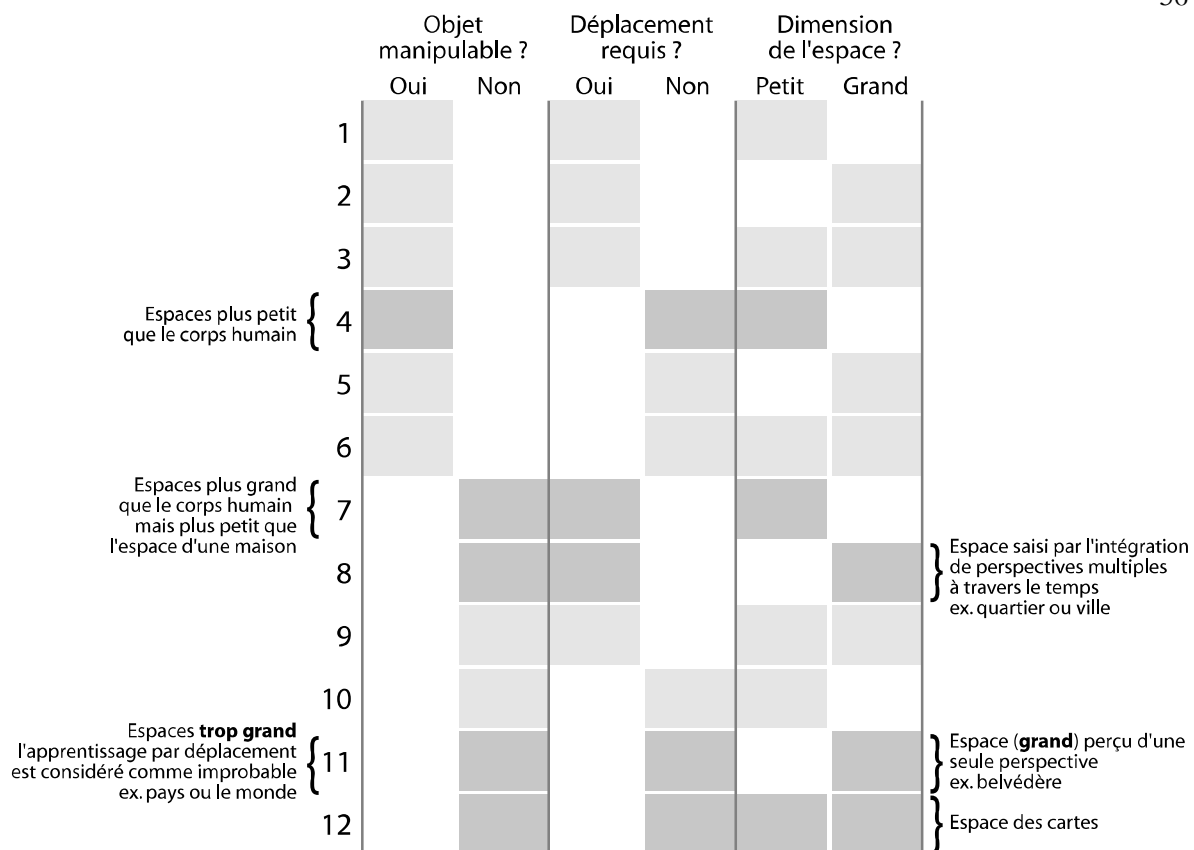


Figure 11. Définition des types d'espaces. Les zones plus foncées indiquent une combinaison de propriétés d'après Freundschuh et Egenhofer (1997)

2.8 Conclusion du chapitre

Dans ce chapitre, nous avons replacé dans un contexte général le cadre de recherche de ce mémoire, puis décrit les différentes étapes de l'acquisition de l'information géographique, la structure de la mémoire et le rappel en langage naturel utilisé pour décrire cet espace. Les différents modes d'acquisition de l'information spatiale permettent de construire des cartes mentales relativement efficaces. Dans le cadre de ce mémoire, la description des lieux est obligatoirement un lieu déjà visité. Une augmentation de la fréquence de visite d'un lieu permet une meilleure capacité de rappel en terme d'exactitude et de point de vue. Les modèles de construction des cartes mentales ont toujours mis l'emphase sur les points de repère, ils servent de point d'ancrage.

Dans la mémoire à long terme, la connaissance est organisée selon un modèle de relation, principalement hiérarchique, catégorielle et prototypale. La reconstruction de l'image mentale

est effectuée lors du rappel et varie selon le contexte à partir de l'information emmagasinée dans la mémoire à long terme. En ajoutant des repères, on permet un rappel plus précis des éléments de localisation.

Maintenant que la base théorique inspirée des sciences cognitives est revue, voyons ce qui en est des bases théoriques des systèmes d'information géographique et de l'intégration des sciences cognitives dans le développement des bases de données.

Deuxième partie

Application à un SIG : aspect théorique et pratique

Chapitre 3

3 Les bases de données spatiales

Les bases de données spatiales sont particulières par la coexistence de deux types de données. « Les données spatiales décrivent l'emplacement, la géométrie et la topologie de l'information à modéliser. D'autres données, de type classique, décrivent les éléments d'information non spatiale et prennent la forme d'enregistrements alphanumériques structurés » (Scholl, 1996). On parle d'objet géographique pour les données spatiales et d'attributs pour les enregistrements alphanumériques.

Les bases de données qui gèrent les attributs peuvent être classifiées selon le modèle de données et selon leur ordre d'apparition. Il en existe 5 types principaux : table simple (1960 environ), hiérarchique (1965), réseau (1969), relationnel (1970) et orienté-objet (fin des années 80). Les tables simples sont certes peu complexes, mais très peu efficaces. Les modèles de données hiérarchiques et de réseau ont été abandonnés, en partie, du à la complexité de maintenir l'intégrité des liens lors de mise à jour et de modification de la base de données. Le modèle relationnel est, encore de nos jours, le plus utilisé parce que la recherche de données est basée sur des opérateurs dans un langage de requête standardisé (SQL) et que la vue conceptuelle du modèle relationnel est constituée seulement de tables et de relations. « Le résultat est un modèle de données flexible et puissant, autant qu'élégamment simple et robuste. Les fondements logicomathématiques du modèle relationnel s'appliquent à plusieurs types de

données avec des opérateurs inclus dans le langage de requête étendu pour s'accommoder aux contextes de nouvelles applications» (Peuquet, 2002). Le modèle orienté objet est encore peu implanté, quoique très prometteur, car il intègre des composantes similaires à celles des modèles cognitifs. Il reste en général utilisé sous les mêmes principes que le modèle relationnel sans en exploiter tout le potentiel. Malgré ces avantages, le modèle orienté-objet (OO) est très prometteur, deux obstacles identifiés par Leung *et al.* (1999) persistent : le manque d'un modèle conceptuel générique et l'intégration insuffisante d'outils de gestion des objets géographiques. La plupart des SIG orientés-objets (SIGOO) sont basées sur les modèles conceptuels traditionnels ou sur l'utilisation du SDTS (spatial data transfer standard) dont le design repose également sur les modèles de données traditionnels. La plupart des SIGOO utilisent des bases de données génériques orienté-objet non spatiales dans lesquels ils gèrent plus ou moins les objets géographiques ce qui constitue essentiellement un système de base de données déguisé, sans plus.

Les modèles de données des objets peuvent se diviser en deux grandes classes, les représentations matricielles et vectorielles. Selon Peuquet (1994), les modèles de données basés sur les représentations vectorielles et matricielles dans les SIG sont inadéquats pour représenter plusieurs phénomènes géographiques complexes, la représentation étant limitée aux formes géométriques primitives de point, de ligne, de polygone et de matrice. Malgré cette critique, les modèles évoluent rapidement et l'on réussit tout de même à réaliser des projets d'envergure.

3.1 Les systèmes d'information géographique

Les SIG sont les outils qui permettent de saisir, emmagasiner, analyser et représenter, en fait gérer, les objets géographiques et leurs relations. On définit aujourd'hui les SIG dans un sens plus large où on inclut aussi les procédures, les individus et les données spatiales qui sont intégrés dans un système.

Bien que des progrès importants aient été réalisés ces dernières années, la modélisation des données spatiales présente encore certaines difficultés. Les premiers modèles de données de SIG mettaient l'accent sur la structure requise pour supporter la géométrie et la topologie des faits géographiques. Malgré l'avancement des sciences de l'information géographique, la

plupart des modèles demeurent très statiques et peu polyvalents; ils ne sont pas adaptés pour refléter l'évolution des phénomènes, le besoin d'interopérabilité des bases de données et l'utilisation du langage naturel dans les requêtes. Il y a un manque flagrant d'ergonomie dans la structure même des bases de données. Plus récemment, les modèles tentent de mettre en relation la thématique et les données géométriques dans une approche plus large qui dépasse les problèmes de structure de données.

3.2 *Intégration des données*

La construction d'un système d'information géographique se fait en grande partie, de nos jours, à partir de données existantes en format numérique. La difficulté est donc non seulement de construire le SIG mais aussi d'intégrer des données géométriques et sémantiques souvent peu compatibles. La fusion des bases de données multisources est un problème : « *tedious batch conversion tasks, import/export obstacles, and distributed resource access barriers imposed by heterogeneous processing environments and heterogeneous data* » (Buehler et McKee, 1998).

La solution proposée depuis le début des années 1990 utilise le concept d'interopérabilité des bases de données (Jensen et Saalfeld, 1998; Worboys, 1998; Fonseca *et al.*, 2000; Worboys et Clementini, 2001). L'interopérabilité « est vue comme une solution aux problèmes d'hétérogénéité syntaxique, structurale et sémantique spécialement l'hétérogénéité spatiale et temporelle entre des données de sources différentes » (Brodeur *et al.*, 2003). Selon le consortium OpenGIS, l'interopérabilité passe par la définition de standards pour la construction des bases de données et leur accès. Il ne s'agit plus de transférer des données, mais d'interrelier plusieurs bases de données conformes à un standard.

Dans ce mémoire, l'approche retenue est l'intégration des données dans une nouvelle base de données parce que les données sources n'étaient pas toujours emmagasinées dans une base de données et que les structures étaient le plus souvent déficientes et très peu documentées ; nous y reviendrons plus tard dans le texte.

3.3 Modélisation des données

La modélisation des bases de données s'effectue habituellement en trois étapes (Shekhar et Chawla, 2003) plus une étape préliminaire : un modèle scientifique de la réalité (Worboys, 1994). Nous en retiendrons ici cinq, pour bien marquer l'importance de l'interface usager lors de la modélisation même des données.

1. Le modèle scientifique de la réalité qui permet de comprendre l'interaction entre les acteurs et le ou les phénomènes étudiés et de saisir les paramètres qui entrent dans le modèle théorique.

Les deux étapes correspondant aux modèles de données « proprement dits » sont:

2. Un modèle conceptuel qui permet de décrire l'information nécessaire à la base de données en terme non-techniques en décrivant les entités (personnes, lieu, etc.), les attributs (caractéristiques et descriptions des entités) et les relations entre les entités.
3. Un modèle logique qui permet l'implantation du modèle conceptuel dans un modèle de type relationnel (entités-relations) (Leung *et al.*, 1999).

La quatrième étape est à l'occasion considérée comme faisant partie intégrante du modèle de données, elle sera traitée ici comme une étape différente, car spécifique au logiciel utilisé.

4. Un modèle physique du modèle logique. C'est à cette étape que l'on se préoccupe du stockage, de l'indexation et de la gestion des données sur support informatique.
5. Le développement de l'interface usager qui fait partie intégrante de cette approche, car cette interface est basée sur le modèle scientifique et non seulement sur l'ergonomie et l'efficacité.

3.3.1 Modèle Entités-Relations

Le modèle relationnel est basé sur une organisation des données sous forme de tables. Comme bien des modèles de données conventionnels, le modèle ER repose sur des entités, des attributs et sur des relations. « Les entités servent à modéliser les objets qui ont une existence indépendante, les attributs décrivent les propriétés des objets et les relations décrivent les liens entre les objets » (Tryfona et Jensen, 1998)

3.4 Modèle cognitif et SIG

Les motivations sous-jacentes d'une approche plus centrée sur l'homme pour les représentations dans un environnement informatique reposent sur trois éléments :

- 1) augmenter la compréhension de l'implication de la cognition humaine dans le contexte d'un

SIG comme moyen d'effectuer une expérience indirecte, pour explorer un monde virtuel 2) augmenter la compréhension de l'utilisation des ordinateurs comme outil afin d'aider les gens à résoudre des problèmes complexes reliés à l'environnement et 3) augmenter nos connaissances sur comment les gens représentent l'espace, apprennent et résolvent des problèmes mal définis afin d'augmenter la capacité de résoudre des problèmes techniques associés à l'utilisation d'ordinateurs dans un environnement contemporain très riche en données (Peuquet, 2002).

Quelques modèles récents intègrent des notions de sciences cognitives. Edwards (2001) propose trois qualificatifs⁹ pour ces modèles selon leur degré d'intégration de la cognition spatiale :

1. Modèle à vraisemblance cognitive (plausibilité) : modèles relativement fidèles à la cognition humaine.
2. Modèle à compatibilité cognitive : modèles qui ne tentent pas de reproduire l'environnement cognitif humain directement, mais plutôt de modéliser cet environnement pour améliorer l'interaction avec l'utilisateur. C'est en fait une stylisation de système de base de données pour anticiper la compréhension humaine.
3. Modèle à portée cognitive : modèles qui reproduisent un aspect très particulier de la cognition.

Par exemple, les modèles développés par Bernard Moulin du département d'informatique de l'Université Laval et son équipe sont de bons exemples de modèle de vraisemblance cognitive (Moulin *et al.*, 1997; Kettani et Moulin, 1999, 1999; Epstein *et al.*, 2001). L'équipe de Peuquet du département de géographie du Penn State University propose des modèles que l'on pourrait qualifier de modèle à compatibilité cognitive et à portée cognitive (Mennis et Peuquet, 2000, 2003).

Mennis propose d'intégrer la cognition dans la stylisation d'une base de données (Mennis, 2003). Il combine la composante donnée et la composante connaissance dans la structure interne d'une base de données (figure 12). La composante connaissance réfère aux principes élaborés dans les sciences cognitives de taxinomie et de partonomie; la taxinomie étant basée sur le principe de catégorisation hiérarchique dérivé de la définition de Rosch et la partonomie¹⁰,

⁹ Ces qualificatifs peuvent être cumulés pour certains modèles.

¹⁰ Le terme méronymie est aussi utilisé pour « définir la relation hiérarchique existante entre deux concepts ou deux signes linguistiques, dans laquelle le premier est une partie d'un tout que constitue le second » (Gouvernement du Québec, 2003).

définissant les parties d'un tout, servant à qualifier la catégorisation selon les propriétés d'appartenance à un tout. D'autres propriétés physiques sont identifiées par l'auteur, comme la taille et la forme.

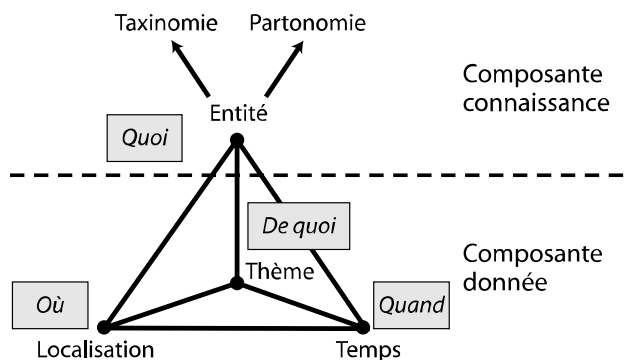


Figure 12. Modèle de Mennis et Peuquet (2000): composante connaissance et composante donnée.¹¹

3.4.1 Relation spatiale

Un aspect important lors de la description de la localisation d'un lieu est la relation spatiale qu'il entretient avec les autres objets localisés. Les relations spatiales peuvent se classer selon Pullar et Egenhofer (1988) en relation topologique qui décrit le voisinage, en relation de direction qui décrit l'ordre (nord, sud) en relation de distance (près, loin), en relation ordinale qui décrit l'inclusion et en relation floue (près de). Ces relations sont exprimées comme nous l'avons vu précédemment en langage naturel, mais sont aussi emmagasinées indirectement dans les objets localisés des SIG.

Papadias et Kavouras (1994) suggèrent qu'un système complet cognitif ou informatique doit comprendre les procédés suivants (figure 13) :

- Construction (perception) : du processus qui crée la relation de représentation basée sur la relation d'un input non interprété, usuellement visuel.
- Transformation (imagination) : processus qui génère les représentations basées sur l'interprétation de l'input, possiblement dans une forme différente.

¹¹ À noter dans la figure 12 que la partonomie semble exprimer une composante de la connaissance dans l'article de 2000. Dans un article plus récent (Mennis, 2003), l'auteur traite la partonomie comme une caractéristique de la taxinomie, ce qu'elle doit être.

- Récupération (raisonnement) : processus qui extrait les relations spatiales explicitement ou implicitement trouvées dans les structures de représentation.

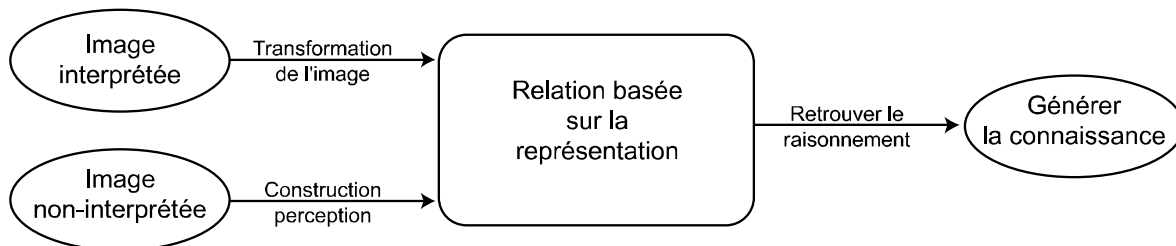


Figure 13. Système de représentation basé sur la représentation spatiale (Papadias et Kavouras, 1994).

3.5 Intégration des sciences cognitives

Du chapitre théorique précédent, nous avons retenu certains éléments qui nous semblent les plus pertinents au développement d'une base de données qui tenait compte de la cognition¹².

Dans la thèse de doctorat de Barkowsky (2002), le problème de la hiérarchisation des objets géographiques est abordé pour construire un modèle informatique à vraisemblance cognitive nommé « Mirage ».

Il part de quatre postulats pour construire le modèle :

- les représentations mentales de l'espace sont construites sur demande ;
- la construction est basée sur des connaissances sous déterminées dans la mémoire à long terme ;
- cette connaissance est emmagasinée dans une forme fragmentaire structurée hiérarchiquement ; et
- le résultat de la représentation dans la mémoire de travail est une image mentale.

Les paramètres que nous avons retenus sont présentés à la figure 14. Nous les avons organisés selon la structure de la mémoire telle que proposée par Kosslyn (1994), tout comme Barkowski (2002). Au concept de hiérarchie de Barkowski, nous proposons d'autres types de relations. Nous ajoutons les relations catégorielles avec prototypicalité, les autres relations et les relations méronymiques telles que définies dans le chapitre précédent.

¹² C'est par l'entremise de la thèse de Barkowski (2002) publiée intégralement dans les Lecture Notes in Artificial Intelligence (volume 2541) que nous abordons cette partie.

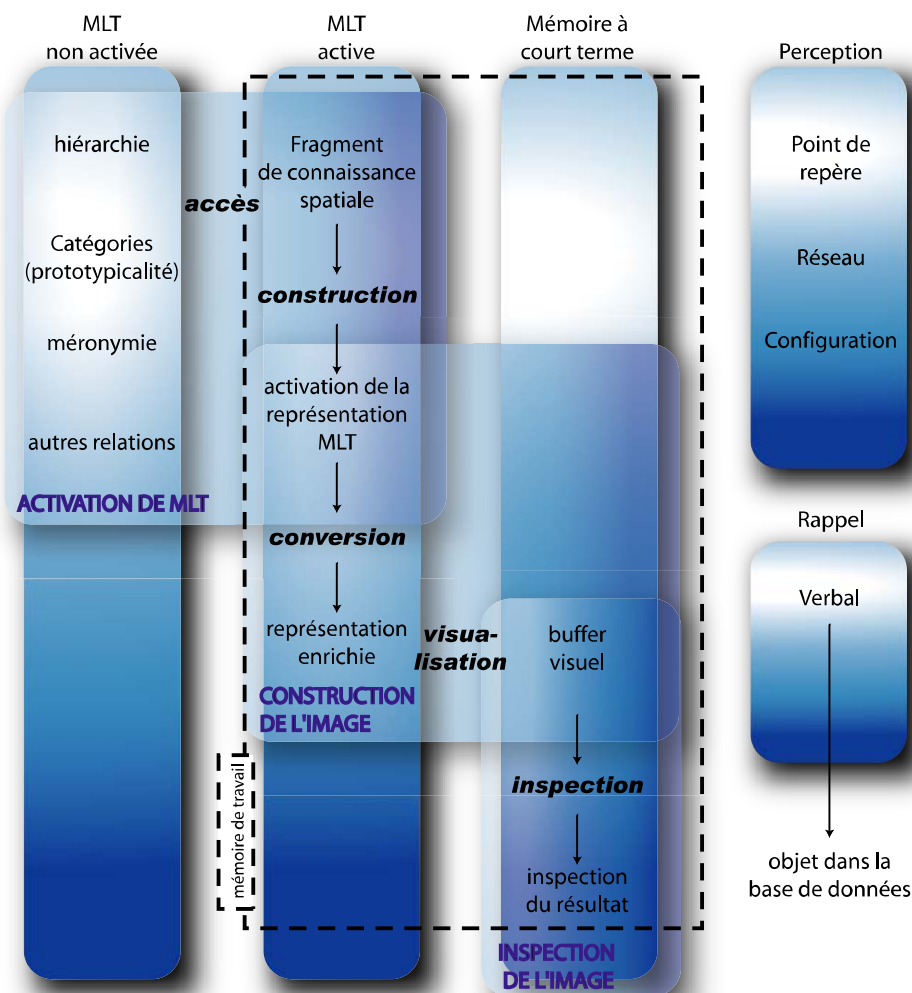


Figure 14. Paramètres intégrés dans le modèle de données tel que représenté dans les sous-systèmes de la mémoire. (Inspiré du modèle de Barkowski (2001)).

L'explication du traitement de l'information est décrite comme un processus sériel et bien défini, ce qui n'est pas le cas dans la réalité. Cette simplification permet de mieux comprendre tous les éléments qui interagissent pour fournir une réponse à une question de localisation. «Le traitement d'un stimulus n'est donc plus considéré comme s'effectuant à travers une succession rigide d'étapes de traitement. Il y a dans le système une constante interaction entre les connaissances en mémoire et les processus de traitement de l'information» (Fortin et Rousseau, 1989).

Dans un contexte de traitement de l'information dans la mémoire, le processus d'acquisition de l'espace géographique peut se diviser en trois grandes étapes. Une première étape d'acquisition de l'information où les éléments de point de repère, de réseau, de configuration, de portail et de relation des objets sont plus importants que les autres éléments et seront emmagasinés dans la mémoire à long terme. Tous les sens sont sollicités, principalement le visuel. Une deuxième étape est celle de « stockage ». Pour optimiser le stockage de l'information, plusieurs méthodes sont utilisées dans la mémoire à long terme soit l'utilisation de hiérarchie, de catégories avec prototype, le stockage des parties d'un objet et leurs relations plutôt qu'un tout et d'autres types de relations. La possibilité d'utiliser les représentations externes permet aussi de minimiser la quantité d'information.

La troisième étape est celle de restitution de l'information. Lorsque le problème à résoudre est simple, l'information circule directement de la mémoire à long terme aux rappels sans passer par la mémoire à court terme. Si le problème est plus complexe, il y a construction de l'information par activation de la représentation, c'est à ce moment que l'on peut parler de « carte mentale », une représentation enrichie qui est construite dans le « buffer » visuel. Le rappel, verbal dans le cas de la description de la localisation d'un lieu dans une enquête téléphonique, sans appui gestuel ou mouvement du faciès, doit à ce moment être retrouvé dans la base de données pour s'assurer de l'exactitude de l'objet décrit, et par extension, de sa localisation.

Dans le cadre de ce mémoire, seront donc retenus les éléments géographiques qui prédominent dans la construction de la carte mentale, principalement les points de repères dans un contexte d'éléments constituant le portail et les méthodes d'emmagasinage.

3.6 Conclusion du chapitre

Le modèle élaboré dans ce mémoire est un modèle à compatibilité cognitive. À partir d'un modèle de données où l'ensemble des informations sur la localisation des lieux est emmagasiné dans une seule table, une modification profonde de la base de données est proposée pour intégrer des paramètres que l'on utilise dans les modèles d'acquisition de l'espace et du rappel de localisation d'un lieu. Le processus automatique de traitement de l'information ne sera pas intégré.

L'effort se situe principalement autour du développement du modèle scientifique et du modèle logique de données. Le modèle scientifique démontre par l'entremise des sciences cognitives, les étapes du traitement de l'information spatiale par l'homme. Dans le chapitre suivant, il est question du modèle de données, de l'intégration des données multisources pour l'enquête OD de 2001 et de la création d'un prototype d'interface usager pour les enquêteurs.

Chapitre 4

4 La construction du SIG

The development of cognitively plausible models of human spatial reasoning may ultimately result in computational systems that are better equipped to meet human needs (Duckham et Worboys, 2001).

La construction de la base de données s'effectue dans un cadre bien précis. Nous rappelons ici seulement le contexte général. La base de données doit servir, lors d'enquêtes téléphoniques, à géoréférencer les lieux visités par les informateurs lors de leurs déplacements durant une journée. Les enquêteurs ne sont pas des spécialistes de base de données, ni de langage d'interrogation, ni de système d'information géographique.

La construction de la base de données s'effectue avec des objectifs connus : nous voulons diminuer au maximum les erreurs lors de la saisie, produire une base de données de localisation potentielle de lieux avec des paramètres qui nous permettent de les retrouver sans utiliser de système d'information géographique. L'information est qualifiée et structurée avec une approche cognitive. L'objectif ultime est de produire une liste des lieux visités par les informateurs et leurs coordonnées géographiques. Étant donné que les gens n'identifient jamais les lieux avec des coordonnées géographiques, il faut donc construire une base de données permettant aux enquêteurs de géoréférencer des lieux décrits en langage naturel, avec possiblement des erreurs ou des omissions.

L'approche retenue pour la construction de la base de données est un modèle de données entités-relations qui nous permet de tenir compte des différents problèmes reliés à la cognition et au rappel. De la cognition, nous retenons principalement les facteurs reliés aux problèmes de catégorisation avec prototypicalité, de hiérarchisation, de méronymie et de changement dans le temps des entités et des nomenclatures sans changement dans la mémorisation des lieux. Pour ce qui est du rappel, nous prenons l'approche d'une interface-usager qui permet de spécifier, en interaction avec le répondant, les différentes alternatives proposées par la base de données si la réponse n'existe pas, ce que nous traitons dans le chapitre suivant.

Comme déjà signalé, nous utilisons une approche de modification de la structure de base de données plutôt qu'une approche algorithmique qui tenterait de simuler la mémoire et les différents paramètres qui influencent la connaissance de l'espace par l'homme. Les relations spatiales sont prétraitées pour permettre de gagner du temps lors de l'interrogation de la base de données et de rendre le système autonome (sans lien avec un système d'information géographique). À cette modification de la base de données, une interface sera ajoutée pour permettre d'afficher une information complémentaire (requête automatique) afin de valider la réponse.

4.1 L'intégration des bases de données¹³

L'intégration des données numériques est un problème qui a évolué dans le temps. D'un problème de non-disponibilité des données numériques des années 60 aux années 80, nous sommes passés à un problème d'intégration de données. Aujourd'hui nous sommes à rechercher des solutions d'interopérabilité des bases de données, où la notion même d'interopérabilité a évolué « en fonction des techniques de stockage, de gestion de l'information et de communication » (Benslimane *et al.*, 1999).

La non-disponibilité des données numériques a fait avorter bien des projets de recherche même au début des années 90, où souvent la totalité des budgets passait dans la collecte de données. L'information, souvent emmagasinée sur un support papier ou tout simplement inexistante, exigeait beaucoup de temps de saisie tant pour les données numériques que pour

¹³ À partir de cette section, les exemples seront en général tirés des bases de données qui ont déjà été utilisées pour la construction de ce projet ; une liste exhaustive est disponible en annexe 1.

les cartes. Dans notre projet les lieudits, les paroisses et les quartiers n'étaient disponibles que sous format papier. Par exemple, pour les paroisses, le diocèse de Québec ne possède pas de carte à jour des paroisses suite aux fusions, la meilleure source selon la personne responsable des archives est chaque curé de paroisse qui est le seul à connaître les frontières réelles de sa paroisse.

Le besoin de standardisation des formats d'échanges est apparu avec la multitude des nouveaux logiciels et de leurs formats de données natives. L'intégration des données a été longtemps très complexe. « Malgré ces obstacles, la réutilisation des données est essentielle parce que le coût d'acquisition à partir de rien est très dispendieux » (Devogele *et al.*, 1998). Plusieurs logiciels de traduction de formats de fichiers pour les données spatiales sont apparus, puis certains formats ont été directement intégrés dans les procédures d'importation et d'exportation des logiciels de SIG.

L'interopérabilité, quant à elle, est un domaine de recherche fort actif depuis une dizaine d'années. « L'objectif principal est d'offrir à un utilisateur les moyens de manipuler de manière transparente des données issues d'un ensemble de sources hétérogènes, autonomes et réparties » (Benslimane *et al.*, 1999). Dans le cadre du projet d'une enquête OD au Québec, il est difficile d'intégrer ce nouveau mode de gestion. Les données ne sont pas gratuites, et elles ne sont pas disponibles entre les ministères d'un même palier de gouvernement. Souvent les données sont distribuées dans un format d'échange très peu efficace, soit en format tableur (MsExcel) ou en format DBF, si ce n'est en format plus particulier. Par exemple, les données de Statistique Canada sont distribuées en format IVT du logiciel Beyond 20/20, format natif non-réutilisable sans exportation.

Le contexte dans lequel les bases de données sont produites doit être compris avant même d'intégrer des données. Il faut connaître : les objectifs, les outils de production et les contraintes de distribution des données. Par exemple, dans les bases de données qui sont utilisées pour ce projet, l'objectif visé par les cartes du MRN du gouvernement du Québec au 1 :20 000 étant la production cartographique. Les cimetières sont représentés par un mot et non une zone, les cours d'eau sont interrompus quand ils passent sous des ponts et la topologie est inexistante. Autre exemple, le ministère de l'Éducation du Québec (MEQ)

produit la liste des établissements scolaires au Québec. (En fait, c'est plutôt la liste des directeurs d'écoles.) Avec les réformes administratives, les directeurs dirigent souvent deux bâtiments. Par conséquent à Sainte-Foy, si on se fie simplement à cette liste, 50% des écoles sont fermées... Sans métadonnées, il est impossible de connaître l'objectif de cette liste.

Les outils de production sont également variables. Les adresses civiques dans les bases de données en sont un bon exemple puisqu'il n'y a aucune standardisation. Par exemple, l'avenue Cartier à Québec est un toponyme officiel, alors que dans le bottin téléphonique, on fait référence à la rue Cartier. Aussi, les producteurs de données ne semblent pas faire référence à des listes officielles d'odonymes. Les cartes numériques de Statistique Canada sont tout aussi particulières : dans la base de données originale en format ArcInfo les aires de diffusion sont composées de points pour les édifices densément peuplés et de zones. Lors du transfert vers MapInfo, ces points sont transformés en petits polygones : outils différents, représentations différentes. Pour ce qui est de la distribution des données, plus d'un problème existe. Le format d'échange utilisé est souvent le format DBF, ou pire encore, le format MsExcel. Trop souvent, on nous présente des tableaux dans Excel et non de réelles bases de données, plusieurs manipulations sont nécessaires et les erreurs sont fréquentes.

Le problème d'intégration des données spatiales comporte certaines singularités par rapport aux données d'attributs (Worboys et Clementini, 2001) :

- représentation incomplète des données : oubli complet lors de la numérisation des données ou omission lors de la numérisation ou conversion imparfaite des données ;
- représentation conflictuelle d'un phénomène spatial ;
- changement de représentation d'un phénomène spatial dynamique. Modification d'une frontière administrative suite à une fusion partielle du territoire.
- observation imprécise d'un phénomène spatial
 - Représentation vague d'objet réel. Ex. frontière floue d'un lieudit
 - Représentation vague d'objet réel résultant d'un changement d'échelle. Ex. arbre vs boisé

4.1.1 Les bases de données d'origines

Les bases de données utilisées dans le projet de l'enquête OD 2001 du RTC sont multiples et très variables en ce qui concerne la production, le format, la sémantique, le type de

géoréférence et la qualité de métadonnées. Il faut se rappeler que le géocodage des origines et des destinations de l'enquête de 1996 se faisait à partir des codes postaux. La localisation pour 2001 est faite à une précision beaucoup plus grande, soit à l'échelle de l'édifice. La complétude des localisations des lieux et leurs références en langage naturel étaient un objectif vers lequel le projet devait tendre.

Les bases de données utilisées pour le projet sont au nombre de 18. On peut les regrouper en trois grands groupes : les répertoires de services, les répertoires de localisation et les répertoires à constituer. Une liste exhaustive est dressée en annexe 1.

Les répertoires de services permettent d'avoir une liste relativement exhaustive des services. Aucune des cinq bases de données n'était géoréférencée, mais elles contenaient toutes des informations pour les géocoder par leur adresse civique ou leur code postal. La base de données de la compagnie Tamec est la plus imposante en quantité d'information (25 769 entrées). Par contre, elle est incomplète et comporte un certain nombre d'erreurs de classification¹⁴.

Les répertoires de localisation sont nombreux, mais très variables pour la précision de la localisation. Le rôle d'évaluation de 1997 de la Communauté Urbaine de Québec est le plus précis, même si étrangement, il comporte quelques erreurs de localisation¹⁵. Lors du géocodage des services, c'était la première des bases de données qui servait à localiser les services dans les édifices selon leur adresse civique. Si la localisation n'était pas trouvée, c'est la base de données de Géocom qui permettait de géocoder des segments de rues dans la région de Québec selon les plages d'adresses. En dernier recours, les services étaient localisés selon leurs codes postaux à partir de la base de données de Statistique Canada (fichier de conversion des codes postaux).

Les producteurs de données ne garantissent en rien la qualité des données. En fait, ce sont plutôt les objectifs des bases de données qui influencent la qualité des données. Le rôle

¹⁴ Un bon exemple de classification erronée est celui de l'Hôpital du Chaudron classé dans les établissements hospitaliers...

¹⁵ L'erreur majeure dans cette base de données est la localisation des maisons de la plage Saint-Laurent à Cap-Rouge qui sont localisées dans le fleuve Saint-Laurent. Pourtant, les numéros de matricules des lots, qui sont en fait le codage des coordonnées géographiques, sont déterminés à partir d'une matrice graphique à grande échelle...

d'évaluation étant fondamental pour les revenus d'une ville, cette base de données est la plus complète et la plus exacte. À l'autre extrémité, Zipcom est un répertoire d'entreprises qui s'enregistrent volontairement.

Les modèles de données des différentes bases de données sont très simples, du moins dans leur format de diffusion. Il s'agit de fichiers plats, sans relation : du fichier Access construit en une seule table au fichier MsExcel qui est plutôt un tableau qu'une base de données, les données doivent être transférées dans une base de données relationnelles pour s'assurer d'un certain niveau de rigueur et d'uniformité en terme sémantique et structurel. Les formats des fichiers spatiaux étant compatibles s'il est relativement simple de transférer toutes ces données spatiales dans un même format de fichier et dans une même projection ; il en est tout autrement sur la structure des éléments. En général, la topologie est inexistante sauf pour le fichier de Géocom conçu pour la modélisation de transport.

4.1.2 Les modalités de l'intégration

La taxonomie proposée par Benslimane *et al.* (1999) permet de qualifier les différentes bases de données utilisées pour le projet. Cette taxonomie (tableau 10) est basée sur les conflits qui peuvent exister dans les systèmes d'information en général et plus particulièrement dans les bases de données à contenu spatial lors de la mise en place d'un système d'interopérabilité. On verra rapidement qu'il est impossible, dans le cas de ce projet, d'utiliser cette manière de gérer l'information compte tenu de la disponibilité de l'information et des formats de données très « pauvres ».

Les conflits structurels entre les objets sont majeurs, par exemple, les routes des différents réseaux routiers ne correspondent pas spatialement, ni géométriquement. Par contre, l'impact pour le géocodage des éléments dans l'espace a été moindre. En effet la procédure établie géocodait les éléments en commençant par la base de données la plus précise (rôle d'évaluation) vers la moins précise (code postal).

Les conflits sémantiques étant fréquents, le principal problème est celui de la dénomination et de l'abréviation. Par exemple, les odonymes comportent peu de problèmes dans leur partie spécifique si ce n'est le problème des « Saint » et « St », mais c'est plutôt pour ce qui est des

génériques (boulevard, chemin, etc.) où les abréviations sont courantes que les problèmes sont les plus fréquents. Les conflits sur l'exactitude du positionnement sont très variables d'une base de données à l'autre. Le cas extrême est probablement la base de données de la Commission de la Toponymie du Québec qui géoréférence les objets en degrés et minutes sans les secondes, soit une précision de plus ou moins 1,5 kilomètres¹⁶. Par exemple, pour localiser une statue en milieu urbain, c'est une grande marge d'erreur.

Tableau 10. Taxonomie des conflits dans les systèmes d'information géographique inspirée de Benslimane *et al* (1999)

	Spatial	Non-spatial
Modèle de données	Représentation matricielle vs vectorielle	Différents modèles de données existent tel le modèle plat, entité-relation et objet
Conflit structurel	Géométrie des objets	Hiérarchisation : codage à des niveaux hiérarchiques différents
<i>Conflit d'échelles ou de généralisation</i>	Représentation différente	Lieux vs quartier
<i>Type de données</i>		Différents formats de date
Conflit sémantique		Interprétation différente en fonction du contexte local d'utilisation
<i>Conflit taxonomique</i>		Dénomination, classification
<i>Conflit temporel</i>	Géométrie modifiée	Changement de nom/date de validation
<i>Conflit de valeurs</i>	Projection, précision et exactitude	Unité et échelle de mesure, abréviation

4.1.3 Le résultat de l'intégration

Avant l'enquête OD 2001, une première base de données a été livrée au MTQ qui était le résultat des fusions des bases de données des services dans la région de Québec. Ces données, de sources multiples, ont été intégrées dans une seule base de données et ont servi de liste de lieux potentiels de destination par les enquêtés. L'objectif pour le ministère était la nomenclature des lieux et leur localisation en coordonnées géographiques. La liste était constituée de 24 375 lieux dont 8 424 ont servi au moins une fois lors de l'enquête (tableau 11).

¹⁶ Aux latitudes de la ville de Québec (47^{ième} parallèle nord), un carré d'une seconde a une dimension d'environ 22,65 mètres d'est en ouest et 31,40 mètres du nord au sud, un carré d'une minute a une dimension de 1359 mètres par 1884 mètres environ.

La nomenclature des lieux a été standardisée pour faciliter la recherche lors de l'enquête, approche qui s'est avérée efficace, mais bien loin de la nomenclature officielle ou de celle du langage naturel. L'analyse de la base de données utilisée par le MTQ lors de l'enquête origine destination de la région de Québec en 2001 nous permet de saisir, du moins en partie, les modifications nécessaires, en terme nominatif, pour repenser une base de données constituée au départ d'une liste des lieux potentiels que les individus des ménages ont pu utiliser soit comme origine, soit comme destination, en une base de données exhaustive et opérationnelle.

Tableau 11. Liste des grands générateurs fréquentés plus de 1000 fois lors de l'enquête OD 2001 dans la région de Québec tels que nommés dans la base de données de l'enquête (9 de 104 657 destinations)

GENER_NOM	NCIV	RUE	SDR_NOM	FREQTOT	XMTM83	YMTM83
PLACE LAURIER	2700	LAURIER BOULEVARD	ST FOY	1609	244942,59	5181436,13
CEGEP ST FOY	2410	ST FOY CHEMIN	ST FOY	1602	244697,55	5183251,02
GALERIES CAPITALE	5401	GALERIES BOULEVARD	QUEBEC	1599	244127,99	5188237,68
UNIV LAVAL CITE UNIVERSITAIRE	0		SILLERY	1423	245963,92	5182476,10
CEGEP GARNEAU FRANCOIS XAVIER	1660	ENTENTE BOULEVARD	QUEBEC	1215	246480,61	5183896,75
BASE MILITAIRE VALCARTIER	0		ST GABRIEL DE VALCARTIER	1175	229324,61	5195418,02
UNIV LAVAL PAV ALPHONSE DESJARDINS	0		SILLERY	1125	246023,00	5182442,06
PLACE FLEUR DE LYS	552	HAMEL BOULEVARD	QUEBEC	1114	247665,02	5187045,58
CEGEP LIMOILLOU CAMPUS QUEBEC	1300	8E AVENUE	QUEBEC	1072	249277,31	5188079,05

4.2 Le modèle de données

Cette section présente la base de données et les modifications apportées qui utilisent une partie des notions de cognition spatiale pour faciliter la recherche de lieu lors d'une enquête. En procédant de la sorte, nous compensons l'information manquante en utilisant des stimuli verbaux pour préciser une localisation floue ou sans correspondance dans la base de données. Ces stimuli sont de type hiérarchique, de méronymie, de référence historique et de localisation principalement.

Idéalement, les modifications permettent une procédure d'enquête de mobilité tendant à s'accommoder de toutes les formes de références spatiales usuelles utilisées par les répondants

afin de retrouver les coordonnées cartésiennes utilisables dans un SIG avec le maximum de rapidité et le minimum d'interventions humaines. Les logiciels de SIG actuels n'offrent pas ce type de service, ce qui limite fortement leur utilisation pour réaliser des enquêtes verbales de mobilité des personnes sans recours à des technologies de positionnement global ou de télémétrie.

4.2.1 Tables de base

La proposition de base de données rend la programmation plus complexe, mais la gestion plus rigoureuse et surtout plus flexible. La base de données pour les localisations telles qu'utilisées lors de l'enquête OD 2001 est constituée d'une seule table (annexe 2) qui liste les générateurs potentiels de déplacement. Ce qui est proposé est une base de données relationnelle (figure 15) qui permet de retrouver un lieu selon ses fonctions catégorisées, d'après sa localisation relative à d'autres éléments spatiaux ou selon les différentes variantes nominales attribuées en langage naturel par les individus.

La base de données est centrée sur les fonctions [FONCTIONS]¹⁷, qui constitue la plus petite unité singulière, car dans un même édifice, il y a souvent plus d'une fonction. Elles correspondent en grandes parties aux entreprises pour une localisation donnée, à une seule adresse. L'unicité sera obtenue par la composition des deux variables. On la nomme fonction, car elle dépasse les entreprises et les individus. On veut être capable d'emmagasiner l'information sur les monuments autant que les parcs publics, en fait, tout lieu qui se définit dans l'espace. Dans cette table, les dates sont emmagasinées aussi bien pour que nous soyons capables de reconstruire le « visage » d'une ville à un moment précis que pour évaluer le roulement des fonctions dans un édifice ou à toute autre échelle spatiale, en fait pour suivre l'évolution spatio-temporelle des phénomènes.

¹⁷ Les noms entre crochets et en majuscules correspondent aux noms des tables, en minuscule aux noms des champs dans la figure 15.

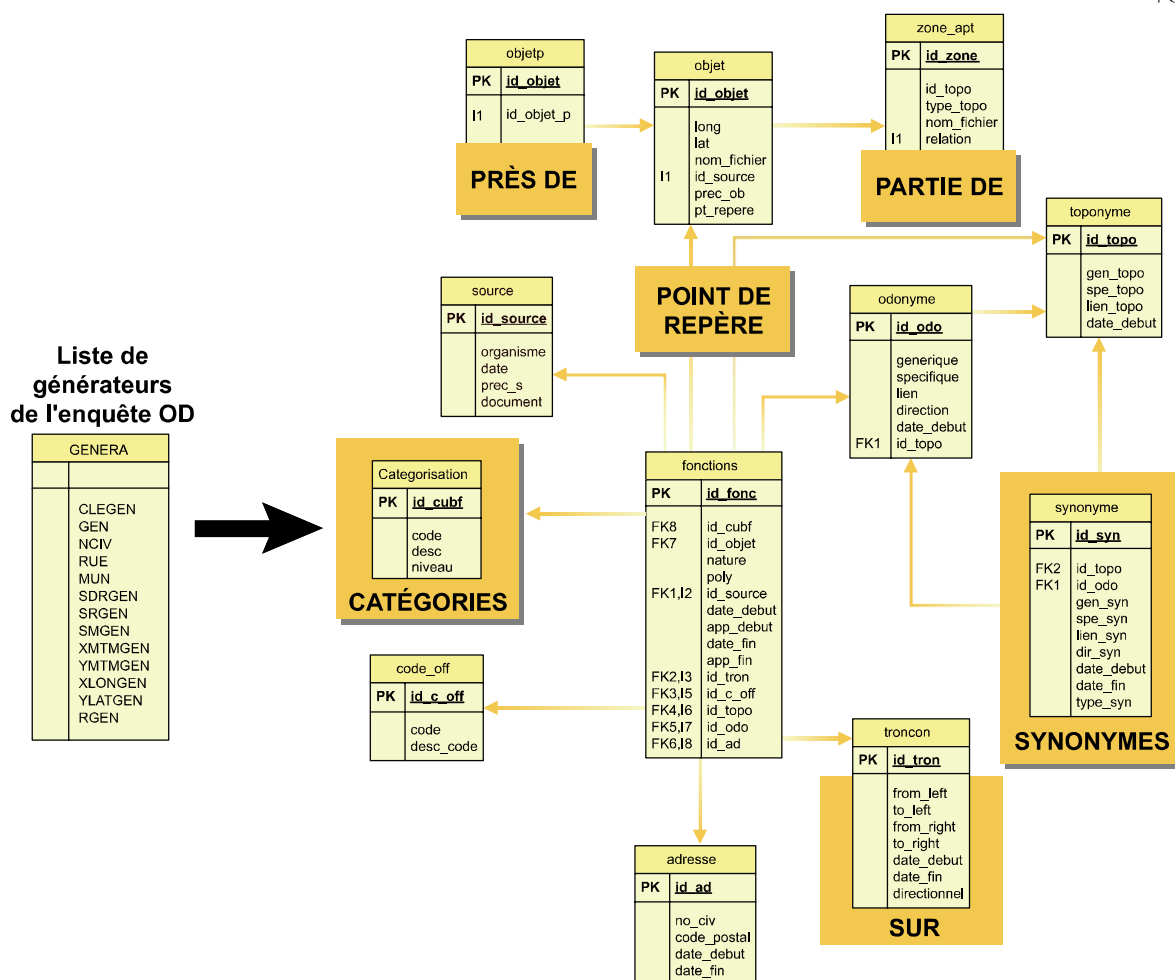


Figure 15. Schéma de base de données adapté pour faciliter la recherche de lieu lors d'une enquête en s'inspirant d'une approche cognitive.

À chaque fonction, existe une localisation obligatoire [OBJET]. Cette localisation est unique, il peut y avoir plusieurs fonctions pour une seule localisation, l'inverse n'est pas possible. Les coordonnées de longitude-latitude assurent l'unicité de la localisation au dixième de seconde. Une précision de plus ou moins trois mètres est suffisante dans une base de données à référence cognitive¹⁸. On se permet une plus grande marge d'erreur. Mais, en milieu urbain, c'est problématique. Une estimation de la précision de la localisation de l'objet est fournie dans la base de données [prec_ob], car il ne faudrait pas rejeter les objets moins bien localisés.

¹⁸ Dans une approche cognitive, la précision de localisation n'est pas la plus importante qualité d'une donnée. La topologie et les diverses relations entre les objets géographiques influencent surtout l'apprentissage et le « stockage » de l'information spatiale comme il a été indiqué dans le second chapitre.

Les autres tables de base sont les codes officiels [CODE_OFF], la source des données [SOURCE]. On entend par code officiel tout code unique qu'un organisme utilise pour différencier les divers établissements de cet organisme. Cela permet de mettre à jour certaines parties de la base de données. Par exemple, le ministère de la Santé utilise un code officiel pour chacun de ses établissements. La source des données permet d'identifier l'organisme qui a diffusé les données, cette table pourrait être beaucoup plus élaborée si on faisait référence à la métadonnée fournie avec les données pour connaître tous les descripteurs des données.

4.2.2 Ajustement cognitif

La structure et les variables des autres tables sont justifiées par l'approche cognitive. On peut regrouper ces ajustements en quatre groupes : les relations spatiales entre les objets, les points de repère, la catégorisation hiérarchique et les synonymes. Les relations spatiales font référence aux vocables de « partie de », « près de » et « sur » qui sont des relations utilisées lors de la description floue de la localisation d'un lieu. De même l'utilisation de points de repère est nécessaire pour le vocable « près de », car en langage naturel, on utilise les points de repère pour spécifier une localisation floue. La catégorisation hiérarchique et les synonymes permettent de retrouver une fonction pour la première ou le nom d'un édifice et l'odonyme pour le second, toujours dans l'optique de retrouver une localisation même si la description est floue ou erronée. Ici on utilise le terme « erroné » pour désigner les vocables non officiels ou les vocables qui ne se retrouvent pas dans la base de données qui sont dans une même catégorie mais à un niveau hiérarchique supérieur.

Les adresses des fonctions sont structurées de manière très particulière pour permettre de gérer efficacement les changements de noms et autres appellations des axes routiers, des zones auxquelles appartient l'édifice (paroisse, quartier, ville) et des édifices. Il est proposé de concentrer les noms officiels dans les tables [TOPONYME] et [ODONYME] et de leur adjoindre une table [SYNONYME] pour conserver les variantes.

4.2.2.1 Point de repère

Les points de repère sont des objets géographiques proéminents dans le paysage. Ce concept permet de limiter le nombre d'objets qu'il faut lister comme voisins potentiels du lieu à décrire.

De plus, les points de repère que les individus utilisent dans le langage courant se situent près des lieux à décrire, en général à l'intérieur du champ visuel. Les points de repère ne sont pas emmagasinés dans une table unique, mais font partie constituante de la table [OBJET] et sont identifiés par le champ [pt_repere] qui indique leur utilisation en référence par la table [OBJETP].

4.2.2.2 *Codage des points de repère*

La recherche des points de repère dans les éléments de la base de données est un élément important dans une approche cognitive de la localisation. Elias (2003) propose une approche d'exploration de données pour sélectionner les bâtiments qui sont proéminents. Il retient plus de 19 critères pour les qualifier, 17 sont des critères géométriques et deux sont qualitatifs soient l'utilisation et le nom du bâtiment. Cette approche, probablement fort opérationnelle, fait fi de l'approche cognitive en excluant l'importance sémantique des objets géographiques. Raubal et Winter (2002) proposent d'emmagasiner les points de repère dans la base de données en les cherchant par analyse spatiale et en cotant ces derniers selon des critères:

- d'attraction visuelle (surface de la façade, forme, couleur, visibilité) ;
- d'attraction sémantique (importance culturelle et historique, marque sémantique) ; et
- d'attraction structurale (intersection et frontière).

Cette approche cognitive exige d'avoir une base de données qui inclut plusieurs paramètres à caractère visuel tels que la surface de la façade ou sa couleur et des paramètres sémantiques qui ne sont pas toujours disponibles dans les bases de données. De plus, quelques questions restent à résoudre. Le point de repère peut être différent selon le mode de transport (piéton, vélo, voiture) et les usagers (touriste, homme d'affaires, personne handicapée). S'il y a plusieurs points de repère très près les uns des autres, comment peut-on les agréger ? Enfin, les points de repère peuvent être différents selon le temps (nuit-jour, été-hiver).

4.2.2.3 *Catégories*

La catégorisation est utilisée dans la mémoire pour optimiser l'encodage et minimiser la quantité d'information. Les codes hiérarchiques sont fréquents dans le codage des entreprises et plusieurs codages existent dont le Système de classification des industries de l'Amérique du

Nord (SCIAN) utilisé par Statistique Canada et le code CUBF utilisé par le ministère des Affaires Municipales du Québec.

La catégorisation devrait se rapproche de ce qui se développe présentement en ontologie. L'utilisation du codage des industries permet de retrouver de manière hiérarchique l'appartenance d'une entreprise à un secteur d'activité, mais dans une nomenclature trop stricte.

« La possibilité d'avoir plusieurs vues d'un même objet géographique est possible en utilisant la hiérarchie et les rôles pour supporter les représentations ontologiques. Donc, un objet géographique peut avoir plus d'une description. Le support de multiples interprétations d'un même espace géographique répond aux questions de différentes applications d'un même territoire. Cette approche permet d'adresser le problème de manipulations d'un objet à différents niveaux de détail par différentes applications. » (Fonseca, 2001, p.9)

L'ontologie est un « ensemble d'informations dans lequel sont définis les concepts utilisés dans un langage donné et qui décrit les relations logiques qu'ils entretiennent entre eux » (Gouvernement du Québec, 2003). L'exemple donné en annexe 4 démontre l'intérêt d'un tel « outil ».

4.2.3 Ajustement géométrique et cognitif

4.2.3.1 Partie de

Dans la base de données, la table [PARTIE DE] permet d'emmagasiner l'appartenance d'un objet à un groupe ou à une zone de niveau hiérarchique supérieur. L'appartenance d'un objet à un groupe ou une zone se détermine de deux façons ; l'appartenance réelle où l'objet est physiquement inclus dans la zone et l'appartenance floue où l'objet est situé à proximité de la zone. Dans le cas présent, cette appartenance floue se limite à une zone tampon de 300 mètres. Il est possible de rendre cette notion de flou beaucoup plus réelle en appliquant les principes développés dans la section 1.2.1.

4.2.3.2 *Près de*

Cette table [PRÈS DE] permet d'inscrire dans la base de données les éléments de repère qui permettent de décrire une localisation floue. Chaque objet peut être associé à plusieurs points de repère.

4.2.3.3 *Sur*

C'est dans la table [TRONCON] que l'on associe la fonction aux tronçons de rue sur lesquels elle se trouve. Dans cette table, on n'identifie que les adresses civiques de début et de fin pour la gauche et la droite de l'axe, la direction et la date d'apparition de début du tronçon et la date de fin s'il y a lieu. L'identification des périodes de validité permet de reconstruire une ville à un moment donné, ne traitant que les axes existant à cette date.

4.2.4 Ajustement nominatif

Les ajustements nominatifs qui ont été effectués pour l'enquête OD de 2001 sont purement une standardisation de la nomenclature. Il n'existait aucune gestion des noms, à savoir si un édifice ou un commerce avait deux noms c'est-à-dire que si dans la base de données, deux noms différents étaient associés avec deux coordonnées géographiques identiques. Il était donc impossible de savoir s'il s'agissait de la même compagnie ou d'une autre compagnie dans le même édifice, l'objectif étant seulement de connaître la coordonnée des lieux de départ et d'arrivée pour l'enquête OD.

Les modifications proposées dans ce mémoire constituent une solution effective de gestion des variations nominatives pour un même objet géographique. Ces variations sont dues au changement du nom avec le temps, à l'utilisation du nom de l'édifice, du principal locataire ou propriétaire, de l'usage actuel ou de l'ancien usage du terrain ou de l'édifice. Trois tables sont dédiées à la gestion des noms : [ODONYME] permet le stockage du nom officiel des rues, [TOPONYME] gère les noms officiels des zones ou édifices, et [SYNONYMES] tous les autres vocables utilisés dans le langage courant pour identifier un lieu, un édifice.

Il faudra approfondir l'utilisation de l'ontologie pour aller beaucoup plus loin dans l'utilisation des catégories. L'annexe 4 démontre un bon exemple de l'utilisation de WebKb¹⁹, un des quelques sites d'ontologie disponibles actuellement, mais seulement en anglais. De même, l'utilisation de Ontoseek semble prometteuse.

Il faudra également développer une typologie pour la table des [SYNONYMES] pour atteindre une plus grande rigueur dans la saisie de nouveaux vocables et pour permettre la reconstitution du territoire avec des vocables d'une autre époque ou des vocables attribués à divers groupes ethniques, par exemple.

4.3 L'interface d'interrogation

La structure de la base de données est importante lors de la saisie de l'information pour faciliter la gestion et assurer l'intégrité des données. Lors de l'utilisation de la base de données, la structure est invisible à l'utilisateur, c'est l'interface-utilisateur qui facilite la consultation de la base de données. C'est souvent lors de l'interrogation de la base de données que l'on peut juger si la base de données est bien construite et possède l'information utile.

La difficulté lors de l'interrogation de la base de données est de traduire la question du langage naturel vers un langage de requête propre aux bases de données afin de trouver l'information que l'on cherche réellement. Du fait que lors d'une enquête, le répondant peut donner une information exacte ou floue, il faut prévoir la recherche à différents niveaux de la hiérarchie catégorielle ainsi qu'une validation de la réponse en utilisant la paronymie et la proximité géographique avec des points de repère.

Le problème de l'interrogation des bases de données spatiales fait l'objet de beaucoup de recherches dans le domaine des SIG depuis plusieurs années et plusieurs avenues de solution sont explorées. Malgré tout, encore de nos jours, l'usage de bases de données de type entité-relation est encore dominant et le langage privilégié est le SQL ou quelques variantes de ce langage.

¹⁹ Voir le site web de WebKb : <http://meganesia.int.gu.edu.au/~phmartin/WebKB/>

4.3.1 Les langages

La formalisation d'un langage dans les sciences de l'information géographique est un thème de recherche important des 20 dernières années. Le cœur de l'effort a été l'identification d'un répertoire de relations spatiales, et comment celles-ci se croisent. Historiquement cette recherche a pris trois voies distinctes selon Peuquet (2002, p. 213). La première utilise directement les principes d'algèbre et de géométrie, les meilleurs exemples étant les travaux de Tomlin (1983; 1990) sur l'algèbre cartographique. Une seconde voie s'est concentrée sur l'amélioration du langage de requête SQL, suite aux nombreuses critiques de ce langage, avec les travaux de Egenhofer (1991) sur les extensions du SQL, le PSQL (Roussopoulos *et al.*, 1988) sur le SQL3 (Snodgrass *et al.*, 1998) et bien d'autres variantes²⁰. La troisième variante en recherche dans la communauté des SIG tourne autour du langage naturel dans le contexte des expressions spatiales, appuyé fortement par la recherche en cognition, en langage et en informatique.

Une classification des langages est proposée par Aufaure et Trépied (2001) et résumée dans le tableau 12. Les auteurs proposent deux grandes classes de langage, soit avec une approche textuelle ou une approche graphique aussi dite par représentation. L'approche ultime est le langage naturel, mais il reste beaucoup d'ambiguïtés à résoudre et les requêtes verbales sont encore difficiles à interpréter particulièrement en français, quoique des solutions récentes soient proposées dans l'article de Ploux et Ji (2003). Ils utilisent la proximité sémantique pour lier des mots en français et en anglais. Les langages artificiels sont les langages les plus répandus. Comme mentionné précédemment, plusieurs variantes du SQL ont été proposées, mais fréquemment l'utilisateur doit prendre plus de temps à identifier la commande qu'il va utiliser plutôt qu'à la tâche elle-même (Egenhofer et Bruns, 1995). Deux aspects limitent le langage SQL, soit l'interrogation des objets spatiaux et la difficulté de formuler les requêtes. L'approche graphique utilise des composantes visuelles comme les diagrammes, les icônes et les images comme métaphores : « Cette approche a connu une utilisation de plus en plus grande ces dernières années pour plusieurs raisons : premièrement parce que les utilisateurs sont autant des spécialistes que des citoyens, et deuxièmement, les développeurs de logiciels

²⁰ Voir l'article synthèse (Egenhofer *et al.*, 1999).

prennent de plus en plus en considération les besoins de l'utilisateur en terme de convivialité et de facilité d'utilisation » (Aufaure et Trépied, 2001).

Tableau 12. Langages de requête d'après Aufaure et Trépied (2001)

			Exemple	Note	Problème
Langage de requête	Approche textuelle	Langage naturel		Favorable pour les utilisateurs	Beaucoup d'ambiguïtés à résoudre. Les requêtes peuvent être verbales et difficiles
		Langage artificiel	TSQL, GISQL, Spatial SQL, SQL3	Extension au SQL pour permettre de traiter correctement l'espace et le temps	Manque de convivialité pour les utilisateurs
	Approche par représentation ou non-textuelle	Tabulaire	QBE	Requête par exemple	Exprimer les jointures
		Graphique	PQBE, SNAP	Meilleur usage du médium graphique	Concepts sous-jacents ne sont pas perçus dans un sens métaphorique
		Langage visuel	Spatial-Query-by-Sketch		
Techniques hypermédias	Carte dynamique	Hypermaps, Argumaps			

Dans un SIG, les modes d'interrogation contraignent fortement les options de structuration des bases de données. Selon Samet (1994), une requête peut prendre une forme spécifique ou exploratoire. Dans le cas d'une requête spécifique, l'utilisateur connaît exactement les critères de recherche et les risques de confusion sont présents, mais maîtrisables. Par contre, dans le cas d'une exploration, il ne sait pas exactement ce qu'il cherche. En général, l'utilisateur a peu d'information sur le contenu de la base de données et il sait encore moins comment l'information y est structurée et représentée. Il utilise une représentation intuitive de la réalité. Souvent, les données spatiales sont retrouvées en combinant des requêtes spatiales et

attributaires. Si la requête comporte plus d'un objet spatial, l'opération définit une jointure spatiale qui utilise une relation topologique ou de proximité entre les entités géographiques représentées dans le SIG. Une requête impliquant plus d'un attribut produit une jointure relationnelle (Samet et Aref, 1994). Plusieurs auteurs proposent de créer des interfaces graphiques (Stacey, 1996; Sebillo *et al.*, 2000) pour faciliter la formulation de ce type d'interrogation et pour présenter le résultat.

Différentes modalités d'interrogation des données sont proposées dans la littérature. La plupart des auteurs conviennent que l'interaction de l'utilisateur avec les données spatiales doit être graphique, la formulation de requêtes utilisant le langage SQL reste limitée en raison de la structure très mathématique qu'il impose et qui entrave une expression naturelle du flot de la pensée de l'opérateur (Egenhofer, 1992). De plus, le langage SQL doit être étendu pour offrir plus que la forme primitive du «select from where». Il reste adéquat pour formuler des requêtes simples qui trouvent une réponse immédiate, mais la structure des requêtes se complexifie rapidement si la représentation de l'espace dans la base de données diffère trop de sa description naturelle ou intuitive.

Dans le cas spécifique de cette application, on peut prévoir une simplification des requêtes en déterminant un nombre limité de requêtes types qui peuvent être activées par des interfaces de dialogue simplifiées lesquelles exécutent les requêtes SQL en arrière-plan.

4.3.2 Le modèle d'interface

Le modèle d'interface proposé permet à l'enquêteur de valider l'existence d'un lieu décrit et, au besoin, de suggérer à l'enquêteur de spécifier des points de repère à proximité du lieu décrit ou l'appartenance du lieu à une catégorie supérieure tant spatiale que fonctionnelle.

L'approche utilise le langage de requête SQL standard pour exécuter les tâches. On propose d'ajouter une gestion des requêtes pour minimiser le temps de réponse. La première réponse à une requête se limite à la recherche d'un seul paramètre tel que défini dans l'interface, soit la recherche du lieu par son adresse ou son nom. Les autres requêtes automatiques qui suivent, servent à la validation de l'unicité de la réponse. Ces autres réponses permettent, au besoin, de stimuler l'enquêteur pour raffiner sa réponse afin de vérifier la validité et l'unicité du lieu nommé.

Cette désynchronisation permet d'augmenter la performance en fournissant une réponse le plus rapidement possible et en poursuivant la recherche d'éléments connexes pour nourrir la validation. C'est à partir de la théorie vue précédemment que l'on peut cerner l'information nécessaire.

L'interface se présente en quatre parties bien distinctes (figure 16). La première partie « Recherche d'un lieu » permet d'entrer l'information précise, telle que décrite par l'enquête. Cette section fait référence au nom officiel tel que l'on retrouve normalement dans une base de données. Elle sert principalement pour décrire les lieux visités le plus fréquemment, soit la maison, le travail ou autres localisations visitées fréquemment et dont la terminologie utilisée pour décrire le lieu est non floue. La raison sociale permet de décrire le lieu selon le nom de l'entreprise (Saint-Hubert, Pizza Royale etc.). La place d'affaires fait référence à la dénomination d'un lieu (Place Laurier, le 200).

Recherche d'un lieu		Spécifier la recherche		nb
Adresse civique	<input type="text"/>	<input type="text"/>	no_civique	<input type="checkbox"/>
Raison sociale	<input type="text"/>	<input type="text"/>	générique	<input type="checkbox"/>
Place d'affaire	<input type="text"/>	<input type="text"/>	spécifique	<input type="checkbox"/>
Description qualitative		Validation ou liste potentiel de lieux		
Usage	<input type="text"/>	fait partie de	<input type="text"/>	
Sur la rue	<input type="text"/>	près de	<input type="text"/>	
Près d'un point de repère	<input type="text"/>	autres noms	<input type="text"/>	
Dans le secteur ou quartier	<input type="text"/>			

Figure 16. Modèle d'interface usager proposé qui intègre des éléments permettant de retrouver un lieu de façon unique et non équivoque.

La seconde partie « Description qualitative » sert à préciser la description en ajoutant des qualificatifs d'appartenance soit à une catégorie hiérarchique supérieure, ou d'appartenance

géographique, « sur », « près de » et « dans ». Cette partie fait référence aux différentes tables de la base de données qui permettent de décrire en langage naturel des lieux dont la localisation est floue dans la mémoire ou dans la connaissance de l'enquêté. Ces paramètres seront utilisés soit pour valider l'information auprès de l'enquêté ou de traiter les spécificités.

La section « Spécifier la recherche » indique le nombre de cas à résoudre pour les adresses civiques, les raisons sociales ou les places d'affaires. C'est par une recherche de cas potentiellement problématiques que l'on peut fournir à l'enquêteur l'information pour s'assurer de l'unicité des lieux (par exemple, route de l'Église ou Saint Hubert que l'on retrouve plusieurs fois dans la base de données).

La dernière section « Validation » suggère une liste potentielle de lieux qui pourraient correspondre à certains éléments entrés lors de la description, par exemple les autres noms associés à ce lieu, toujours dans le même but.

En fait, ce que est proposé dans la construction de l'interface usager consiste à ajouter des informations corolaires pour permettre un dialogue lors de l'entrevue téléphonique. Cette approche fournit des informations contextuelles à la recherche d'un lieu dans une description en langage naturel entre deux individus. Il faudrait probablement regarder du côté de l'intelligence artificielle avec la théorie du « sharedPlan » (Grosz et Kraus, 1993) pour pousser plus loin l'intégration de ce type de requête dans la recherche de concept flou comme « près de ». L'article de Cai *et al.* (2003) permet de croire que l'on pourrait aller beaucoup plus loin.

Conclusion

L'approche par la cognition spatiale est riche en apports pour comprendre les processus d'apprentissage, de traitement et de restitution de l'information géographique. En analysant chacune des étapes, il a été possible de saisir les facteurs qui caractérisent les objets spatiaux facilitant l'apprentissage et la mémorisation des lieux par l'homme.

Nos deux hypothèses de départ se sont avérées justes; par contre, certains ajustements devront être faits. Notre première hypothèse se formulait ainsi « l'ajout d'informations corolaires dans la base de données afin de répliquer la logique des références spatiales relatées par les personnes permet de pallier certaines lacunes inhérentes aux bases de données institutionnelles et aux langages informatiques, parfois inadaptés pour refléter tous les attributs phénoménologiques et relationnels de l'espace géographique. » Cette information corolaire est en partie disponible ou en partie générée dans un SIG. Par contre, la catégorisation hiérarchique avec prototype demandera d'être ajustée au développement en cours sur Webkb, selon les travaux de Jones (2001) et au développement du web sémantique. La difficulté résidera dans l'implantation en français.

Notre seconde hypothèse était « qu'une structure entité-relation est suffisante pour opérer, entretenir et exploiter une telle base de données intégrée dans un SIG vectoriel opérant en mode géorelationnel ». Elle s'est avérée juste et opérationnelle. Nous avons proposé une structure particulière pour nous permettre de gérer les informations corolaires qui rendent cette base de données plus utile. Il faudra poursuivre le développement de la structure de la base de données pour implanter de manière efficiente la hiérarchie catégorielle.

Les principaux objectifs fixés dans l'introduction du mémoire ont été atteints et de nouvelles avenues de recherche ont émergé. La réflexion sur des faits géographiques corolaires par une approche cognitive nous a permis de comprendre les éléments structurant de l'espace qui sont mémorisés. Ce n'est pas l'ensemble des éléments du paysage qui sont emmagasinés, mais bien certains éléments, tels les points de repère, et leur agencement qui servent à diminuer le nombre d'éléments à mémoriser. Cette réflexion a permis de proposer une structure de base de données et une interface usager plus compatibles avec la cognition humaine spatiale.

La structuration de la base de données proposée permet une meilleure gestion des vocables des lieux en gérant les « synonymes » et les relations spatiales entre les objets. Elle permet une intégration des éléments qui permettent de retrouver un lieu même défini de manière imprécise. De plus, l'ajout de la gestion temporelle des noms permet de reconstruire une ville à une époque donnée, avec ses vocables et ses lieux existants ou persistants dans la mémoire.

L'intégration des informations de sources multiples permet une grande diversité de vocables existants pour désigner des lieux à condition, bien entendu, de les gérer pour s'assurer de l'exactitude et de l'unicité géographique des lieux constituant la ville. L'intégration ne doit pas se faire en terme d'uniformisation des vocables, mais en terme d'uniformisation spatiale. L'intégration de sources multiples est perçue, ici, comme un avantage et non comme un inconvénient. Si la phase préparatoire des données pour l'enquête OD 2001 « s'obstinait » à une uniformisation à tout prix, l'approche proposée est celle d'une intégration harmonieuse en gérant les vocables différents pour un même lieu.

L'avantage d'une structure de base de données basée sur une approche cognitive permet d'emmagasiner plus d'information que l'on pourra réutiliser pour d'autres projets et permet également de faire évoluer celle-ci en terme de diversité des noms d'usages pour les différents lieux. En datant les noms de lieux officiels ou non, cela nous permet de reconstruire une ville à un moment donné, de comprendre l'usage d'odonymes anachroniques, de calculer la persistance d'usage de noms de lieu et de comprendre l'usage des points de repère dans le langage courant. En codant les points de repère, cela permet de diminuer les relations entre les objets. Sans point de repère, on devrait lister l'ensemble des relations entre tous les objets, ce qui générerait une base de données gigantesque. Sachant de plus que les gens utilisent seulement les points de repère situés à proximité des lieux à décrire, cela permet encore de réduire les relations entre lieux et points de repère.

La faiblesse du modèle actuel réside dans un usage très limité de la hiérarchisation catégorielle. Il faudra orienter la suite des travaux vers des approches de proximité et de similarité sémantique afin d'obtenir un outil encore plus performant. De plus, en modifiant la base de données pour y inscrire les composantes hiérarchiques et méronymiques, celle-ci devient plus

lourde et plus lente à l'usage. Il faudrait développer des solutions purement informatiques pour optimiser la base de données.

Malgré le fait que la préparation des bases de données pour l'enquête OD 2001 constituait une liste d'objets avec dénomination à standardiser, elle a déjà permis d'augmenter de manière substantielle la précision de la localisation des origines et des destinations et de diminuer énormément le temps de traitement des géoréférences. La base de données proposée dans ce mémoire amène une plus grande rigueur et une utilisation plus variée de cette « liste d'objets géographiques ». L'avantage de l'approche cognitive est de permettre de joindre plus facilement la rigueur exigée par les bases de données et la souplesse exigée par le langage naturel. On pourra gagner du temps sur la préparation des bases de données en gardant des nomenclatures plus diversifiées et en s'assurant de l'unicité et l'exactitude de la géoréférence.

Certains points techniques restent à valider, comme la rapidité d'exécution et la taille de la base de données. Il faudra tester cette nouvelle approche pour mieux évaluer la rapidité d'exécution et le gain réel à utiliser ce système. La première base de données permettait déjà de géoréférencer près de 90% des données. Le développement proposé permettrait un usage plus diversifié de ce type de données, mais à quel prix?

Cette approche de complémentarité cognitive doit être poussée plus loin pour devenir un modèle plus dynamique. L'implantation des processus d'acquisition et de traitement de l'information seraient possiblement plus prometteurs, mais serait à réaliser après des études expérimentales et théoriques plus poussées à propos des étapes du traitement de l'information géographique par l'homme. Déjà, les articles très récents en neuropsychologie sur le rôle de l'hippocampe dans la mémoire (McNamara et Shelton, 2003) permettent de croire que l'on révisera le tout dans quelques années.

L'approche par les catégories hiérarchiques et l'usage de la proximité catégorielle permet une approche computationnelle qui n'est pas obligatoirement quantitative. Le développement d'ontologie spatiale de même type que Wordnet permettra une meilleure interopérabilité des bases de données et une implantation beaucoup plus rapide et plus proche du langage naturel.

Quelques articles récents nous orientent vers des pistes intéressantes (Jones *et al.*, 2001; Cai *et al.*, 2003; Wang, 2003; Rodríguez et Egenhofer, 2004) en terme sémantique.

Avant de continuer dans une approche cognitive, il faudrait tester certains concepts théoriques à partir des enquêtes de déplacements du groupe de Martin Lee-Gosselin. Par exemple, voir s'il existe vraiment des niches, établir à quoi correspondent les points de repère dans la ville de Québec et par une enquête sur le terrain, observer comment naissent des points de repère et leur usage dans la description des lieux en langage naturel.

Une fois la base de données totalement opérationnelle, il faudrait l'utiliser dans un contexte de recherche sur les biographies individuelles qui pourraient permettre d'aller encore plus loin. L'idéal serait dans ce contexte d'ajouter à la base de données des éléments qui nous permettraient de comprendre le rappel d'objets géographiques non visités sur du très long terme (5 à 30 ans).

Il pourrait être intéressant d'impliquer la Commission de la toponymie sur une gestion plus cognitive et plus géographique des noms de lieux afin d'ajouter en précision, en définition géométrique (plus qu'un point) et en sémantique. Il faudra certes sensibiliser celle-ci à développer un répertoire dans une approche d'interopérabilité des bases de données pour permettre d'enrichir leurs bases de données des noms de lieu, de leur usage et de leur histoire...

Il reste difficile d'appréhender cette approche cognitive, car elle touche plusieurs domaines scientifiques et c'est une approche en plein développement. Plusieurs articles émettent des hypothèses contradictoires souvent basées sur des expériences en dehors des laboratoires. De ce mémoire, il restera une ouverture fort enrichissante et un constat de l'importance de l'interdisciplinarité. Et plein d'idées pour se tourner vers le traitement et la gestion de l'information géographique dans une approche cognitive...

Bibliographie

- Allen, G.L. (1999) Spatial abilities, cognitive maps, and wayfinding: Bases for individual differences in spatial cognition and behavior. *In: Wayfinding behavior: Cognitive mapping and other spatial processes* (ed. Golledge, R.G.), pp. 46-80. Baltimore, Johns Hopkins University Press.
- Amorin, M.-A. (1999) A neurocognitive approach to human navigation. *In: Wayfinding behavior: Cognitive mapping and other spatial processes* (ed. Golledge, R.G.), pp. 152-167. Baltimore, Johns Hopkins University Press.
- Aufaure, M.-A. et Trépiéd, C. (2001) What approach for searching spatial information? *Journal of Visual Languages and Computing*, 12(4), 351-373.
- Baddeley, A.D. (1990) *Human memory: Theory and practice*. London, Lawrence Erlbaum Associates. 515 p.
- Barkowsky, T. (2001) Mental processing of geographic knowledge. *In: Proceeding of the COSIT 2001, The fifth international Conference On Spatial Information Theory* (ed. Montello, D.R.), Californie, Springer-Verlag Berlin Heidelberg. Lecture Notes in Computer Science, 2205, pp. 371-386.
- Barkowsky, T. (2002) *Mental representation and processing of geographic knowledge: A computational approach*. Ph. D., Department for Informatics, Hamburg, University of Hamburg. p. 168.
- Benelli, G., Caporali, M., Rizzo, A. et Rubegni, E. (2002) Design concepts for learning spatial relationships. *In: Proceeding of the 19th annual international conference on Computer documentation*, Sante Fe, New Mexico, USA, University of Siena. pp. 22-30.
- Benslimane, D., Jouanot, F., Yétongnon, K., Cullot, N. et Savonnet, M. (1999) Interopérabilité de sig: Un état de l'art. *Revue internationale de géomatique*, 9(3), 279-316.
- Berendt, B. (1998) *Representation and processing of knowledge about distances in environmental spaces. A computational model of inferred route distances investigating their qualitative and quantitative determinants*. Ph. D., Doctoral Program in Cognitive Science, Hamburg, Humboldt-University at zu Berlin. p. 271.
- Bittner, T. et Stell, J.G. (2002) Vagueness and rough location. *GeoInformatica*, 6(2), 99-121.
- Brodeur, J., Bédard, Y., Edwards, G. et Moulin, B. (2003) Revisiting the concept of geospatial data interoperability within the scope of human communication processes. *Transactions in GIS*, 7(2), 243-265.
- Buehler, K. et McKee, L. (1998) *The.opengis guide: Introduction to interoperable geoprocessing and the.opengis specification*. Third edn. Wayland, Massachusetts, Open GIS Consortium Technical Committee. 103 p.
- Cadet, B. (1998) *Psychologie cognitive*. Paris, In Press éditions. 330 p.
- Cai, G., Wang, H. et MacEachren, A.M. (2003) Communicating vague spatial concepts in human-gis interactions: A collaborative dialogue approach. *In: Proceeding of the COSIT 2003, Spatial Information Theory: Foundations of Geographic Information Science* (eds. Kuhn, W., et al.), Kartause Ittingen, Switzerland, Springer-Verlag. Lecture Notes in Computer Science, 2825, pp. 287-300.
- Campari, I. (1996) Uncertain boundaries in urban space. *In: Geographic objects with indeterminate boundaries* (eds. Burrough, P.A. et Frank, A.U.), pp. 57-69. London, Taylor & Francis.
- Canter, D.V. (1977) *The psychology of place*. New York, N.Y, St. Martin's Press. 198 p.

- Car, A. et Frank, A.U. (1994) General principles of hierarchical spatial reasoning - the case of wayfinding. *In: Proceeding of the 6th International Symposium on Spatial Data Handling, SDH'94* (eds. Waugh, T.C. et Healey, R.G.), Edinburgh, Scotland, IGU. 2, pp. 646-664.
- Cauvin, C. (1999) Propositions pour une approche de la cognition spatiale intra-urbaine. *In: Cybergeo. no 72* [En ligne].
<http://www.cybergeo.presse.fr/geocult/texte/cognima.htm> (Page consultée le 21/02/2003).
- Chown, E. (1999) Making predictions in an uncertain world: Environmental structure and cognitive maps. *Adaptive Behavior*, 7(1), 1-17.
- Chown, E. (1999) Error tolerance and generalization in cognitive maps: Performance without precision. *In: Wayfinding behavior: Cognitive mapping and other spatial processes* (ed. Golledge, R.G.), pp. 349-369. Baltimore, Johns Hopkins University Press.
- Chown, E. (2000) Gateways: An approach to parsing spatial domains. *In: Proceeding of the ICML 2000 Workshop on Machine Learning of Spatial Knowledge*, Stanford University, CA.
- Chown, E., Kaplan, S. et Kortenkamp, D. (1995) Prototypes, location, and associative networks (plan): Towards a unified theory of cognitive mapping. *Cognitive Science*, 19(1), 1-51.
- Codd, E.F. (1970) A relational model of data for large shared data banks. *Communications of the Association for Computing Machinery*, 13(6), 377-387.
- Cohen, R., Baldwin, L. et Sherman, R. (1978) Mental representations of spatial relations. *Child Development*, 49, 1216-1218.
- Cohn, A.G. et Hazarika, S.M. (2000) A spatio-temporal theory of physical objects. *In: Proceeding of the ECAI 2000 Workshop on Current Issues on Spatio-Temporal Reasoning*.
- Cohn, A.G. et Hazarika, S.M. (2001) Qualitative spatial representation and reasoning: An overview. *Fundamenta Informaticae*, 43, 2-32.
- Couclelis, H. (1999) Space, time, geography. *In: Geographical information systems; principle and technical issues* (eds. Longley, P.A., et al.), pp. 29-38. New York, John Wiley & Sons.
- Courgeau, D. (2000) New approaches and methodological innovations in the study of partnership and fertility behaviour. *In: Proceeding of the Flagship Conference; Partnership and Fertility - a Revolution? Brussels, Belgium*.
- Croize, A.-C. (2001) Analyse des dynamiques spatio-temporelles des processus de mémoire de travail: Apport de la meg et de l'eeg. *In: Proceeding of the École pratique MEG/EEG de l'Hôpital Pitié-Salpêtrière, Paris, France*.
- Cutmore, T.R.H., Hine, T.J., Maberly, K., Langford, N. et Hawgood, G. (2000) Cognitive and gender factors influencing navigation in a virtual environment. *International Journal of Human and Computer Studies*, 53(2), 223-249.
- Davies, C. (2002) When is a map not a map? Task and language in spatial interpretation with digital map displays. *Applied cognitive psychology*, 16(3), 273-285.
- Davies, C. et Pederson, E. (2001) Grid patterns and cultural expectations in urban wayfinding. *In: Proceeding of the COSIT 2001, The fifth international Conference On Spatial Information Theory* (ed. Montello, D.R.), Californie, Springer-Verlag Berlin Heidelberg. Lecture Notes in Computer Science, 2205, pp. 400-414.
- Denis, M. (1997) *Langage et cognition spatiale*. Paris, Masson. 182 p.
- Devlin, A.S. (2001) *Mind and maze: Spatial cognition and environmental behavior*. Westport, Connecticut, Praeger Publishers. 278 p.
- Devogele, T., Parent, C. et Spaccapietra, S. (1998) On spatial database integration. *International Journal of Geographical Information Science*, 12(4), 335-352.

- Dorion, H. et Poirier, J. (1975) *Lexique des termes utiles à l'étude des noms de lieux*. Québec, Presses de l'Université Laval. 162 p.
- Duckham, M. et Worboys, M.F. (2001) Computational structure in three-valued nearness relations. *In: Proceeding of the COSIT 2001, The fifth international Conference On Spatial Information Theory* (ed. Montello, D.R.), Californie, Springer-Verlag Berlin Heidelberg. Lecture Notes in Computer Science, 2205, pp. 76-91.
- Duckham, M., Mason, K., Stell, J.G. et Worboys, M.F. (2001) A formal approach to imperfection in geographic information. *Computers, Environment and Urban Systems*, 25(1), 89-103.
- Edwards, G. (2001) A virtual test bed in support of cognitively-aware geomatics technologies. *In: Proceeding of the COSIT 2001, The fifth international Conference On Spatial Information Theory* (ed. Montello, D.R.), Californie, Springer-Verlag Berlin Heidelberg. Lecture Notes in Computer Science, 2205, pp. 140-155.
- Egenhofer, M.J. (1991) Extending sql for cartographic display. *Cartography and Geographic Information Systems*, 18, 230-245.
- Egenhofer, M.J. (1992) Research under ncgia research initiative 2: Languages of spatial relations. *In: NCGIA*. <http://www.spatial.maine.edu/~max/i-2-research.html> (Page consultée le 15 septembre 2003).
- Egenhofer, M.J. (1997) Query processing in spatial-query-by-sketch. *Journal of Visual Languages and Computing*, 8(4), 403-424.
- Egenhofer, M.J. et Bruns, T. (1995) Visual map algebra: A direct-manipulation user interface for gis. *In: Proceeding of the Third IFIP 2.6 Working Conference on Visual Database Systems* (eds. Spaccapietra, S. et Jain, R.), Lausanne, Switzerland, Chapman & Hall. pp. 235-253.
- Egenhofer, M.J. et Shariff, A.R.B.M. (1998) Metric details for natural-language spatial relations. *ACM Transactions on Information Systems*, 16(4), 295-321.
- Egenhofer, M.J., Glasgow, J., Gunther, O., Herring, J. et Peuquet, D.J. (1999) Progress in computational methods for representing geographic concepts. *International Journal of Geographical Information Science*, 13(8), 775-796.
- Elias, B. (2003) Extracting landmarks with data mining methods. *In: Proceeding of the COSIT 2003, Spatial Information Theory: Foundations of Geographic Information Science* (eds. Kuhn, W., et al.), Kartause Ittingen, Switzerland, Springer-Verlag. Lecture Notes in Computer Science, 2825, pp. 375-389.
- Encyclopaedia Universalis (2001) Universalis multimédia. *In: Encyclopaedia Universalis. version 7 [DVD]*.
- Epstein, S.L., Moulin, B., Chaker, W., Glasgow, J. et Gancet, J. (2001) Pragmatism and spatial layout design. *In: Proceeding of the COSIT 2001, The fifth international Conference On Spatial Information Theory* (ed. Montello, D.R.), Californie, Springer-Verlag Berlin Heidelberg. Lecture Notes in Computer Science, 2205, pp. 189-205.
- Erwig, M. et Schneider, M. (1997) Vague regions. *In: Proceeding of the Advances in Spatial Databases, 5th International Symposium, SSD '97* (eds. Scholl, M.O. et Voisard, A.), Berlin, Germany, Springer. Lecture Notes in Computer Science, 1262, pp. 298-320.
- Eschenbach, C. (1999) Geometric structures of frames of reference and natural language semantics. *Spatial Cognition and Computation*, 1(4), 329-348.
- Fiore, S., M et Schooler, J.W. (2002) How did you get here from there? Verbal overshadowing of spatial mental models. *Applied cognitive psychology*, 16(8), 897-910.
- Fonseca, F. (2001) *Ontology-driven geographic information systems*. Ph. D., Department of Spatial Information Science and Engineering, Orono, University of Maine. p. 117.

- Fonseca, F., Egenhofer, M.J. et Davis Jr., C.A. (2000) Ontology-driven information integration. *In: Proceeding of the AAAI-2000 Workshop on Spatial and Temporal Granularity*, Austin TX.
- Fortin, C. et Rousseau, R. (1989) *Psychologie cognitive: Une approche de traitement de l'information*. Sainte-Foy, Télé-université; Presses de l'Université du Québec. 434 p.
- Frank, A.U. (1992) Qualitative spatial reasoning about distances and directions in geographic space. *Journal of Visual Languages and Computing*, 3, 343-371.
- Freundschuh, S.M. et Egenhofer, M.J. (1997) Human conceptions of spaces: Implications for gis. *Transactions in GIS*, 2(4), 361-375.
- Frutos, S. et Berger, C. (2001) Étude du rôle du niveau hiérarchique des catégories lors de la généralisation de propriétés. *In Cognito - Cahiers Romains de Sciences Cognitives*, 23, 43-50.
- Gerevini, A. et Renz, J. (1998) Combining topological and size constraint for spatial reasoning. *In: Proceeding of the Proceedings of the 4th International Conference on Principles and Practice of Constraint Programming*, Springer-Verlag. pp. 220-234.
- Golledge, R.G. (1976) Methods and methodological issues in environmental cognition research. *In: Environmental knowing* (eds. Moore, G.T. et Golledge, R.G.), pp. 300-314. Dowden, Hutchinson & Ross Inc.
- Golledge, R.G. (1999) Human wayfinding and cognitive maps. *In: Wayfinding behavior: Cognitive mapping and other spatial processes* (ed. Golledge, R.G.), pp. 5-45. Baltimore, Johns Hopkins University Press.
- Golledge, R.G. (1999) *Wayfinding behavior: Cognitive mapping and other spatial processes*. Baltimore, Johns Hopkins University Press. 428 p.
- Golledge, R.G. et Stimson, R.J. (1997) *Spatial behavior: A geographic perspective*. New York, The Guilford Press. 620 p.
- Goodchild, M.F. (1990) Keynote address: Spatial information science. *In: Proceeding of the 4th International Symposium on Spatial Data Handling, SDH'90* (eds. Kishimoto, H. et Brassel, K.), Zurich, pp. 3-14.
- Goodchild, M.F. (2001) A geographer looks at spatial information theory. *In: Proceeding of the COSIT 2001, The fifth international Conference On Spatial Information Theory* (ed. Montello, D.R.), Californie, Springer-Verlag Berlin Heidelberg. Lecture Notes in Computer Science, 2205, pp. 1-13.
- Gould, P. et White, R. (1986) *Mental maps*. 2 edn. Harmondsworth, Penguin. 172 p.
- Gouteux, S. et Spelke, E.S. (2001) Children's use of geometry and landmarks to reorient in an open space. *Cognition*, 81, 119-148.
- Gouvernement du Québec (2003) Le grand dictionnaire terminologique. *In: Office de la langue française*. [En ligne]. <http://www.granddictionnaire.com/> (Page consultée le 26/02/2003).
- Grabowski, J. (1999) A uniform anthropomorphological approach to the human conception of dimensional relations. *Spatial Cognition and Computation*, 1(4), 349-363.
- Grosz, B.J. et Kraus, S. (1993) Collaborative plans for group activities. *In: Proceeding of the IJCAI '93* (eds. Fisher, M. et Owens, R.), Chambéry, France, Springer. pp. 367-373.
- Haber, L., Haber, R.N., Penningroth, S., Novak, K. et Radgowski, H. (1993) Comparison of nine methods of indicating the direction to objects: Data from blind adults. *Perception*, 22, 35-47.
- Harnad, S. (1990) The symbol grounding problem. *Physica D*, 42, 335-346.

- Hazarika, S.M. et Cohn, A.G. (2001) A taxonomy for spatial vagueness, an alternative egg-yolk interpretation. *In: Proceeding of the COSIT 2001, The fifth international Conference On Spatial Information Theory* (ed. Montello, D.R.), Californie, Springer-Verlag Berlin Heidelberg. Lecture Notes in Computer Science, 2205, pp. 92-107.
- Herskovits, A. (1986) *Language and spatial cognition: An interdisciplinary study of the prepositions in english*. Cambridge, Cambridge University Press. 208 p.
- Hillier, B. (2001) A theory of the city as object. Or, how spatial laws mediate the social construction of urban space. *In: Proceeding of the Space Syntax 3rd International Symposium* (eds. Peponis, J., et al.), Atlanta, University of Michigan. pp. 2.1-2.28.
- Hochmair, H. et Frank, A.U. (2002) Influence of estimation errors on wayfinding-decisions in unknown street networks - analyzing the least-angle strategy. *Spatial Cognition and Computation*, 2(4), 283-313.
- Hoel, E., Menon, S. et Morehouse, S. (2003) Building a robust relational implementation of topology. *In: Proceeding of the 8th International Symposium, SSTD, Advances in Spatial and Temporal Databases* (eds. Hadzilacos, T., et al.), Santorini Island, Greece, Springer. Lecture Notes in Computer Science, 2750, pp. 508-524.
- Hunt, E. et Waller, D. (1999) *Orientation and wayfinding: A review*. Arlington, VA, Office of Naval Research. Report ONR technical N00014-96-0380.
- Huttenlocher, J., Hedges, L.V. et Duncan, S. (1991) Categories and particulars: Prototype effects in estimating spatial location. *Psychological Review*, 98(3), 352-376.
- Jensen, J. et Saalfeld, A. (1998) Spatial data acquisition and integration. *In: UCGIS White Paper*. [En ligne]. <http://www.cla.sc.edu/geog/ucgis/> (Page consultée le 07/07/2002).
- Jones, C.B. (1997) *Geographical information systems and computer cartography*. Harlow, England, Longman. 319 p.
- Jones, C.B., Alani, H. et Tudhope, D. (2001) Geographical terminology servers - closing the semantic divide. *In: Proceeding of the Meeting on Fundamental Questions in Geographical Information Science* (eds. Duckham, M. et Worboys, M.F.), Palace Hotel, Manchester, UK, pp. 141-160.
- Jones, C.B., Alani, H. et Tudhope, D. (2001) Geographical information retrieval with ontologies of place. *In: Proceeding of the COSIT 2001, The fifth international Conference On Spatial Information Theory* (ed. Montello, D.R.), Californie, Springer-Verlag Berlin Heidelberg. Lecture Notes in Computer Science, 2205, pp. 322-335.
- Kelso, J.A.S. (1995) *Dynamic patterns: The self-organization of brain and behavior (complex adaptive systems)*. Cambridge, MA, MIT Press. 360 p.
- Kettani, D. et Moulin, B. (1999) A spatial model based on the notions of spatial conceptual map and object's influence areas. *In: Proceeding of the COSIT 99 Spatial Information Theory - Cognitive and Computational Foundations of Geographic Information Science* (eds. Freksa, C. et Mark, D.M.), Stade, Germany, Springer-Verlag. Lecture Notes in Computer Science, 1661, pp. 401-416.
- Kettani, D. et Moulin, B. (1999) A wayfinding application based on the influence area spatial model. *In: Proceeding of the 12th International Florida Artificial Intelligence Research Society Conference, Special Track on Spatiotemporal Reasoning* (eds. Kumar, A.N. et Russell, I.), Orlando, Florida, pp. 422-430.
- Kettani, D. et Moulin, B. (1999) A spatial model based on the cognitive concept of influence area. *In: Proceeding of the Sixth International Colloquium on Cognitive Science (ICCS'99)*, Donostia (Spain), pp. 171-182.
- Kitchin, R. et Blades, M. (2002) *The cognition of geographic space*. London, I.B. Tauris. 241 p.

- Kosslyn, S.M. (1980) *Image and mind*. Cambridge, MA, Harvard University Press. 500 p.
- Kosslyn, S.M. (1994) *Image and brain - the resolution of the imagery debate*. Cambridge, MA, MIT Press. 528 p.
- Kosslyn, S.M., Cacioppo, J.T., Davidson, R.J., Hugdahl, K., Lovallo, W.R., Spiegel, D. et Rose, R. (2002) Bridging psychology and biology: The analysis of individuals in groups. *American Psychologist*, 57(5), 341-351.
- Kuipers, B. (1978) Modeling spatial knowledge. *Cognitive Science*, 2, 129-153.
- Kuipers, B. (1982) The map in the head metaphor. *Environment & Behavior*, 14(2), 202-220.
- Kuipers, B. (2000) The spatial semantic hierarchy. *Artificial Intelligence*, 119(1-2), 191-233.
- Kuipers, B., Tecuci, D.G. et Stankiewicz, B.J. (2003) The skeleton in the cognitive map: A computational and empirical exploration. *Environment & Behavior*, 35(1), 81-106.
- Kulik, L. (2001) A geometric theory of vague boundaries based on supervaluation. In: *Proceeding of the COSIT 2001, The fifth international Conference On Spatial Information Theory* (ed. Montello, D.R.), Californie, Springer-Verlag Berlin Heidelberg. Lecture Notes in Computer Science, 2205, pp. 44-59.
- Ladrière, J. (2002) Cybernétique. In: *Universalis multimédia version 7*. Paris, Encyclopaedia Universalis.
- Landau, B. et Jackendoff, R. (1993) What and where in spatial language and spatial cognition. *Behavioral and Brain Sciences*, 16, 217-238.
- Lang, E. (1993) The meaning of german projective prepositions: A two-level approach. In: *The semantics of prepositions: From mental processing to natural language processing* (ed. Zelinsky-Wibbelt, C.), p. 249-294. Berlin; New York, Mouton de Gruyter.
- Leung, Y., Leung, K.S. et He, J.Z. (1999) A generic concept-based object-oriented geographical information system. *International Journal of Geographical Information Science*, 13(5), 475-498.
- Levelt, W.J.M. (1982) Cognitive styles in the use of spatial direction terms. In: *Speech, place, and action: Studies in deixis and related topics* (eds. Jarvella, R.J. et Klein, W.), pp. 251-268. Chichester; New York, Wiley.
- Levinson, S.C. (1996) Frames of reference and molyneux's question: Crosslinguistic evidence. In: *Language and space* (eds. Bloom, P., et al.), pp. 109-169. Cambridge, MA, MIT-Press.
- Levinson, S.C., Kita, S., Haun, D.B.M. et Rasch, B.H. (2002) Returning the tables: Language affects spatial reasoning. *Cognition*, 84(2), 155-188.
- Ligozat, G. (1998) Reasoning about cardinal directions. *Journal of Visual Languages and Computing*, 9(1), 23-44.
- Lo, C.P. et Yeung, A.K.W. (2002) *Concepts and techniques of geographic information systems*. Upper Saddle River, Prentice Hall. 476 p.
- Lynch, K. (1960) *The image of the city*, The Technology Press and the Harvard University Press. 208 p.
- Mark, D.M. et Frank, A.U. (1991) *Cognitive and linguistic aspects of geographic space*. Dordrecht, Kluwer. 519 p.
- Mark, D.M. et Egenhofer, M.J. (1994) Calibrating the meanings of spatial predicates from natural language: Line-region relations. In: *Proceeding of the 6th International Symposium on Spatial Data Handling, SDH'94*, Edinburgh, Scotland, pp. 538-553.
- Mark, D.M. et Frank, A.U. (1996) Experiential and formal models of geographic space. *Environment and Planning, B*, 23, 3-24.
- Mark, D.M., Freksa, C., Hirtle, S.C., Lloyd, R. et Tversky, B. (1999) Cognitive models of geographical space. *International Journal of Geographical Information Science*, 13(8), 747-774.

- McNamara, T.P. (2003) How are the locations of objects in the environment represented in memory? *In: Spatial cognition iii, routes and navigation, human memory and learning, spatial representation and spatial reasoning* (eds. Freksa, C., et al.), pp. 174-191. Berlin, Springer.
- McNamara, T.P. et Diwadkar, V.A. (1997) Symmetry and asymmetry of human spatial memory. *Cognitive Psychology*, 34(2), 160-190.
- McNamara, T.P. et Shelton, A.L. (2003) Cognitive maps and the hippocampus. *Trends in Cognitive Sciences*, 7(8), 333-335.
- Medin, D.L. et Aguilar, C. (1999) Categorization. *In: The mit encyclopedia of the cognitive sciences (mitacs)* (eds. Wilson, R.A. et Keil, F.). Cambridge, Mass, Bradford Books.
- Mennis, J.L. (2003) Derivation and implementation of a semantic gis data model informed by principles of cognition. *Computers, Environment and Urban Systems*, 27(5), 455-479.
- Mennis, J.L. et Peuquet, D.J. (2000) A conceptual framework for incorporating cognitive principles into geographical database representation. *International Journal of Geographical Information Science*, 14(6), 501-520.
- Mennis, J.L. et Peuquet, D.J. (2003) The role of knowledge representation in geographic knowledge discovery: A case study. *Transactions in GIS*, 7(3), 371-392.
- Miller, G.A. (1956) The magic number seven, plus or minus two: Some limits on our capacity for processing information. *The Psychological Review*, 63(2), 81-97.
- Montello, D.R. (1997) The perception and cognition of environmental distance: Direct sources of information. *In: Proceeding of the COSIT 97, Spatial Information Theory: A theoretical basis for GIS* (eds. Hirtle, S.C. et Frank, A.U.), Pittsburgh, USA, Springer-Verlag. Lecture Notes in Computer Science, 1329, pp. 297-311.
- Montello, D.R. (1997) Unit 006 - human cognition of the spatial world. *In: NCGIA Core Curriculum in Geographic Information Systems*. [En ligne]. <http://www.ncgia.ucsb.edu/education/curricula/giscc/units/u006/u006.html> (Page consultée le 07/07/2002).
- Montello, D.R. (1998) A new framework for understanding the acquisition of spatial knowledge in large-scale environments. *In: Spatial and temporal reasoning in geographic information systems* (eds. Egenhofer, M.J. et Golledge, R.G.), pp. 143-154. New York, Oxford University Press.
- Montello, D.R., Lovelace, K.L., Golledge, R.G. et Self, C.M. (1999) Sex-related differences and similarities in geographic and environmental spatial abilities. *Annals of the Association of American Geographers*, 89(3), 515-534.
- Montello, D.R., Richardson, A.E., Hegarty, M. et Provenza, M. (1999) A comparison of methods for estimating directions in egocentric space. *Perception*, 28(8), 981-1000.
- Moratz, R., Renz, J. et Wolter, D. (2000) Qualitative spatial reasoning about line segments. *In: Proceeding of the 14th European Conference on Artificial Intelligence* (ed. Horn, W.), Berlin, Germany.
- Moulin, B., Gryl, A. et Kettani, D. (1997) Route descriptions based on the notions of spatial conceptual map and of object's influence areas. *In: Proceeding of the AAAI-97 Workshop on Temporal and Spatial Reasoning, Fourteenth National Conference on Artificial Intelligence*, Providence, Rhode Island, pp. 95-104.
- Nadel, L. (1999) Neural mechanisms of spatial orientation and wayfinding: An overview. *In: Wayfinding behavior: Cognitive mapping and other spatial processes* (ed. Golledge, R.G.), pp. 313-327. Baltimore, Johns Hopkins University Press.
- Openshaw, S. (1984) The modifiable areal unit problem. *In: Concepts and techniques in modern geography*. Norwich, Geobooks.

- Openshaw, S. (1991) A view on the gis crisis in geography, or, using gis to put humpty-dumpty back together again. *Environment and Planning, A*, 23, 621-628.
- Paivio, A. (1986) *Mental representations: A dual coding approach*. New York, Oxford University Press. 322 p.
- Palmer, S.E. (1978) Fundamental aspects of cognitive representation. In: *Cognition and categorization* (eds. Rosch, E. et Lloyd, B.B.), pp. 259-303. Hillsdale, N.J., Lawrence Erlbaum Associates.
- Papadias, D. et Kavouras, M. (1994) Acquiring, representing and processing spatial relations. In: *Proceeding of the 6th International Symposium on Spatial Data Handling, SDH'94*, Edinburgh, U.K., Taylor & Francis.
- Peuquet, D.J. (1994) It's about time: A conceptual framework for the representation of temporal dynamics in geographical information systems. *Annals of the Association of American Geographers*, 84(3), 441-461.
- Peuquet, D.J. (2002) *Representations of space and time*. New York, The Guilford Press. 380 p.
- Piaget, J. et Inhelder, B. (1948) *La représentation de l'espace chez l'enfant*. Paris, PUF. 581 p.
- Ploux, S. et Ji, H. (2003) A model for matching semantic maps between languages (french/english, english/french). *Computational Linguistics*, 29(2), 155-178.
- Pocock, D.C.D. (1978) The cognition of intra urban distances: A summary. *Scottish Geographical Magazine*, 94, 31-35.
- Portugali, J. (1996) *The construction of cognitive maps*, Kluwer. 365 p.
- Portugali, J. (2002) The seven basic propositions of sirn (synergetic inter-representation networks). *Nonlinear Phenomena in Complex Systems*, 5(4), 428-444.
- Portugali, J. et Omer, I. (2003) Systematic distortions in cognitive maps: The north american west coast vs. The (west) coast of israel. In: *Proceeding of the COSIT 2003, Spatial Information Theory: Foundations of Geographic Information Science* (eds. Kuhn, W., et al.), Kartause Ittingen, Switzerland, Springer-Verlag. Lecture Notes in Computer Science, 2825, pp. 93-100.
- Pullar, D. et Egenhofer, M.J. (1988) Toward formal definitions of topological relations among spatial objects. In: *Proceeding of the 3rd International Symposium on Spatial Data Handling, SDH'88* (ed. Marble, D.F.), Sydney, Australia, pp. 225-241.
- Pylyshyn, Z. (2003) Return of the mental image: Are there really pictures in the brain? *Trends in Cognitive Science*, 7(3), 113-118.
- Raper, J. (2000) *Multidimensional geographic information science*. London, Taylor & Francis. 320 p.
- Raubal, M. (2001) Ontology and epistemology for agent-based wayfinding simulation. *International Journal of Geographical Information Science*, 15(7), 653-665.
- Raubal, M. et Winter, S. (2002) Enriching wayfinding instructions with local landmarks. In: *Proceeding of the GIScience 2002, Second International Conference* (eds. Egenhofer, M.J. et Mark, D.M.), Boulder, CO, USA, Springer-Verlag Berlin Heidelberg. Lecture Notes in Computer Science, 2478, pp. 243-259.
- Rodríguez, A. et Egenhofer, M.J. (2004) Comparing geospatial entity classes: An asymmetric and contextdependent similarity measure. *International Journal of Geographic Information Science*, 18(3), 229-256.
- Rosch, E. (1978) Principles of categorization. In: *Cognition and categorization* (eds. Rosch, E. et Lloyd, B.B.), pp. 27-48. Hillsdale, N.J., Lawrence Erlbaum Associates.
- Roussopoulos, N., Faloutsos, C. et Sellis, T. (1988) An efficient pictorial database system for psql. *IEEE Transactions on Software Engineering*, 14(5), 639-650.

- Roy, A.J. et Stell, J.G. (2001) Spatial relations between indeterminate regions. *International Journal of Approximate Reasoning*, 27, 205-234.
- Samet, H. et Aref, W.G. (1994) Spatial data models and query processing. In: *Modern database systems: The object model, interoperability, and beyond* (ed. Kim, W.), pp. 339-360. Reading, MA, Addison Wesley/ACM Press.
- Sasse, M.A. (1997) *Eliciting and describing users' models of computer systems*. Ph. D., Faculty of Science of the University of Birmingham, Birmingham, The University of Birmingham.
- Scholl, M.O. (1996) *Sgbd géographiques: Spécificités*. Paris; Toronto, International Thomson Pub. France. 185 p.
- Schuurman, N. (2000) Trouble in the heartland: Gis and its critics in the 1990s. *Progress in Human Geography*, 24(4), 569-590.
- Sebillo, M., Tortora, G. et Vitiello, G. (2000) The metaphor gis query language. *Journal of Visual Languages and Computing*, 11, 439-454.
- Séguin, A.-M., Thériault, M., Thomas, C. et Aubé, Y. (2000) Une approche géomatique pour l'analyse rétrospective des lieux de résidence et de travail. In: *Proceeding of the Géomatique 2000*, Montréal.
- Shekhar, S. et Chawla, S. (2003) *Spatial databases: A tour*. Upper Saddle River, N.J., Prentice Hall. 262 p.
- Shemyakin, F.N. (1962) General problems of orientation in space and space representations. In: *Psychological science in the u.S.S.R.* (ed. Ananyev, B.G.), pp. 186-251. Washington, D.C, U.S. Office of Technical Reports.
- Sheppard, E., Couclelis, H., Graham, S., Harrington, J.W. et Onsrud, H. (1999) Geographies of the information society. *International Journal of Geographical Information Science*, 13(8), 797-823.
- Sholl, M.J. (1996) From visual information to cognitive maps. In: *The construction of cognitive maps* (ed. Portugali, J.), pp. 157-186. Netherlands, Kluwer.
- Sholl, M.J. et Nolin, T.L. (1997) Orientation specificity in representations of place. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 23(6), 1494-1507.
- Siegel, A.W. et White, S.H. (1975) The development of spatial representations of large-scale environments. In: *Advances in child development and behavior* (ed. Reese, H.W.), pp. 9-55. New York, Academic Press.
- Smith, B. (1996) Mereotopology: A theory of parts and boundaries. *Data and Knowledge Engineering*, 20, 287-303.
- Smolensky, P. (1988) On the proper treatment of connectionism. *Behavioral and Brain Sciences*, 11, 1-74.
- Snodgrass, R.T., Böhlen, M.H., Jensen, C.S. et Steiner, A. (1998) Transitioning temporal support in sql2 to sql3. In: *Temporal database management* (ed. Jensen, C.S.), pp. 615-663
- Sorrows, M.E. et Hirtle, S.C. (1999) The nature of landmarks for real and electronic spaces. In: *Cosit 99 spatial information theory - cognitive and computational foundations of geographic information science* (eds. Freksa, C. et Mark, D.M.), pp. 37-50. Stade, Germany, Springer-Verlag.
- Spencer, J.P. et Hund, A.M. (2002) Prototypes and particulars geometric and experience-dependent spatial categories. *Journal of Experimental Psychology: General*, 131(1), 16-37.
- Spiekermann, K. et Wegener, M. (2000) Freedom from the tyranny of zones: Toward new gis-based spatial models. In: *Spatial models and gis; new potential and new models* (eds. Fotheringham, A.S. et Wegener, M.), pp. 45-61. London, Taylor & Francis.

- Stacey, M. (1996) Spatial expressions in design idea capture languages. *In: Proceeding of the AID'96 Workshop in Visual Representations, Reasoning and Interaction in Design* (eds. Damski, J.C.B. et Narayanan, N.H.).
- Stell, J.G. (1999) Granulation for graphs. *In: Proceeding of the COSIT 99 Spatial Information Theory - Cognitive and Computational Foundations of Geographic Information Science* (eds. Freksa, C. et Mark, D.M.), Stade, Germany, Springer-Verlag. Lecture Notes in Computer Science, 1661, pp. 417-432.
- Stevens, A. et Coupe, P. (1978) Distortions in judged spatial relations. *Cognitive Psychology*, 10, 422-437.
- Talmy, L. (1983) How language structures space. *In: Spatial orientation: Theory, research, and application* (eds. Pick, H.L. et Acredolo, L.P.), p. 378. New York, Plenum Press.
- Talmy, L. (2001) How spoken language and signed language structure space differently. *In: Proceeding of the COSIT 2001, The fifth international Conference On Spatial Information Theory* (ed. Montello, D.R.), California, Springer-Verlag Berlin Heidelberg. Lecture Notes in Computer Science, 2205, pp. 247-262.
- Taylor, H.A. et Tversky, B. (1996) Perspective in spatial descriptions. *Journal of Memory and Language*, 35(3), 371-391.
- Taylor, P.J. (1990) Gks. *Political Geography Quarterly*, 9, 211-212.
- Theobald, D.M. (2001) Topology revisited: Representing spatial relations. *International Journal of Geographical Information Science*, 15(8), 689-705.
- Thériault, M., Séguin, A.-M., Lee-Gosselin, M. et Thomas, C. (2000) Improving geomatic approaches for geocoding locations using addresses: Its use for socio-economic and transportation surveys. *In: Proceeding of the GEOIDE Network Annual Conference*, Calgary.
- Thériault, M., Joerin, F., Villeneuve, P. et Bégin, F. (2001) Modelling accessibility using fuzzy logic within transportation geographical information system. *In: Proceeding of the IASTED International Conference on Modelling Identification and Control 2001*, Anaheim, California, ACTA Press. 1, pp. 268-275.
- Thorndyke, P.W. et Hayes-Roth, B. (1982) Differences in spatial knowledge acquired from maps and navigation. *Cognitive Psychology*, 14(4), 560-589.
- Timpf, S. et Kuhn, W. (2003) Granularity transformations in wayfinding. *In: Spatial cognition iii, routes and navigation, human memory and learning, spatial representation and spatial reasoning* (eds. Freksa, C., et al.), pp. 77-88. Berlin, Springer.
- Tolman, E.C. (1948) Cognitive maps in rats and men. *The Psychological Review*, 55(4), 189-208.
- Tomlin, C.D. (1983) *Digital cartographic modeling techniques in environmental planning*. Ph. D., School of Forestry and Environmental Studies, New Haven, CT, Yale University.
- Tomlin, C.D. (1990) *Geographic information systems and cartographic modelling*. Englewood Cliffs, NJ, Prentice Hall. 249 p.
- Tryfona, N. et Jensen, C.S. (1998) *A component-based conceptual model for spatiotemporal design*, ChoroChronos, A Network for Spatiotemporal Database Systems. *Report Technical report ch-98-10*.
- Turing, A.M. (1937) On computable numbers, with an application to the entscheidungsproblem. *Proceedings of the London Mathematical Society*, 2(42), 230-265.
- Turing, A.M. (1950) Computing machinery and intelligence. *Mind*, 59(236), 433-460.
- Turvey, M.T., Carello, C. et Kim, N.-G. (1990) Links between active perception and the control of action. *In: Synergetics of cognition* (eds. Haken, H. et Stadler, M.), pp. 269-295. Heidelberg, Germany
- Tversky, B. (1992) Distortions in cognitive maps. *Geoforum*, 23(2), 131-138.

- Tversky, B. (1993) Cognitive maps, cognitive collages, and spatial mental models. *In: Proceeding of the COSIT 93, European Conference on Spatial Information Theory* (eds. Frank, A.U. et Campari, I.), Marciana Marina, Italy, Springer-Verlag. Lecture Notes in Computer Science, 716, pp. 14-24.
- Tversky, B. et Lee, P.U. (1998) How space structures language. *In: Spatial cognition, an interdisciplinary approach to representation and processing spatial knowledge* (eds. Freksa, C., et al.), pp. 157-176. Berlin Heidelberg, Springer-Verlag.
- Tversky, B., Taylor, H.A. et Mainwaring, S. (1997) Langage et perspective spatiale. *In: Langage et cognition spatiale* (ed. Denis, M.), pp. 25-49. Paris, Masson.
- Tversky, B., Franklin, N., Taylor, H.A. et Bryant, D.J. (1994) Spatial mental models from descriptions. *Journal of the American Society for Information Science*, 45(9), 656-669.
- van der Zee, E. et Slack, J.M. (2003) *Representing direction in language and space*. Oxford; Toronto, Oxford University Press. 282 p.
- Vandersmissen, M.-H. (2000) *Mobilité géographique et professionnelle des femmes dans la région urbaine de québec, 1977-1996*. Ph. D., Aménagement du territoire et développement régional, Québec, Université Laval. p. 154.
- Varela, F.J., Thompson, E.T. et Rosch, E. (1991) *The embodied mind: Cognitive science and human experience*. Cambridge, Mass, MIT Press. 299 p.
- Vishton, P.M. et Cutting, J.E. (1995) Wayfinding, displacements, and mental maps: Velocity fields are not typically used to determine one's aimpoint. *Journal of Experimental Psychology: Human Perception and Performance*, 21(5), 978-995.
- Vorweg, C. (2003) Use of reference directions in spatial encoding. *In: Spatial cognition iii, routes and navigation, human memory and learning, spatial representation and spatial reasoning* (eds. Freksa, C., et al.), pp. 321-347. Berlin, Springer.
- Vorweg, C. et Rickheit, G. (1998) Typicality effects in the categorization of spatial relations. *In: Spatial cognition, an interdisciplinary approach to representation and processing spatial knowledge* (eds. Freksa, C., et al.), pp. 203-222. Berlin, Springer-Verlag.
- Wang, F. (2003) Handling grammatical errors, ambiguity and impreciseness in gis natural language queries. *Transactions in GIS*, 7(1), 103-121.
- Waxman, S.R., Lynch, E.B., Casey, K.L. et Baer, L. (1997) Setters and samoyeds: The emergence of subordinate level categories as a basis for inductive inference in preschool-age children. *Developmental Psychology*, 33(6), 1074-1090.
- Weisman, J. (1981) Evaluating architectural legibility: Wayfinding in the built environment. *Environment & Behavior*, 13(2), 189-204.
- Wender, K.F., Haun, D.B.M., Rasch, B.H. et Blümke, M. (2003) Context effects in memory for routes. *In: Spatial cognition iii, routes and navigation, human memory and learning, spatial representation and spatial reasoning* (eds. Freksa, C., et al.), pp. 209-231. Berlin, Springer.
- Werner, S. et Habel, C. (1999) Spatial reference systems. *Spatial Cognition and Computation*, 1(1), iii-vii.
- Werner, S. et Schmidt, K. (1999) Environmental reference systems for large-scale spaces. *Spatial Cognition and Computation*, 1(4), 447-473.
- Wiener, N. (1948) *Cybernetics, or control and communication in the animal and the machine*. Cambridge, MA, MIT Press. 194 p.
- Winter, S. (2000) Uncertain topological relations between imprecise regions. *International Journal of Geographical Information Science*, 14(5), 411-430.
- Worboys, M. (1994) Object-oriented approaches to geo-referenced information. *International Journal of Geographical Information Systems*, 8(4), 385-399.

- Worboys, M.F. (1998) Computation with imprecise geospatial data. *Computers, Environment and Urban Systems*, 22, 85-106.
- Worboys, M.F. et Clementini, E. (2001) Integration of imperfect spatial information. *Journal of Visual Languages and Computing*, 12, 61-80.
- Zhan, F.B. (2001) A fuzzy set model of approximate linguistic terms in descriptions of binary topological relations between simple regions. *In: Soft computing in defining spatial relations* (eds. Sztandera, L. et Matsakis, P.), pp. 179-199. Heidelberg, Germany, Physica-Verlag.

Annexes

Annexe 1 – Liste des bases de données intégrées pour l'enquête OD de la région de Québec en 2001 par le RTC et le MTQ

	Nom de la base de données	Fournisseur	Année	Type de géoréférence
Répertoire de services	<ul style="list-style-type: none"> Répertoire des entreprises Zipcom 	Tamec inc.	2001	Adresse civique
	<ul style="list-style-type: none"> Répertoire téléphonique* 	Vernon Business Source ver. 4	2000	Adresse civique
	<ul style="list-style-type: none"> Répertoire des grands générateurs de déplacement 	ministère des Transports du Québec et de la RTC	2001	Adresse civique
	<ul style="list-style-type: none"> Répertoire des centres commerciaux 	CUQ et CRAD	2000	Adresse civique
	<ul style="list-style-type: none"> Répertoire des écoles 	CRAD	2000	Adresse civique
Répertoire de localisation	<ul style="list-style-type: none"> Rôle d'évaluation des propriétés foncières 	Communauté urbaine de Québec (CUQ)	1997	Adresse civique et positionnement géographique
	<ul style="list-style-type: none"> Codes postaux à six positions (version révisée) 	Statistique Canada révisée par le CRAD	1994/1996	Centroides des zones postales
	<ul style="list-style-type: none"> Réseau routier 	CRAD 1 : 20 000	1996	Topologique
	<ul style="list-style-type: none"> Réseau routier ** 	Statistique Canada	1996	Avec nomenclature
	<ul style="list-style-type: none"> Réseau routier 	Géocom 1 : 50 000	2000	Avec nomenclature et plages d'adresses
	<ul style="list-style-type: none"> Topos et toponyme populaire 	Commission de la toponymie du Québec	2001	Localisation en longitude et latitude en degré, minute (centroides et description textuelle des zones)
	<ul style="list-style-type: none"> Carte topographique 1 :20 000 	MRN Québec	1992/1998	Carte topographique en longitude et latitude non-structurée
	<ul style="list-style-type: none"> Découpage du recensement 1996 (SD, SR et municipaux) 	Statistique Canada révisé par le CRAD 1 : 250 000 et 1 :20 000 pour l'hydrographie	1996	longitude et latitude
Répertoire à construire	<ul style="list-style-type: none"> Lieux-dits 	Cartes 1 : 20 000, Topos, parc, immeubles		
	<ul style="list-style-type: none"> Lieux publics 	Villes, Cartes 1 : 20 000 et Topos	2001	Toponyme et coordonnées géographiques
	<ul style="list-style-type: none"> Quartier 	Villes	2001	(papier) Non géoréférencé
	<ul style="list-style-type: none"> CLSC et services de santé 	MTQ	2001	Adresse civique
	<ul style="list-style-type: none"> Paroisse 	Diocèse de Québec	2001	(papier) Non géoréférencé

* permettra d'enrichir et de valider le répertoire des entreprises

** nomenclature sans plage d'adresse

Annexe 2 – Modèle de données

Modèle original de l'enquête OD 2001. Liste des variables.

CHAMP	TYPE	DESCRIPTION
IdunikA	Int	Identifiant unique de générateur
GeneratA	C60	Nom du générateur (source)
Gener_nom	C60	Nom validé du générateur [pour les enregistrements utilisés soit ceux pour lesquels FreqTot>0 seulement]
Nciv	SmInt	Numéro civique du générateur
Rue	C36	Nom de rue normalisé (format selon séquence "spécifique générique direction")
sdr_nom	C45	Nom de municipalité simplifié
sdr2001C	C5	Code de municipalité 2001 (format caractères)
sdr2001N	Int	Code de municipalité 2001 (format numérique)
FreqTot	Int	Fréquence d'utilisation du générateur (compilation préliminaire, travaux en cours)
Xmtm83	D9,2	Longitude en format MTM (Nad83)
Ymtm83	D10,2	Latitude en format MTM (Nad83)
Longx83	D11,6	Longitude en degrés décimaux (Nad83)
Laty83	D10,6	Latitude en degrés décimaux (Nad83)

Annexe 3 – Glossaire

- Axiomatique** : La méthode axiomatique est un mode d'exposition des sciences exactes fondé sur des propositions admises sans démonstration et nettement formulées et des raisonnements rigoureux. (Encyclopaedia Universalis, 2001)
- Déictique** : Unité linguistique dont la signification est déterminée par la situation de communication. Voir aussi Embrayeur (Par exemple, je, nous, vous, aujourd'hui et ici sont des embrayeurs puisqu'ils renvoient au locuteur, au destinataire, au moment ou au lieu de l'énonciation. Alors qu'autrefois le terme français déictique et les termes anglais deictic et deictic word étaient employés pour désigner un embrayeur renvoyant plus particulièrement à la situation spatiale et temporelle, ils sont aujourd'hui considérés comme synonymes d'embrayeur et de shifter. (Gouvernement du Québec, 2003)
- Modèle** : le modèle est l'incarnation concrète de la théorie. La relation avec la théorie est qu'il satisfasse les postulats de la théorie. Parce qu'il y a plusieurs façons pour qu'une théorie soit satisfaite, il y a plusieurs modèles possibles (Palmer, 1978).
- Méréotopologie** : Amalgame de la méronymie et de la topologie qui permet de formuler certaines lois ontologiques concernant les frontières, l'intérieur d'un tout, les relations de contact et de connexité aux concepts de surfaces, points et voisinages entre autres (Smith, 1996).
- Méronymie** : relation conceptuelle partie-tout (Gouvernement du Québec, 2003)
- Ontologie** : Partie de la métaphysique qui s'applique à « l'être en tant qu'être » (Aristote), indépendamment de ses déterminations particulières. [Office de la langue française, 1991]. Selon Smith (Smith, 1996) l'ontologie est une science qui traite de la nature et de l'organisation de la réalité.
- Paronymie** : ensemble des relations méronymiques implantée de la même façon que la taxinomie. (Griot 1997)
- Taxinomie** : Construction d'un plan de classification de concepts utilisant des classes disjointes de concepts agrégés. (L'emploi du terme taxinomie est à éviter en informatique) (Gouvernement du Québec, 2003)
- Taxonomie** : Science de la classification caractérisée par une structure arborescente ou hiérarchisée dont les différentes catégories sont prénommées « taxons », surtout utilisées pour la classification des organismes vivants en bactériologie, en botanique et en zoologie, mais dont l'utilisation s'étend aujourd'hui à d'autres sciences, telles les sciences humaines et la psychologie. (Gouvernement du Québec, 2003)
- Théorie** : une théorie est essentiellement une description d'un phénomène à un certain niveau d'analyse. Elle exprime les lois structurelles qui sous-tendent l'objet d'étude à un niveau d'abstraction approprié. La théorie n'inclut pas les aspects plus concrets qui peuvent être vérifiés par observations empiriques, c'est le rôle du modèle (Palmer, 1978).

Topologie : Branche des mathématiques traitant des relations de voisinage qui s'établissent entre des figures géométriques, relations qui ne sont pas altérées par la déformation des figures. L'application des principes de la topologie en géomatique assure la cohérence des données géométriques et facilite leur utilisation parce qu'elle élimine la redondance des points, lorsqu'il s'agit de former des lignes, et la redondance des lignes, lorsqu'il s'agit de former des polygones. (Gouvernement du Québec, 2003)

Annexe 4 – WebKb

Exemple de résultat de la recherche du mot « building » et de l'utilisation de la catégorisation hiérarchique et de « partie de » dans cet outil.

4 categories (meanings) are recorded for "building".

Each category is related to other categories by links, e.g. "generalization".

To continue your search, click on the relevant source category or one linked to it.

Some links are hyperlinked too (see underlines). If you click on such a link,

you will also see the categories *indirectly* accessible by this link.

The identifier of each source category is given (the creator is specified

before the '#' except when the category come from WordNet 1.7).

For each destination category, the first name is displayed (not the identifier).

1. #building *the occupants of a building; "the entire building complained about the noise"*

supertype: gathering

No statement uses or specializes building; click here to add one.

2. #construction.creating_from_raw_materials (or: building) *the act of constructing or building something; "during the construction we had to take a detour"; "his hobby was the building of boats"*

supertype: creating_from_raw_materials

subtype: erecting, house-raising, fabrication, dry_walling, leveling, road_construction, shipbuilding, rustication.construction

No statement uses or specializes

construction.creating_from_raw_materials; click here to add one.

3. #construction.business_enterprise (or: building) *the commercial activity involved in constructing buildings; "their main business is home construction"; "workers in the building trades"*

supertype: business_enterprise

subtype: jerry-building

part: masonry, house_painting, plumbing, roofing, sheet-metal_work, shingling

No statement uses or specializes construction.business_enterprise; click here to add one.

4. #edifice (or: building) *a structure that has a roof and walls and stands more or less permanently in one place; "there was a three-story building on the corner"; "it was an imposing edifice"*

supertype: physical_structure

subtype: abattoir, airdock, apartment_building, architecture.edifice, aviary, bathhouse, bowling_alley, center, chapterhouse, clubhouse, farm_building, feedlot, firetrap, gambling_house, gazebo, government_building, greenhouse, hall.edifice, hall.academic_building,

[Hall_of_Fame](#), [hotel](#), [hotel-casino](#), [house](#), [house.edifice](#),
[library.edifice](#), [medical_building](#), [ministry](#), [morgue](#), [observatory](#),
[office_building](#), [opium_den](#), [outbuilding](#), [packinghouse](#),
[place_of_worship](#), [planetarium.edifice](#), [presbytery](#), [rest_house](#),
[skating_rink](#), [Roman_building](#), [rotunda.edifice](#), [ruin](#), [schoolhouse](#),
[shooting_gallery](#), [signal_box](#), [skyscraper](#), [student_union](#), [tavern](#),
[temple.edifice](#), [theater](#), [whorehouse](#)
instance: [Houses_of_Parliament](#)
part: [annex](#), [anteroom](#), [corner.structure](#), [wall_corner](#), [cornerstone](#),
[cornerstone.stone](#), [courtyard](#), [cullis](#), [elevator.lifting_device](#),
[exterior_door](#), [storey](#), [foundation_stone](#), [heating_system](#), [interior_door](#),
[roof](#), [room](#), [scantling](#), [shaft.passageway](#), [skeleton](#), [stairway](#), [wall](#),
[window](#), [crawlspce](#)

178 statements are about indirect instances of [edifice](#)

[click here to display them](#) or click [here for a search form](#) or [here to add a statement](#)

source consulté le 15 mars 2004 :

<http://meganesia.int.gu.edu.au/~phmartin/WebKB2/bin/categDisplay.cgi?categ=building>

SVP complétez ce formulaire qui servira au catalogage de votre thèse

Étudiant	Nom :	Girard
	Prénom :	Étienne
Directeur de recherche	Nom :	Thériault
	Prénom :	Marius
Codirectrice	Nom :	Trépanier
	Prénom :	Cécyle
Mots-clés	Géographie cognitive; Perception et nomenclature de l'espace; Système d'information géographique; Modélisation des références géographiques; Structuration et interrogation des références spatiales dans un SIG; Représentation des localisations dans une enquête de déplacements	