

Madeleine ARNOLD & Claude LEBRUN

Utilisation d'une langue pour la création de scènes architecturales en image de synthèse. Expérience et réflexions*

Introduction

La traduction d'un énoncé verbal en une image, mentale ou matérialisée à travers une représentation graphique, est un problème qui suscite un intérêt grandissant dans le champ des sciences cognitives, en particulier en psychologie cognitive, en linguistique, en sémiotique et en intelligence artificielle. Parallèlement aux recherches d'ordre théorique et méthodologique, des applications expérimentales sont mises en œuvre sur des systèmes informatiques où des énoncés verbaux sont transcrits automatiquement sous la forme d'images. La communication homme-machine en conception architecturale assistée par ordinateur (CAAO) constitue un terrain favorable pour tester les hypothèses, et pour opérationnaliser les résultats issus des recherches en traduction verbal-image. En effet si l'architecte utilise beaucoup l'image pour élaborer un

* Langue/langage : "Langage" est un terme générique. Un langage peut être défini comme un mode d'expression symbolique. Ainsi on parlera du langage verbal par opposition au langage iconique, gestuel ou architectural. Une langue — l'arabe, le chinois, le français, par exemple — est une manifestation particulière du mode d'expression qu'est le langage verbal. Elle est le produit d'une élaboration socio-culturelle.

projet, la parole écrite et orale est aussi un moyen de communication et d'expression majeur pour lui car il se trouve sans cesse confronté à la nécessité d'établir des correspondances entre énoncés verbaux et images, et de traduire ceux-ci dans celles-là (et vice versa). Un environnement de conception informatisé devrait donc, pour être convivial, comporter une interface permettant de passer aisément d'énoncés verbaux à des images (et vice versa).

Dans le texte qui suit, nous rappellerons tout d'abord les caractéristiques de la conception en général et de la conception architecturale en particulier, qui amènent à penser que l'utilisation d'une langue, en relation avec l'image, faciliterait la communication homme-machine en conception architecturale assistée par ordinateur. Puis nous aborderons la problématique de la traduction verbal-image telle qu'elle apparaît actuellement dans le champ de quatre sciences cognitives : la psychologie cognitive, la linguistique, la sémiotique et l'intelligence artificielle. Nous examinerons ensuite les questions qu'une interface iconique et verbale en CCAO pose aux sciences cognitives. Pour mieux cerner ces questions nous rendrons compte d'une expérience que nous avons entreprise et qui consiste à traduire automatiquement la description verbale d'une scène en un schéma de scène tridimensionnelle. Enfin les résultats et les perspectives de cette expérience seront discutés en liaison avec les questions posées précédemment.

***1. La conception architecturale assistée par ordinateur (CAAO).
De l'ergonomie des systèmes techniques à des questions cognitives***

Actuellement les systèmes infographiques mis à la disposition de l'architecte ne sont pas "intelligents". Ainsi ils ne fournissent pas d'aide à la planification non plus qu'aux raisonnements géométrique, spatial et temporel, et a fortiori au raisonnement architectural. Par ailleurs

l'architecte qui utilise un système infographique communique avec la machine par l'intermédiaire de menus, d'un curseur permettant de désigner des parties de l'image affichée sur un écran, ou d'un langage informatique. Ce dernier de même que le langage de commande, a peu à voir avec le langage opératif à la fois iconique et verbal, que l'architecte a l'habitude d'employer pour réaliser ses projets. Ceci tient à la difficulté d'intégrer des connaissances proprement architecturales dans les systèmes qui sont limités à un savoir d'ordre géométrique et algébrique, à faible contenu sémantique (Quintrand et al. 1985, p. 80 s. ; Jézéquellou 1989, pp. 4-5).

La recherche cognitive à visée ergonomique vise à développer un environnement de conception informatisé mieux adapté à l'opérateur, c'est-à-dire qui satisfasse aux spécificités d'une part, du processus de conception, et d'autre part à celles de la communication entre le concepteur et son projet via un langage opératif (Falzon 1989). La recherche d'une solution en conception est considérée comme un type particulier de résolution de problème. Le concepteur procède par cycles itératifs problème-solution-évaluation. Le problème, incomplètement défini au départ, est élaboré conjointement à la solution. Le concepteur fait appel à des schémas de solutions existants, personnels ou non, qu'il applique tels quels ou bien en les combinant ou en les adaptant. Sa recherche est soumise à des contraintes (de site, de forme, financières, etc.) qui lui ont été données au départ ou qu'il a construites au fil de son travail ; elles sont impératives ou facultatives et dans ce dernier cas c'est à lui de leur attribuer un ordre de préférence. Il planifie son activité de façon opportuniste en ce sens qu'il n'hésite pas à s'écarter d'un plan initial si cela lui paraît plus profitable (Hoc 1987, pp. 146-51 ; Falzon, Bisseret et al. 1990 ; Visser, 1990a et b ; Falzon et Darses 1991 ; Guéna 1991 ; Zreik 1991).

Dans le cas de la conception architecturale l'image joue un rôle majeur (Lebahar 1983 pour le dessin ; Porada 1990 et 1991 pour l'image de synthèse). L'image sert à simuler des faits, des idées, des choses. Elle représente les caractéristiques de l'objet architectural en les liant à des points de vue déterminés sur cet objet (morphologie, relations entre espaces, système constructif, etc.). En même temps, elle "exprime les structures intellectuelles de l'architecte (classes, relations, certitude/incertitude)" (Lebahar 1983, p. 86). Elle est support d'évaluations et "déclencheur[s]" d'idées (id. p. 81). Des schémas conceptuels sont matérialisés dans des images intermédiaires permettant ainsi l'élaboration et l'exploration visuelles d'hypothèses (Porada 1990, 1991). L'image devient l'instrument d'une "vision finalisée d'une scène architecturale". Le regard reconnaît et évalue, par exemple, des "effets architecturaux [...], un effet de profondeur, un effet d'axe de parcours, un effet de place ou encore de volumétrie d'espaces libres par opposition aux espaces construits" (Zeitoun 1988).

Conjointement à l'image, la parole joue un rôle important dans la conception. Le cahier des charges est écrit, éventuellement illustré par quelques schémas, et tout au long du projet l'architecte utilise la langue pour présenter son travail, en discuter et préciser les modifications à faire. Il est donc sans cesse confronté à la nécessité d'établir des correspondances entre des énoncés verbaux et des images ou de traduire les uns dans les autres. On attendrait donc d'une interface conviviale en CAO qu'elle soit multimodale (Binot, Falzon et al. 1990) et qu'elle permette en particulier de passer aisément d'un moyen d'expression à un autre, d'établir des correspondances entre un énoncé verbal et une image ou une partie d'image, et qu'elle assiste le concepteur dans la traduction verbal-image. Mais si l'intérêt pour le multimodal grandit dans le domaine de la CAO, l'interface verbal-image a reçu, à notre connaissance, peu d'attention jusqu'à ce jour . Nous évoquerons cependant deux

expériences, l'une néerlandaise et l'autre française. Une équipe d'Amsterdam a cherché à déterminer des termes du langage commun qui permettraient aux différents spécialistes intervenant dans un projet d'architecture de communiquer entre eux à travers un système de CAO (Lapré et Hudson 1988). A Rennes, en France, D. Jézequellou et ses collègues (Delepine et al. 1991) produisent des schémas de scènes tridimensionnelles à partir de la description verbale de ce que voit un individu, supposé circuler dans un environnement architectural. Pour le moment, le passage des phrases en français aux instructions en Prolog, qui créent la scène, n'est pas automatisé.

La création de scènes architecturales à l'aide d'une langue pose des questions d'ordre pluridisciplinaire et qui relèvent largement des sciences cognitives. Parmi celles-ci, la linguistique, la psychologie cognitive, la sémiotique et l'intelligence artificielle sont particulièrement concernées.

2. La problématique de la traduction verbal-image dans les sciences cognitives

Dans ce qui suit, il est fait état d'apports des sciences cognitives et de questions que ceux-ci provoquent ou qu'ils amènent à se poser en ce qui concerne les relations entre ces deux modes d'expression que sont le verbal et l'image. On notera que ces réflexions ne prétendent pas à l'exhaustivité et renvoient à des problèmes génériques sur lesquels existe une abondante littérature dans les domaines scientifiques concernés. On remarquera, par ailleurs, que si la nature de l'image, produite par un système informatique à partir d'un énoncé verbal, est physiquement explicite, elle doit être précisée lorsqu'il s'agit des autres domaines de recherche. En effet, ou bien l'image est mentale et échappe à la perception, ou bien elle est matérialisée — sous la forme d'un dessin,

d'un schéma ou d'une image vidéo par exemple —, et elle est alors observable directement.

Pour déterminer la signification spatio-visuelle d'expressions verbales — prépositions, locutions, modificateurs, verbes d'action — les linguistes comparent des phrases ou des fragments de phrases, où entrent ces expressions. Si la description de la signification spatio-visuelle est tout d'abord d'ordre topologique, géométrique, physique, projectif ou/et métrique (Herskovits 1986, p. 55) elle inclut aussi des connaissances sur le monde et sur le contexte extra-linguistique (id., p. 73 s. ; Vandeloise p. 22). La signification spatio-visuelle dans une langue présente des phénomènes analogues à ceux qui ont été observés pour la catégorisation cognitive. Une expression verbale a un noyau de signification d'où dérivent des types apparentés (Herskovits 1986). Certaines des recherches sur la signification spatio-visuelle d'énoncés verbaux se situent explicitement à la limite des phénomènes linguistiques et cognitifs (Borillo 1991). Ainsi le sujet d'A. Herskovits est-il le langage et la cognition, et l'auteur se réfère aux travaux de E. Rosch, de Piaget et Inhelder ainsi qu'aux recherches sur l'image mentale visuelle. J.-P. Desclés (1985 p. 7-8), pour sa part, considère un prédicat, verbe de mouvement ou de changement d'état en particulier, comme le produit de l'immersion dans un système linguistique, d'un schème conceptuel issu d'un archétype cognitif. Pour résoudre les problèmes de description que pose la partie figurable des expressions verbales, il est envisagé d'exploiter des modèles élaborés en compréhension automatique de l'image ou en robotique (Herskovits 1986, pp. 100-103 ; Jackendoff 1987).

En psychologie cognitive, beaucoup d'attention a été portée aux images mentales associées aux énoncés verbaux. Dans le cadre d'une étude comparée des deux systèmes de représentation symbolique dont

dispose l'homme, le système verbal et le système imagé, on s'est demandé quels étaient les natures et les fonctionnements relatifs de ceux-ci et s'ils étaient connectés ou non par l'intermédiaire d'un langage de haut niveau d'abstraction (Denis 1979, 1989). L'image mentale visuelle est considérée comme la figuration de la partie figurable de la signification. Elle facilite alors certains traitements (Denis 1989, pp. 129, 212-13) et constitue un facteur d'intégration des connaissances (id. p. 187). En ce qui concerne la compréhension des textes, l'image mentale apparaît comme un modèle de la situation, de nature spatio-visuelle ; elle est construite parallèlement à la représentation, de nature propositionnelle, du texte (Denhière et Baudet 1989 ; Denis 1989). Dans les recherches sur la catégorisation, des expressions verbales et des images matérialisées sont utilisées comme identificateurs de catégories cognitives relatives à des objets naturels ou artificiels, à des relations spatiales, à des actions ou à des scènes (Tversky et Hemenway 1983, Denis et Dubois 1988, Fleury et Dubois 1988, Cordier 1989, Mazet 1991). Le sujet établit une correspondance entre une expression verbale et une image, ou vice versa. Un des apports de ces expérimentations consiste en la détermination de caractéristiques des images mentales stéréotypées qui sont associées aux mots, ainsi que des modes d'organisation, d'activation et d'utilisation de celles-ci. Parmi les problèmes soulevés par les recherches menées on en relèvera quatre. Le premier concerne la traduction de la structure linéaire du texte en une structure spatiale bi- ou tridimensionnelle, et le deuxième, les influences respectives des contextes linguistique et d'activité sur la construction de l'image mentale. Les deux derniers se rapportent à la catégorisation : la catégorie se définit-elle intensionnellement ou bien par un spécimen typique auxquels ressemblent plus ou moins les autres éléments de la catégorie ? Il semble, par ailleurs que la catégorisation opère par référence à un niveau de base privilégié par les sujets auxquels s'ajoutent

des niveaux plus généraux ou plus fins, moins utilisés et variables selon l'expertise des sujets (Dubois 1991, pp. 31-54).

En psychologie cognitive comme en linguistique, et en sémiotique également, des recherches sont donc consacrées à la partie "figurable" des expressions verbales. Une distinction importante doit être faite cependant entre les approches de la psychologie et celle des deux autres disciplines. En effet, tandis que la première s'intéresse à la représentation mentale associée à l'énoncé verbal, la deuxième et la troisième sont concernées par la signification de celui-ci telle qu'elle est déterminée au sein d'un système d'expression (verbal pour la linguistique ; verbal, iconique ou autre, pour la sémiotique). Représentation mentale et signification ne sont pas deux phénomènes identiques (Rastier 1991, chap. 3 "Du concept au signifié"). Les catégories spatio-visuelles mentales ne se confondent pas avec les catégories spatio-visuelles associées à des énoncés verbaux ou bien à des dessins, des photos, des films, ou tout autre représentation matérielle imagée. On s'en rend compte lorsque l'on veut, par exemple, exprimer par un texte ou par des dessins ce que l'on a dans l'esprit. La langue et l'image s'affirment alors comme des instruments ayant leurs caractéristiques propres en termes de structures matérielles auxquelles sont corrélées des structures de signification. L'expression de la représentation mentale spatio-visuelle doit composer avec les contraintes imposées par ces structures.

Dans une approche analogue à celle de la linguistique du discours (et non seulement du mot et de la phrase) , la sémiotique vise à décrire les structures qui sous-tendent les deux aspects corrélés du discours : le sens et la manifestation matérielle de celui-ci à travers des sons, des traces graphiques, etc. La traduction d'un discours selon un mode d'expression en un discours selon un autre mode d'expression est considérée comme la transcription selon deux modalités différentes

d'une même structure profonde. Cela n'est pas sans rappeler la théorie propositionnelle en psychologie cognitive dans laquelle on considère que les systèmes cognitifs imagé et verbal dérivent d'un système commun plus abstrait de nature propositionnelle (Denis 1979, p. 45 ; 1989, p. 45). Mais, insistons sur ce point, il en est des phénomènes sémiotiques comme des phénomènes linguistiques : ils ne se confondent pas avec les phénomènes cognitifs...

Le passage de la structure profonde commune aux structures de surface spécifiques, relevant d'un mode d'expression ou d'un autre, s'appuie sur les différents niveaux d'abstraction du parcours génératif de la signification (Greimas et Courtés 1979 pp. 157-60 ; Floch 1990, pp. 121-126). Ainsi, en 1984, l'identité visuelle, verbale et architecturale de l'entreprise bancaire le Crédit du Nord a été redéfinie à partir du concept clé de clarté, dont le contenu a été établi à travers des relations d'opposition ou de proximité avec d'autres concepts. La trame du sens à transmettre par le logo, la papeterie, les affiches, la campagne de presse, l'architecture des agences, etc., ainsi que les principes spatio-visuels qui guidèrent le travail des créatifs, furent élaborés à partir du concept clé en utilisant le parcours génératif ainsi que le schéma narratif, modèle intermédiaire dans la génération des structures sémiotiques (Floch 1990, pp. 49-81). Un autre concept sémiotique, celui d'énonciation, joue un rôle important lorsqu'un discours porte sur l'espace. En effet, l'énonciateur, cette entité abstraite qui produit le discours et dont on trouve des traces dans celui-ci, introduit un point de vue sur les événements qu'il met en scène verbalement, iconiquement ou autrement. Le point de vue est lié à un observateur dans la scène ou hors scène et répond à une stratégie de la part de l'énonciateur vis-à-vis de l'énonciataire — lecteur, spectateur ou autre (Fontanille 1989). Les linguistes, qui étudient les prépositions spatiales en particulier recourent

également à la notion d'observateur pour rendre compte de certaines prépositions (Herskovits 1986 ; Vandeloise 1986).

En intelligence artificielle ou dans d'autres domaines informatiques tels que la reconnaissance des formes, des réflexions sont menées sur la transcription automatique verbal-image (et vice versa), ou dans cette perspective ; des systèmes expérimentaux sont réalisés (Larkin et Simon 1987 ; Arnold 1990 ; Faure 1991). Ceci donne lieu à l'exploitation de modèles ou de résultats issus de la linguistique et de la psychologie cognitive ou de la psychologie de la perception. Un problème clé, déjà soulevé à propos de la compréhension du texte en psychologie cognitive, est le passage d'une représentation séquentielle de l'information à une représentation spatiale. Le contenu informatif des deux types de représentation diffèrent, de même que les traitements de connaissances qu'ils permettent ou au contraire contrecarrent ou interdisent (Larkin et Simon 1987 ; Faure 1991, p. 8-14). Les "procédures perceptuelles" (Larkin et Simon 1987) par lesquelles un être humain "enrichit " ce qu'il voit, en identifiant un angle ou une symétrie ou en distinguant une figure de son fond, devraient jouer un rôle primordial non seulement pour articuler "représentations séquentielle et spatiale pour la résolution de problème" (Faure 1991, p. 14) mais aussi pour la transcription automatique verbal-image.

La transcription automatique d'un énoncé verbal en une image bi- ou tridimensionnelle (2D ou 3D), fixe ou animée (et vice-versa), est un sujet d'étude transversal à un certain nombre de domaines en intelligence artificielle ou proches de celle-ci : animation assistée, compréhension automatique de l'image, conception et fabrication assistée par ordinateur, robotique, communication homme-machine (Arnold 1990). Les systèmes réalisés ont une visée générale ou bien sont destinés à apporter une aide à un professionnel. Dans le premier cas, le but est d'élucider les rapports

entre les deux modes de représentation, verbal et iconique, du monde physique (perçu ou imaginé). Dans le second, l'image et l'énoncé verbal sont utilisés pour : enregistrer, organiser et fournir des connaissances dans un domaine déterminé ; concevoir un objet ; résoudre un problème.

La traduction d'un énoncé verbal en une image, ou l'inverse, est effectuée par l'intermédiaire de représentations internes. Les problèmes clé consistent à réduire l'écart entre la représentation de l'énoncé et celle de l'image, et à trouver des procédures pour passer de l'une à l'autre, sachant que la représentation de l'image est essentiellement d'ordre quantitatif (géométrique et algébrique), et que la représentation de l'énoncé verbal est surtout d'ordre qualitatif (spatial et non spatial se mêlant, Arnold 1990).

Langue et image ont des rôles complémentaires. Une langue permet de formuler ce qui n'est pas visible dans le visuel, une relation de partie à tout par exemple. Elle peut décrire l'aspect fonctionnel d'une morphologie, l'action d'un humain dans un environnement physique, par exemple, ou simuler qualitativement des situations statiques ou dynamiques. Elle est aussi un moyen de raisonner à différents niveaux d'abstraction. Par contre une représentation spatio-visuelle d'un objet ou d'une situation facilite l'intégration et la récupération des connaissances fournis par l'énoncé verbal sur eux. Elle supporte par ailleurs des inférences. Sachant qu'une personne était à la plage on en déduit qu'elle ne travaillait pas à ce moment là (Arnold 1989, p. 326). Dans l'exemple qui vient d'être donné et qui est illustratif d'un phénomène courant, catégorisation et enrichissement réciproque des données verbales et spatio-visuelles sont étroitement mêlés : une scène stéréotypée déclenche un événement stéréotypé qui lui est attaché.

3. Questions posées aux sciences cognitives par une interface verbale et iconique en CAAO

Pour aborder les problèmes que pose une interface verbale et iconique en CAAO nous nous appuyerons sur les triades sujet-instrument-objet et objet de pensée-objet graphiques-objet matériel, posées et exploitées en psychologie cognitive et plus particulièrement dans le domaine de l'apprentissage et de l'utilisation du dessin technique (Weill-Fassina et Rabardel 1987). Dans le cas d'une activité finalisée où le sujet recourt à ce type d'instrument particulier que sont les "outils sémiotiques" (Weill-Fassina 1982, p. 217) pour se représenter des objets ou des situations, ou pour agir sur ceux-ci, des interactions surviennent entre sujet, instrument et objet (Vérillon et Rabardel 1987, p. 115). En effet objets de pensée, objets graphiques et objets matériels ont les uns et les autres leurs structures et leurs modes de fonctionnement propres. Des concepts opératoires et des méthodes sont disponibles pour analyser les trois types d'objets. Ainsi les niveaux de représentation de l'espace définis génétiquement par Piaget — agi, pré-opératoire, opératoire concret, opératoire formel — ont été repris, retravaillés et étendus au sujet adulte (Weill-Fassina 1979, pp. 215-17 ; Vérillon et Rabardel 1987, p. 110). De la géométrie sont tirées les distinctions entre "propriétés topologiques, projectives ou métriques" de l'espace graphique (Weill-Fassina 1989). Piaget ayant aussi eu recours à la géométrie pour définir ses niveaux de représentation conceptuelle, les comparaisons entre objets graphiques et de pensée peuvent se faire à l'aide de catégories communes issues de cette discipline mathématique. La sémiotique a permis d'une part de faire le partage entre la signification spatiale ou non spatiale (fonctionnelle par exemple) d'un objet graphique et l'objet de pensée correspondant, et d'autre part de définir corrélativement les caractéristiques des deux aspects liés de l'objet graphique, le sens et la manifestation matérielle de celui-ci (Cuny et Hoc 1974 ; Weill-Fassina 1979, pp. 212-14).

Des traces observables, graphismes et propos tenus par le sujet pendant ou après son activité, est inférée la nature des représentations et des opérations mentales mises en jeu ainsi que des difficultés rencontrées dans l'ajustement entre objets de pensées, objets graphiques et objets matériels. Les décalages entre le niveau de représentation spatiale requis par le dessin et le niveau de représentation spatiale dont est capable le sujet est une source de difficultés qui a été bien étudié en ce qui concerne le dessin technique. Cette question pourrait aussi être considérée sous un autre angle, celui du concepteur qui a à choisir un type de dessin adapté au problème qu'il veut traiter. En effet selon qu'il s'agit de formes ou de dimensions, ou bien de fonctionnements, par exemple, un graphisme à caractère projectif et métrique ou au contraire topologique, conviendra mieux (Weill-Fassina 1989). On rejoint ici la réflexion de J. Bertin (1970) sur le bon usage des variables visuelles pour communiquer graphiquement différents types d'information. Un problème voisin et également important est celui de l'influence du type de dessin sur l'activité du sujet. Ainsi, un conducteur de voiture, confronté à différents types d'images, relèvera dans la représentation d'une scène routière ce qui l'intéresse en tant que conducteur ou bien des caractéristiques perceptives détachées de la situation de conduite (Dubois 1991, p. 45-47). Il s'attachera tantôt à l'aspect fonctionnel tantôt à l'aspect perceptif de la scène, selon qu'il s'agit de photos ou de dessins au trait, ces derniers étant proches des photos ou schématiques.

Des recherches sur l'impact des nouvelles technologies (conception assistée par ordinateur, machines outils à commandes numériques), indiquent des évolutions dans les relations sujet-instrument-objet par rapport à la pratique traditionnelle du dessin technique. On observe que les nouveaux systèmes demandent un haut niveau d'abstraction, et de coordination et de raisonnement spatiaux, ainsi qu'une planification

beaucoup plus stricte de l'activité (Lebahar 1987 pp. 367-76 ; Rabardel et Weill-Fassina 1987). Un type de géométrie particulier apparaît "qu'on pourrait appeler géométrie des facettes" (Osta 1987, p. 333) ; il caractérise un mode de construction des objets spatiaux très commun dans les logiciels. On remarquera, par ailleurs, qu'en conception assistée par ordinateur l'objet matériel non seulement est absent mais n'existe pas encore. Il est remplacé par une maquette virtuelle construite progressivement.

D'autres questions essentielles sont à ajouter à celles qu'étudient les psychologues spécialistes du dessin technique lorsqu'on s'intéresse à une interface en CAO permettant de créer des scènes architecturales à l'aide d'une langue. Elles concernent l'articulation de l'image matérialisée à l'énoncé verbal et à l'image mentale dans le cadre de l'activité finalisée qu'est l'élaboration d'un projet architectural. Il est probable que l'activité de catégorisation joue un grand rôle dans cette articulation. Dans la conception architecturale, en effet, les schémas spatio-visuels ont une fonction importante, qu'ils soient individuels ou collectifs, génériques ou attachés à des spécimens considérés comme exemplaires. Par ailleurs il existe un langage opératif verbal de l'architecture (Hanrot et al. 1988, p. 389 ; Lapré et Hudson 1988). Pour un architecte créant des scènes à l'aide d'une langue, quelle image mentale éveille tel énoncé verbal (d'un ou plusieurs mots) ou bien à quelle formulation recourt-il pour décrire telle image mentale ou tel entité graphique ? On peut supposer que les contextes linguistique, iconique et de l'activité, ont des effets soit rétroactifs soit anticipateurs sur les différents types de catégories impliquées et leur interaction.

Un autre problème important est celui de l'intégration des représentations et des traitements linguistiques, graphiques et mentaux dans une représentation globale de l'activité finalisée du concepteur.

Faut-il adopter une approche en termes de résolution de problème et de planification comme on le fait généralement en intelligence artificielle et en psychologie cognitive, ou bien faut-il considérer l'enchaînement des événements graphiques, verbaux (et autres éventuellement), comme un récit en souscrivant alors aux options de la sémiotique et de la compréhension automatique du récit écrit ? Il est probable qu'il n'y a pas à choisir entre les deux approches mais plutôt à s'inspirer de l'une et de l'autre comme cela se passe déjà en intelligence artificielle ou en psychologie cognitive où les modèles établis par R.C. Schank et ses collaborateurs pour le récit écrit sont appliqués à d'autres domaines tels que la catégorisation et l'acquisition de connaissances (Schank et Leake 1989 ; Tversky et Hemenway 1983, p. 141).

Pour mieux cerner les problèmes que poserait la création, en CAAO, de scènes architecturales à l'aide d'une langue, nous avons procédé à une application expérimentale. Notre but, ce faisant, n'a pas été d'élaborer une maquette informatique ou de vérifier méthodiquement des hypothèses cognitives définies, mais plutôt de préciser des questions déjà soulevées et éventuellement d'en découvrir de nouvelles. Cette expérience est un préalable et vise à déterminer des axes de recherche à effectuer avec des spécialistes de différents domaines des sciences cognitives. Elle se situe dans la perspective d'un environnement de conception informatisé doté d'une interface multimodale conviviale. Pour créer des scènes l'utilisateur aurait la possibilité de créer son propre vocabulaire spatio-visuel et architectural (objets, relations, états et actions), de le mémoriser et de le réutiliser dans des contextes variés, ou bien il pourrait faire appel à des éléments de vocabulaire élaborés par d'autres.

4. Une expérience

Nous nous sommes placés dans la situation d'un utilisateur qui voudrait se faire une première idée rapide de la scène qu'il veut créer. Pour ce faire il générerait un schéma de scène tridimensionnelle en décrivant verbalement la scène. Nous avons utilisé le langage de synthèse d'image IKOGRAPH, réalisé par Michel Bret au CIMA (fig. 1), et qui comporte des commandes de traitement de chaînes de caractères en plus de ses opérateurs graphiques. IKOGRAPH permet de fabriquer une maquette virtuelle d'une scène et d'afficher des images de celle-ci à partir de points de vue déterminés.

Figure 1. L'univers d'IKOGRAPH

Unité de mesure : le pixel, autrement dit un carreau de la mosaïque qu'est l'image.

Lum Position par défaut de la lumière qui éclaire la scène.

0,0,0 Coordonnées de l'origine des axes x, y et z, par rapport à laquelle sont construits et positionnés les objets de la scène.

Oeil Position par défaut de l'œil de l'observateur.

Fuite Point de visée de l'œil. Ses coordonnées sont 0,0,0 par défaut.

En général on choisit le plan xOy comme sol, les formes architecturales s'élevant du sol vers l'observateur, c'est-à-dire de zéro vers -z. Le résultat équivaut à une vue aérienne de la scène. Pour avoir une perspective avec un sol horizontal on change la position de l'œil, ou bien on fait une rotation du sol autour de l'axe des x.

Exemples d'opérateurs graphiques d'IKOGRAPH

Création d'objets : boule 50,10 (Créer une sphère dont le rayon est long de 50 pixels et qui compte 10 tranches) .

Modification d'objets : dilx.for 1 ; coe=30 (Dilater la forme 1 suivant l'axe des x ; coefficient de dilatation = 30) .

Opérations spatiales sur les objets : dep.for 3 ; poi=100,200,50 (Déplacer la forme jusqu'au point de coordonnées 100,200,50) .

4.1 Construction de la scène à partir du texte

L'expérience consiste à traduire automatiquement la description verbale d'une scène en une représentation graphique schématique par l'intermédiaire des instructions d'IKOGRAPH. Le texte de départ compte onze phrases spécifiant, en français, les positions relatives des composants de la scène (fig. 2). On y lit par exemple : "L'église est au centre de la place. La mairie est à gauche de l'église.[...] Des maisons bordent le côté droit de la rue."

- 1 L'église est au centre de la place.
- 2 La mairie est à gauche de l'église.
- 3 La place est entourée d'arbres.
- 4 La rue arrive à la place.
- 5 Des maisons bordent le côté droit de la rue.
- 6 A gauche il y a trois immeubles.
- 7 Il y a un magasin entre deux immeubles.
- 8 Une voiture est sur la place.
- 9 Une autre voiture est au centre de la place.
- 10 Trois voitures roulent dans la rue.
- 11 Un piéton marche sur le trottoir de droite.

Figure 2. Texte décrivant la scène

Figure 3. Tirage d'une image affichée à l'écran. La scène correspondant aux quatre premières phrases :

L'église est au centre de la place.

La mairie est à gauche de l'église.

La place est entourée d'arbres.

La rue arrive à la place.

La construction de la scène, décrite brièvement par le texte, est progressive (fig. 3). De chaque phrase sont extraites les informations nécessaires à la production infographique de la partie de la scène correspondante : entités architecturales, relation de positionnement entre les entités, nombre d'entités. Les informations issues de la phrase considérée sont enregistrées selon un moule pré-déterminé. Lorsque les objets et la relation sont déjà en mémoire ils sont rappelés et mis en scène automatiquement. Dans le cas contraire le concepteur rédige les procédures qui créeront les objets et établiront la relation spatiale entre eux lorsqu'elles seront exécutées. Si le résultat ne satisfait pas l'utilisateur celui-ci a la possibilité d'adapter les objets ou la relation ou d'en créer de nouveaux (annexes 1 et 2 : exemples de procédures linguistique et graphique).

Analyse linguistique du texte

MOULE	Quantification 1	Entité 1	Relation 1	Quantification 2	Entité 2
-------	------------------	----------	------------	------------------	----------

Construction infographique de la scène

On soulignera que le schéma de scène créé à l'aide d'IKOGRAPH n'exploite pas et de très loin toutes les possibilités plastiques et d'animation du langage. Il ne s'agit pas d'une scène présentant de riches effets de coloration, de texture, d'éclairage, ou des transparences, des reflets, des brouillards, mais d'une simple maquette "fil de fer" et statique. En effet, le but n'était pas de produire une représentation réaliste ou des effets perceptifs ou esthétiques, mais de se limiter au cas où le concepteur désire visualiser rapidement une scène pour avoir une première idée de l'ensemble. Par ailleurs, d'un point de vue méthodologique il est nécessaire de simplifier la scène pour réduire le nombre des problèmes à traiter.

4.2. *La modélisation et la formalisation des connaissances spatiales*

L'accent a été mis sur les relations spatiales, les entités architecturales — "maison", "mairie", "piéton", etc. — étant réduites à des formes géométriques simples. On remarquera que, dans le texte choisi, les expressions verbales concernant les relations, sont de nature descriptive alors que le langage IKOGRAPH est de nature opérative. Ne disposant que de commandes de création, de modification, de suppression et de déplacement, d'entités géométriques, il ne peut pas construire une scène à partir de sa description. Par contre il permet d'indiquer quelles actions doit effectuer le système pour réaliser la scène. Ainsi pour IKOGRAPH, la description "La maison est au bord de la rue" doit être traduite en des commandes correspondant à l'opération complexe "Mettre la maison au bord de la rue". Mais ces deux étapes - description verbale et commandes infographiques - ne sont que les extrémités du processus de modélisation et de formalisation qui mène du texte à sa transcription graphique. Le processus comporte en effet plusieurs stades intermédiaires qui interagissent fréquemment au fur et à mesure que s'élabore la transcription.

Analyse du contenu spatial des expressions verbales

A une expression verbale d'ordre spatial, est attaché un contenu spatial représentable par un modèle tridimensionnel partiel (Jackendoff 1987, p. 101 s.). Que signifie, par exemple, "border" dans "Des maisons bordent la rue", "Des arbres bordent la place" ? Des objets sont au bord, à la limite considérée comme un contour, d'un autre objet ; un intervalle assez régulier les séparent s'ils ne se touchent pas. S'il s'agit d'objets de la même sorte il se dégage une impression de répétition... Mais être à la limite est-ce être à l'intérieur ou à l'extérieur du contour, ou sur lui, en débordant de part et d'autre ? Etre au bord implique-t-il un contact avec le bord ou une distance peut-elle être ménagée entre le bord et les objets, mais alors quelle grandeur peut atteindre cette distance ? Dans le cas des arbres et de la place les objets sont dressés sur le bord et leur base est en contact avec la surface de celui-ci . Par ailleurs en quoi consiste le bord ? Pour une place c'est son contour dans son entier grosso modo, alors qu'une rue a deux bords qui sont ses deux plus longs côtés (fig. 4).

Différentes représentations possibles des deux phrases :

"Des maisons bordent la rue"

"Des arbres bordent la

place"

Figure 4. Schémas mentaux associés à "border"

Une fois élucidé le contenu spatial de l'expression verbale, il faut en préciser l'équivalent opératif. Comment "mettre la maison au bord de la rue", par exemple ? La maison est déplacée jusqu'au bord de la rue de telle sorte que la base de la façade de la maison se confonde avec le bord de la rue. Ou bien la base est amenée à une petite distance du bord de la rue parallèlement à celui-ci. Dans les deux cas la maison reste extérieure à la rue (fig. 5.1). On notera que la représentation spatiale, élaborée par un utilisateur à partir d'une expression verbale, peut être uniquement mentale ou être matérialisée de temps à autre sous la forme de schémas graphiques.

Figure 5. Modélisation et formalisation des connaissances spatiales. Les éléments pertinents selon les niveaux de modélisation sont marqués en gras

Géométrisation de l'image mentale

La représentation mentale des objets, des relations et des actions doit être géométrisée pour pouvoir être enregistrée et traitée par le système. L'utilisateur décide alors de la forme et de la dimension approximatives des objets : un rectangle allongé pour la rue et un parallélépipède rectangle pour une maison, par exemple, la longueur de la rue dépassant la plus grande dimension de la maison. Il identifie les opérations à effectuer et les éléments nécessaires à l'accomplissement de celles-ci. Le déplacement est décomposé en deux opérations successives : la translation amène le sommet A de la maison au sommet E de la rue ou à proximité du bord ; la rotation fait pivoter la maison autour de A jusqu'à ce que AB se confonde avec EF ou lui soit parallèle (fig. 5.2).

Formalisation de la représentation géométrique

Lorsque l'on passe à la représentation formelle dans le système de coordonnées du système infographique IKOGRAPH de nouveaux facteurs doivent être pris en compte. Formes et objets étant numérotés les numéros de la forme rue et du sommet E de celle-ci sont cherchés pour indiquer E comme point d'aboutissement de la translation. Segments et directions sont orientés et sont traités comme des vecteurs. Les valeurs des angles de rotation sont précisées. Alors que dans l'image mentale et sa géométrisation le fond de la maison et la rue apparaissent dans le même plan, il n'en est pas de même lorsque l'on passe à l'espace du système infographique. En effet, dans celui-ci, la position d'un objet est variable à la façon de celle d'un astronaute dans le vide. Il faut donc ajouter des rotations qui amèneront le fond de la maison dans le plan xOy de la rue. Celles-ci demandent d'une part que soit clairement identifiée la facette correspondant au fond de la rue, et d'autre part que soient précisément définies les rotations à effectuer (fig. 5.3).

L'arête AB ne suffit pas pour désigner la facette ABCD. En effet une seule arête ne permet pas de distinguer une facette des deux autres facettes qui forment un coin d'un parallélépipède avec elle. Deux autres arêtes doivent donc lui être ajoutées, AD et AP, qui forment coin avec elle (fig. 6). Une fois la facette ABCD identifiée, un moyen de la placer dans le plan xOy est tout d'abord d'amener A en O, puis de faire coïncider la perpendiculaire AP à la facette avec l'axe Oz. La facette est alors perpendiculaire à Oz en O et se trouve donc confondue avec le plan xOy . Cette seconde opération se décompose en deux rotations successives, l'une autour de l'axe des y, l'autre autour de celle des z .

Les angles de rotation sont calculées à partir des projections de OP sur les plans xOz et xOy. Les rotations posent des problèmes de représentation mentale et de raisonnement géométrique en trois dimensions du fait, en particulier, qu'elles se font dans des sens déterminés autour des axes x, y et z. Une fois que le plan des différents calculs et opérations à effectuer a été établi, l'utilisateur doit rédiger les instructions correspondantes dans le langage IKOGRAPH.

Figure 6. Formalisation de la représentation géométrique. Détails

5. Résultats de l'expérience et perspectives

L'expérience souligne la distinction à faire entre les deux aspects liés du langage opératif et leur enracinement dans le texte et dans l'image, mentale ou matérialisée. Les phrases et les images mentales de fragments de scènes schématiques qu'elles suscitent se rapportent à l'aspect descriptif du langage. Les opérations spatiales à imaginer sous une forme naïve, puis géométrique, et finalement infographique, se rapportent à l'aspect opératoire, autrement dit à la fabrication du schéma de scène à partir de sa description.

L'extraction, des phrases du texte, des informations nécessaires à la construction infographique de la scène posent un certain nombre de difficultés. Les principales portent sur les points suivants : la référence, la relation de partie à tout, la quantification, la synonymie, l'observateur.

- *La référence* : Lorsqu'il est fait référence à une entité directement (dénomination, pronom personnel, etc.) ou indirectement (adjectif possessif, dénomination d'une partie la dénomination de l'entité étant sous-entendue, etc.) il est nécessaire d'établir de quelle entité il s'agit et si elle n'a pas déjà paru dans les phrases précédentes.

- *Relation de partie à tout* : Un nom commun désigne soit une entité soit une partie d'entité comme le montrent les expressions "le côté droit de la rue" et "le trottoir de droite". Le traitement infographique requiert que la partie soit distinguée du tout, et que l'on détermine aussi, éventuellement, à quel objet elle appartient.

- *La quantification* : La construction infographique ne peut se faire que si toute quantité (nombre d'objets, grandeur d'une surface, localisation) apparaissant dans le texte est précise. Si la quantification est indéterminée dans le texte il faut la préciser soit par défaut soit en interrogeant l'utilisateur.

- *La synonymie* : Certaines expressions ont le même contenu spatial et d'autres des contenus voisins ou dérivés. Il est indifférent du point de vue de la construction infographique de dire *la rue arrive à, débouche sur, ou part de la place*. Dans le cas de *des arbres bordent la place* et *des arbres entourent la place*, *border* et *entourer* ont un sens identique. Par contre, dans *des arbres bordent la rue*, *border* a seulement un sens voisin de celui de *entourer*. Par ailleurs *border* et *entourer* se présentent comme des dérivés de *arriver à* du fait qu'ils intègrent le contenu équivalent à *être au bord de* et de *arriver à*.

- *L'observateur* : Certaines relations spatiales sont dépendantes d'un observateur comme dans le cas de la phrase "A gauche [de la rue] il y a trois immeubles". Il faut donc déterminer la position de l'observateur et la direction dans laquelle il regarde. On remarquera que ce problème ne se pose pas dans les maquettes construites par D.

Jézéquellou et ses collègues (Delepine et al. 1991) : la scène étant construite morceau par morceau en fonction de ce que voit un observateur avançant et découvrant la scène, des expressions comme “en face” ou “derrière” sont sans ambiguïté.

Ne disposant pas actuellement d'un analyseur syntaxique et sémantique, nous avons résolu les difficultés avec des moyens simples. Ainsi lorsqu'une entité est sous-entendue dans une phrase, le système demande si celle-ci est l'une ou l'autre entité de la phrase précédente. Sinon l'utilisateur indique quelle est l'entité concernée. Dans le cas d'une relation liée à un observateur, le problème est compliqué par la nécessité de positionner l'observateur et son regard par rapport aux axes de coordonnées du système. Nous avons simplifié la situation en posant que tout objet a un avant et qu'un observateur se tient debout à l'intérieur de l'objet (maison, rue, voiture, etc.) en regardant vers l'avant. Le côté droit de l'objet est à la droite de l'observateur et le côté gauche, à la gauche de celui-ci.

Les différents types de représentation et de traitement mis en œuvre par le langage opératif iconique, mental et matérialisé, sont définissables en recourant à des géométries diverses (topologique, descriptive, métrique) comme dans le cas du dessin technique. On note aussi l'importance d'une “géométrie des facettes”. Au stade le plus formel directement lié au langage infographique intervient la géométrie analytique à travers l'usage qui est fait des vecteurs. A ce niveau on relève également une distanciation par rapport au donné observable comme cela a été relevé pour la programmation des machines outils à commande numérique. Le sujet fait appel à des correspondances automatiques entre opérations et désignations symboliques de celles-ci (Lebahar 1987, p. 374). Ainsi lorsque l'on veut faire glisser une forme dans la direction d'un vecteur déterminé il arrive un moment où l'on sait, sans plus avoir à s'imaginer la situation spatiale, qu'il faut faire tels calculs dans tel ordre. Par ailleurs les contraintes de l'instrument, système de synthèse d'image, sur le fonctionnement cognitif du sujet sont sensibles en termes d'anticipation. On constate aussi l'influence du caractère schématique donné à la scène, but de l'activité. L'image mentale est pauvre iconiquement et ses formes sont fortement influencées par les possibilités géométriques du système.

Les traitements de représentations spatio-visuelles effectuées au cours de l'expérience, nous ont amenés à constater un curieux phénomène : la réduction du problème spatial tridimensionnel à un problème bidimensionnel. On a vu, en effet, que les images mentales attachées aux énoncés linguistiques, ainsi que leurs géométrisations, sont en trois dimensions. Mais lorsque l'on passe à la formalisation des représentations géométrisées on s'aperçoit qu'il est inutile, en général, d'agir sur les volumes et d'utiliser la troisième dimension de l'espace. Il suffit de travailler sur les facettes qui constituent les fonds des objets (fond d'un magasin ou d'une voiture, surface d'une place ou d'une rue, base d'un arbre ou d'un piéton, etc.) après les avoir placées dans le plan xOy . Ceci tient en fait aux

limites de notre univers. Le sol y est horizontal et confondu avec le plan xOy . Les objets sont soumis à la gravité : les murs des bâtiments sont verticaux de même que le piéton et l'arbre. Tous les coins sont à angle droit. Ainsi lorsque l'on pose le fond horizontal d'un bâtiment contre le fond horizontal d'un autre bâtiment, on est sûr que les murs mitoyens sont juxtaposés et ne s'interpénètrent pas. On remarquera que si ces différentes contraintes simplifient considérablement l'univers de la ville, elles sont aussi des caractéristiques fréquentes de celui-ci.

Un second phénomène s'est imposé à nous progressivement au cours de l'expérience : la réduction de plusieurs descriptions linguistiques à une seule description infographique, et vice versa. Connaissant le phénomène de la synonymie et le rapport qui existe entre structures superficielles et structures profondes en linguistique nous nous attendions à ce que des expressions linguistiques apparemment différentes recouvrent en fait la même situation spatiale ou inversement. Mais nous ne soupçonnions pas qu'un aussi petit nombre d'opérations nous permettraient de traiter l'ensemble du texte. Ainsi les quatre phrases suivantes :

“La mairie est à gauche de l’église.”

“La rue arrive à la place.”

“Des maisons bordent le côté droit de la rue.”

“Trois voitures roulent dans la rue.”

renvoient à la même relation spatiale en profondeur, et celle-ci est réalisée par la même suite d’opérations infographiques : le système fait coïncider un côté du fond de l’objet à positionner avec un côté du fond de l’objet de référence. Ensuite des variations sont introduites selon les objets et les situations concernés, ou bien en fonction du choix du concepteur : l’objet à positionner est écarté ou non de l’autre objet, ou/et il est multiplié et positionné le long du bord de l’objet de référence. Ainsi l’avant de la rue reste accolée au côté de la place ; par contre l’utilisateur a le choix entre accoler les maisons au bord de la rue ou bien les écarter de celui-ci. Dans le cas d’une voiture qui roule, celle-ci sera obligatoirement écartée du bord de la rue. On notera que le moyen d’expression par lequel on passe, ici l’image fixe, influe sur la réalisation graphique de la scène : les mouvements sont ramenés à des situations fixes. “Rouler dans la rue”, “marcher sur le trottoir”, reviennent à “être posé sur la rue, sur le trottoir, à une petite distance du bord de ceux-ci”.

Le petit nombre des opérations requises pour traduire graphiquement des énoncés verbaux déjà complexes sémantiquement, ne doit pas faire illusion cependant. La simplicité du niveau formalisé résulte de raisonnements complexes attachés aux niveaux intermédiaires de modélisation. En effet il ne s’agit pas de n’importe quels côtés des objets à positionner et de référence. Par ailleurs l’objet à positionner est tantôt à l’intérieur, tantôt à l’extérieur de l’autre objet. L’identification des côtés concernés et la définition précises des relations spatiales requièrent la mise en œuvre de connaissances du domaine, sur ce que sont une rue et un immeuble, sur le fait que c’est en général la façade de l’immeuble qui borde la rue, par exemple.

Si l’on a pu constater que différents énoncés sont transcrits en profondeur par les mêmes opérations, l’inverse arrive aussi. Ainsi en est-il des expressions “le côté droit de la rue” et “le trottoir de droite [de la rue]” qui manifestent verbalement la même relation spatiale de partie à tout. Dans le contexte linguistique où se trouvent ces deux énoncés, le côté droit est transcrit par une ligne qui limite la rue à droite et par rapport à laquelle sont situées des maisons, tandis que le trottoir est une surface sur laquelle marche un piéton. Le côté droit est un élément de base de la rue alors que le trottoir n’existe pas encore quand il apparaît dans le texte. Il faut donc créer une forme *Trottoir* qui a le même statut infographique que la forme *Rue*. Infographiquement le côté droit est une partie de la rue alors que le trottoir n’en est pas une. Pour mettre les maisons au bord de la rue, à droite, on dispose tout de suite du repère nécessaire qu’est le côté droit. Mais pour poser le

piéton sur le trottoir, il faut d'abord identifier le côté droit du trottoir qui sert de repère pour effectuer l'opération. Comme on le voit, le passage au niveau perceptif formalisé se traduit tantôt par une simplification tantôt par une complication de la modélisation. Ceci dépend du contexte linguistique mais aussi du contexte opératif. En effet, un autre texte aurait pu imposer, par exemple, de créer le trottoir en même temps que la rue et de le traiter infographiquement comme une partie de la rue. Le problème consistant à poser le piéton sur le trottoir aurait été différent car, dans ce cas-là, une forme et une partie de forme ne peuvent pas recevoir un traitement infographique identique.

La définition des représentations spatio-visuelles associées aux expressions verbales et leur traitement graduel pour passer de l'image mentale à la description infographique posent un double problème de catégorisation et de niveaux de modélisation. Ainsi, y a-t-il une représentation type d'une rue, d'un immeuble, de la relation "border", ou de l'opération "mettre contre" ? Une représentation une fois définie, comment va-t-elle être adaptée à un nouveau contexte linguistique ou opératif ? Qu'en est-il, par ailleurs, de l'interaction entre les différents types de catégories : celles qui relèvent de l'expression linguistique, de l'image mentale naïve ou géométrisée, du langage infographique ?

Le parcours génératif de la signification, en sémiotique, offre un dispositif utile pour distinguer les différents niveaux de représentation de l'espace mental et graphique ainsi que les différents types de catégories qui leur sont associées. Il nous semble devoir jouer un rôle complémentaire de celui des types de géométrie et des "registres de fonctionnement cognitifs" inspirés de Piaget (Weill-Fassina, Vermersch et Zougari 1987, p. 101) et qui sont utilisés en psychologie cognitive en particulier dans le domaine du dessin technique. Le parcours génératif comporte quatre niveaux d'abstraction applicables à tout objet ayant une signification (dessin technique, scène architecturale, protocole d'observation, etc.). Les deux niveaux les plus concrets, le figuratif et le thématique, sont spatio-temporalisés. Au niveau figuratif, acteurs, actions, relations et états, sont décrits comme on pourrait le faire dans la vie courante. Ce niveau n'est pas sans affinités avec le domaine du sens commun en intelligence artificielle ou le niveau de base en psychologie cognitive. Le niveau thématique constitue une première phase de généralisation des catégories du niveau figuratif. Au troisième niveau, narratif, les dimensions spatiale et temporelle disparaissent. Les états et les transformations qui affectent les abstractions d'acteurs et d'objets de ce niveau sont d'ordre logique. Le dernier niveau, le plus abstrait, est constitué par les deux types d'opposition qui structurent fondamentalement tout objet ayant une signification : l'opposition qualitative (entre éléments différents) et l'opposition privative (entre la présence et l'absence d'un élément).

A ces quatre niveaux, nous en avons ajouté deux. L'un est d'ordre perceptif et s'apparente au domaine de la reconnaissance des formes. Nous qualifierons l'autre de

“perceptif formalisé” si une telle expression peut être acceptée. A ce palier, les catégories de nature perceptive reçoivent la description formalisée qu’impose l’utilisation d’un système infographique. Dans le cas de notre expérience nous ne considérerons que les trois niveaux les plus concrets, le thématique, le figuratif et le perceptif, ainsi que le niveau perceptif formalisé. Au niveau thématique, “border”, par exemple, est un état semi-abstrait qui définit une relation spatiale schématique liant deux objets, autrement dit deux acteurs, vides. Au niveau figuratif, les configurations spatiales thématiques sont plongées dans des univers sémantiques définis, l’univers de l’urbain par exemple. Les acteurs vides, qui n’étaient jusqu’alors caractérisés que par leur fonction (border et être bordé), sont investis par des entités déterminés : la rue et les maisons, ou la place et les arbres. Le schéma spatial semi-abstrait “border” est spécifié et ajusté en fonction des particularités des entités intégrées. Au niveau perceptif, les configurations figuratives perdent la plus grande partie de leur sémantisme : une façade devient une surface rectangulaire, un arbre, une forme approximativement sphérique posée sur une forme approximativement cylindrique. La formalisation géométrique revient à une mise en œuvre, traitable informatiquement, du niveau perceptif. Sont pertinents : d’une part les opérations (déplacements et rotations) qui réalisent l’état dénommé “border” ; d’autre part les sommets et les facettes des surfaces polyédriques constituant les objets, et les vecteurs et les angles, qui interviennent dans les opérations. Le passage d’un niveau à un autre s’opère par enrichissement ou par appauvrissement des catégories spatio-visuelles concernées. Des traits descriptifs qui étaient pertinents à un niveau ne le sont plus à un autre, ou bien des traits qui ne l’étaient pas le deviennent.

En ce qui concerne l’intelligence artificielle nous retiendrons un point en particulier qui relève de l’acquisition des connaissances. Au fur et à mesure que l’on rédige des procédures, en effet, on se rend compte que l’on peut réutiliser certaines d’entre elles, entièrement ou partiellement. Cette prise de conscience peut d’ailleurs être antérieure à la formalisation elle-même sous l’influence d’un problème précédemment résolu. Ainsi dans le texte, la phrase “La rue arrive à la place” précède immédiatement la phrase “Des maisons bordent le côté droit de la rue”. Dans la construction infographique de la première phrase, une extrémité de la rue est posée contre un côté de la place. Le simple fait, à propos de la seconde phrase, de formuler la première étape de “border” comme un “poser contre” provoque un rapprochement immédiat avec l’opération définie pour “arriver à”. Par ailleurs certaines parties de procédures ont une portée générale qui amène à penser que l’on aura à utiliser souvent ces fragments. Cependant, pressé par le temps et désireux de voir tourner la procédure au plus vite, on ne les extrait pas pour en faire une procédure à part qu’il suffirait alors de rappeler par son nom pour qu’elle s’exécute. Tel est le cas, par exemple, du calcul de l’angle que font deux vecteurs.

Il serait intéressant que le système soit capable de repérer des fragments identiques et d'établir s'ils forment un tout correspondant à des entités ou à des opérations infographiques déterminées de façon à proposer leur extraction à l'utilisateur et leur remplacement par un appel à procédure dans les procédures où ils auraient à intervenir par la suite. Plus encore, serait-il en mesure de discerner des procédures analogues et d'analyser les raisons spatiales de leurs différences ? Ainsi l'expression "poser contre" pour un arbre se traduira par une tangence entre la section circulaire de la base du tronc et le bord contre lequel est dressé l'arbre, alors qu'elle sera transcrite par la fusion de deux vecteurs pour une forme à base parallélépipédique. De plus on utilisera un sommet du fond du parallélépipède comme repère dans le déplacement alors que ce sera le centre du cercle de base qui sera utilisé pour l'arbre. Les procédures correspondantes différeront d'autant. Des procédures analogues pourraient être enregistrées en précisant leurs conditions d'utilisation, inférées de leurs différences. Ainsi, confronté à une nouvelle relation spatiale impliquant un "poser contre", le système pourrait faire des suggestions en tenant compte des conditions.

6. Conclusion

Un problème essentiel auquel on s'est heurté au cours de l'expérience est celui de la modélisation qu'impose toute implantation informatique de connaissances quelles qu'elles soient. C'est alors que l'on peut tester la validité des modèles que l'on a élaborés. Si le système pose un immeuble en travers de la rue au lieu de le mettre au bord de celle-ci, ou bien on a négligé des contraintes techniques, ou bien des éléments pertinents de la relation spatiale "border" ont été oubliés dans le modèle, ou au contraire on a introduit à tort des caractéristiques propres à une autre situation spatiale. Un système informatique à base de connaissances est un instrument épistémologique redoutable comme le soulignent Marie-Salomé Lagrange et Monique Renaud (1987, pp. 217-18). Ainsi elles disent de leur système-expert sur les discours d'archéologues qu'il "a peut-être un rôle démonstratif : il expose machinalement les faiblesses du discours archéologique qui lui a donné naissance". Leur observation, qu'elles ont vérifiée dans d'autres domaines telles que l'histoire et l'architecture, et qu'elles étendent aux sciences humaines en général, vaut entièrement pour notre application expérimentale.

La réflexion déclenchée par l'expérience menée débouche sur plusieurs voies de recherche possibles. Ces recherches nécessitent le concours de compétences pluridisciplinaires, essentiellement en linguistique, en psychologie cognitive, en sémiotique et en intelligence artificielle, pour résoudre les problèmes d'ordre symbolique, ergonomique et technique qui se posent. Elles concernent en particulier les caractéristiques du langage opératif iconique qu'est l'image de synthèse utilisée en conception architecturale, la réutilisation automatique d'anciennes solutions et l'apprentissage, l'implantation à l'aide de techniques d'intelligence artificielle de la représentation et du traitement du savoir architectural.

Une première recherche possible porte sur la détermination du langage opératif verbal qui est utile à l'architecte réalisant un projet à l'aide d'un système de synthèse d'image. Une fois ce langage défini la question se pose de l'implanter en étudiant les correspondances et les décalages entre les représentations conceptuelles élaborées par les analyseurs linguistiques et les représentations conceptuelles exigées par les systèmes de synthèse d'image. Une autre recherche envisageable porte sur les procédures, désignées par des expressions verbales, et qui permettent à l'architecte de définir les objets et la composition d'une scène virtuelle et d'animer celle-ci. Il est intéressant alors de voir si des techniques de réutilisation des connaissances et d'apprentissage étudiées pour la conception architecturale sont applicables à la réutilisation et à l'apprentissage de telles procédures (suggestion de F. Guéna).

Le passage de l'image mentale à l'image géométrique puis à l'image de synthèse, soulève des problèmes de représentation et de relation entre différents types de représentation, mais aussi de raisonnement géométrique et plus généralement spatial. Par ailleurs quel rôle joue le signifié des expressions verbales dans ces passages ? Si l'on s'attache à la formalisation requise par la mise en image de synthèse, on est amené à se demander si l'utilisation de l'image de synthèse n'impose pas un enseignement sensiblement différent en ce qui concerne la représentation iconique de l'espace architectural et les outils mathématiques correspondants.

Madeleine ARNOLD & Claude LEBRUN
CIMA (Centre d'informatique et de méthodologie en architecture)
9 rue Barbanègre
75019 Paris

Remerciements. Nos premiers lecteurs furent Jean Zeitoun, directeur du CIMA, et nos collègues François Guéna, Sabine Porada et Khaldoun Zreik. Puis vinrent les évaluations des deux lecteurs, pressentis par la revue, et les commentaires de Danièle Dubois et de Claudie Faure. Nous avons ainsi été amenés à approfondir et à remanier substantiellement notre réflexion. Ce fut une tâche, rude certes, mais extrêmement enrichissante. Que tous nos lecteurs soient assurés de notre vive reconnaissance pour leur attention soutenue et leur perspicacité.

Bibliographie

- ARNOLD M. (1990) Transcription automatique verbal-image et vice versa - Contribution à une revue de la question. *EuropIA-90*, Liège, Paris, Hermès, pp. 30-37.
- ARNOLD M. (1989) La dimension spatiale : Contribution à une approche comparative de la sémiotique et de l'intelligence artificielle. *Semiotica*, 77-1/3, special issue : semiotics, cognition, and artificial intelligence, pp. 317-38.
- BERTIN J. (1970) La graphique. *Communications*, 15, pp. 169-85.
- BINOT J.L., FALZON P., PEREZ R., PEROCHE B., SHEEHY N., ROUAULT J., WILSON M. (1990) Architecture of a multimodal dialogue interface for knowledge-based systems. *ESPRIT technical week*.
- BORILLO M. (1991) Sémantique de l'espace et raisonnement spatial. *Sciences de la cognition*, Paris, 28-31 janvier 1991. Ministère de la recherche et de la technologie, Grands colloques de prospective, pp. 125-28.
- CLAYSSSEN D., LOBSTEIN D., ZEITOUN J. (1987) *Les nouvelles images. Introduction à l'image informatique*. Paris, Dunod.
- CORDIER F. (1989) Process representation and typicality.- *Cahiers de psychologie cognitive. European bulletin of cognitive psychology*, vol.9, n° 3, pp. 323-36.
- CUNY X., HOC J.-M. (1974) Les intermédiaires graphiques dans le travail : principes de caractérisation des codes.-*Le travail humain*, t. 37, n° 2, pp. 213-28.
- DELEPINE O., DONIKIAN S., HEGRON G., JEZEQUELLOU D., SOTINEL F. (1991) Une méthode déclarative de description de scènes architecturales. Polistina, A. (a cura di), *La citta' interattiva*, pp. 307-21.
- DENHIÈRE G., BAUDET S. (1989) Lecture et compréhension de textes : aspects cognitifs. *Questions de logopédie*.
- DENHIÈRE G., DENIS M. (1989) The processing of texts describing spatial configurations. *Knowledge acquisition from text and pictures*, éd. H. Mandl, J.R. Levin, Elsevier-North-Holland, pp. 249-61.
- DENIS, M. (1979) *Les images mentales*. Paris, Presses Universitaires des France.
- DENIS M. (1989) *Image et cognition*. Paris, Presses Universitaires des France.

- DENIS M., DUBOIS D. (1988) Knowledge organization and instantiation of general terms in sentence comprehension. *Journal of experimental psychology : Learning, memory, and cognition*, vol.14, n° 4, pp. 604-611.
- DESCLÉS J.-P. (1985) Représentation des connaissances. *Actes sémiotiques. Documents*, VII, n° 69-70, 51 p.
- DUBOIS D. dir. (1991) *Sémantique et cognition. Catégories, prototypes, typicalité*. Paris, Editions du CNRS.
- DUBOIS D. (1991) Les catégories sémantiques "naturelles" : prototype et typicalité. *Sémantique et cognition*, sous la dir. de D. Dubois, pp. 31-54.
- EUROPIA (association). (1991) Sur la modélisation des processus de conception créative. *Actes de 01 Design - 90*. Paris, EUROPIA.
- FALZON P. (1989) *Ergonomie cognitive du dialogue*. Grenoble, Presses universitaires de Grenoble, 176 p.
- FALZON P., BISSERET A., BONNARDEL N., DARSESES F., DETIENNE F., VISSER W. (INRIA). (1990) Les activités de conception : l'approche de l'ergonomie cognitive. Communication présentée au colloque *Recherches sur le design*, Compiègne, 17-19 oct. 1990.
- FALZON P., DARSESES F. (1991) La conception de réseaux informatiques : une approche psychologique. EUROPIA, *Sur la modélisation des processus de conception créative*, pp. 49-59.
- FAURE C. (1991) Représentations verbales et imagées. Support de cours. Paris, Ecole Nationale Supérieure des Télécommunications.
- FLEURY D., DUBOIS D. (1988). Sécurité routière. Représentations mentales de scènes urbaines. *Les Annales de la recherche urbaine*, n° 40, pp. 97-104.
- FLOCH J.-M. (1990) *Sémiotique, marketing et communication. Sous les signes, les stratégies*. Paris, Presses Universitaires de France.
- FONTANILLE J. (1989) *Les espaces subjectifs. Introduction à la sémiotique de l'observateur*. Paris, Hachette.
- GREIMAS A.J., COURTES J. (1979) *Sémiotique. Dictionnaire raisonné de la théorie du langage*. Volume 1. Paris, Hachette.
- GUENA F. (1991) *Réutilisation de solutions génériques pour résoudre des problèmes de conception*. Thèse de l'Université Paris 6.
- GUENA F., ZREIK Kh. (1988) Un système de conception architecturale assistée par architecte. *EuropIA-88*. Paris, Hermès, pp. 133-50.
- HANROT S., QUINTRAND P., CHOURAQUI E., DUGERDIL P., FRANCOIS P. (1988) Un système de CAO intelligent en architecture. *EuropIA 88*. Paris, Hermès, pp. 387-95.
- HERSKOVITS A. (1986) *Language and spatial cognition*. Cambridge, Cambridge University press.

- HOC J.-M. (1987) *Psychologie cognitive de la planification*. Grenoble, Presses Universitaires de Grenoble.
- JACKENDOFF R. (1987) On beyond zebra : The relation of linguistic and visual information. *Cognition*, vol. 26, n° 2, pp. 89-114.
- JEZEQUELLOU D. (1989) Vers une scénographie architecturale assistée par ordinateur. Approche déclarative pour la création de scènes 3D. Mémoire de fin d'études, Ecole d'architecture de Bretagne.
- LAGRANGE M.-S., RENAUD M. (1987) De l'analyse logiciste aux systèmes experts, et : Superikon, essai de cumul de six expertises en iconographie. *Systèmes experts et sciences humaines*, par J.C. Gardin et al., Paris, Eyrolles, pp. 27-55, 191-229.
- LAPRE L., HUDSON P. (1988) Talking about design : supporting the design process with different goals. *CAAD futures '87*. Elsevier, pp. 127-136.
- LARKIN J.H., SIMON H.A. (1987) Why a diagram is (sometimes) worth ten thousand words. *Cognitive science*, 11 : 65-99.
- LEBAHAR, J.-C. (1983) *Le dessin d'architecte. Simulation graphique et réduction d'incertitude*. Roquevaire (France), Ed. Parenthèses.
- LEBAHAR, J.-C. (1987) Quelques aspects significatifs de la compétence graphique des techniciens sur machines-outils à commande numérique en situation de formation. Weill-Fassina, A., Rabardel, P., éd., *Le dessin technique*, pp. 367-76.
- MAZET C. (1991) Fonctionnalité dans l'organisation catégorielle. *Sémantique et cognition*, D. Dubois, dir., pp. 89-100.
- OSTA I. (1987), L'outil informatique et l'enseignement de l'espace. Weill-Fassina, A., Rabardel, P., éd., *Le dessin technique*, pp. 325-45.
- POLISTINA A., éd. (1991) La citta' interattiva... *Atti del 3 convegno internazionale. Associazione Europea Gestione e Rappresentazione delle Trasformazioni Urbane*, Milano.
- PORADA S. (1990) Modèles et songes. Communication au *1er congrès de l' AISIM* (Association internationale de sémiologie de l'image), Blois, 1-3 nov. 1990.
- PORADA S. (1991) Architecture de l'écran : Esthétique et langage de l'image de synthèse. Polistina, A. (éd.), *La citta' interattiva...*, pp. 279-94.
- QUINTRAND P., AUTRAN, J., FLORENZANO, M., FREGIER M., ZOLLER J. (1985) *La conception assistée par ordinateur en architecture*. Paris, Hermès, 257 p.
- RABARDEL P., WEILL-FASSINA A. (1987) Nouvelles technologies et dessins techniques. Enjeux pour la recherche et la formation. *Actes du colloque de dessin technique au musée*, pp. 39-42.
- RASTIER F. (1991) *Sémantique et recherches cognitives*, Paris, Presses Universitaires de France.
- SCHANK R.C., LEAKE D.B. (1989) Creativity and learning in a case-based explainer. *Artificial intelligence*, 40, pp. 353-85.

- TVERSKY B., HEMENWAY K. (1983) Categories of environmental scenes. *Cognitive psychology*, 15, pp. 121-49.
- VANDELOISE C. (1986) *L'espace en français. Sémantique des prépositions spatiales*. Paris, Le Seuil, 245 p.
- VERILLON P., RABARDEL P. (1987) L'intériorisation du système projectif du dessin. Weill-Fassina, A., Rabardel, P., éd., *Le dessin technique...*, pp. 109-17.
- VISSER W. (1990a). Acquisition de connaissances : L'approche de la psychologie cognitive illustrée par le recueil d'expertise en conception. Texte présenté à la *Journée Acquisition de Connaissances*, Lannion, 27 avril 1990. Le Chesnay, INRIA (Institut National de recherche en informatique et en automatique).
- VISSER W. (1990b). Evocation and elaboration of solutions : different types of problem-solving actions. An empirical study on the design of an aerospace artifact. *Proceedings of Cognitiva 90*, vol. 2, Paris, AFCET, 1990, pp. 689-96.
- WEILL-FASSINA A. (1979) Présentation spatiale des données de travail et traitement des informations : points de vue et hypothèses. *Psychologie française*, vol. 24, n° 3/4, pp. 205-27.
- WEILL-FASSINA A. (1982) Représentation de données spatiales symbolisées : la lecture des intermédiaires graphiques en situation de travail et d'apprentissage professionnels. *Psychologie française*, 27, 3/4, pp. 215-27.
- WEILL-FASSINA A., RABARDEL P., éd. (1987) *Le dessin technique. Apprentissage, utilisations, évolutions*. Actes du colloque de la RCP 722/CNRS, 19-21 nov. 1986. Paris, Hermès, 397 p.
- WEILL-FASSINA A., RABARDEL P. (1989) Les graphismes comme instruments fonctionnels de communication et de pensée dans le travail scientifique et technique. Communication aux *Premières journées de psychologie du travail, ergonomie et psychopathologie du travail*, PIRTTEM-CNRS, Paris, 13-14 juin 1989.
- WEILL-FASSINA A., VERMERSCH P., ZOUGGARI G. (1987) L'évolution des compétences dans la lecture de formes élémentaires présentées en vues orthogonales. Weill-Fassina, A., Rabardel, P., ed., *Le dessin technique*, pp. 101-108.
- ZEITOUN J. Simulation visuelle et conception spatiale en architecture. Les apports de l'image de synthèse en conception architecturale. *MICAD 88*.
- ZREIK K. (1991) Sur la créativité assistée par ordinateur. EUROPIA, *Sur la modélisation des processus de conception créative*, pp. 15-27.

Annexe 1. Exemple de procédure linguistique

La procédure **rphra** remplit le moule avec les résultats de l'analyse linguistique d'une phrase. Elle est appliquée à chaque mot de la phrase, la variable **quoi** représentant le mot.

Le moule est considéré comme une suite de chaînes de caractères dénommée **phrase2**.

```
rphra quoi
  local w,quoui,rela
  quoui=quoi," "
  w=%"defini.var ".,quoi
  si w
    phrase2=phrase2,quoui," [ "
    phrase2=phrase2,informer quoui,"definition:"
    phrase2=phrase2," ] "
    ret
  fsi
  rela=informer "relation ",quoui
  si rela~""
    quoui="rel",rela
    w=%"defini.var ".,quoui
    si w
      phrase2=phrase2,quoui," re "
      phrase2=phrase2,rela{0}," "
      phrase2=phrase2,informer quoui," forme:"
      phrase2=phrase2," ] "
    ret
  fsi
  fsi
!"remplissage non fait" ;!
ret
/end rphra

end
```

Annexe 2. Exemple de procédure graphique

La procédure *macborde* commence par appeler et exécuter la procédure *posercontre*.

Posercontre place une forme *ff* contre le bord d'une forme *f*.

Puis *macborde* duplique *ff* tout le long du bord de *f*.

```

macborde f ff s ss sss p r rr rrr
local g v d a b c e fff is iss
posercontre f ff s ss sss p r rr rrr
s=is ; ss=iss
!s ; !ss
g = f
v = (som.for ff,rr)-(som.for ff,r)
v = v/(module v)
dd= dist.som f,s,ss
a=som.for f,s
!"quelle distance entre objets au bord?\n"
:dist
!"d";?d
:dupli
b=(som.for g,ss);!"b=";!b
c = dist.var a,b;!"c=";!c
c=c+dd+d;!"c=";!c
e=dist.som for ff,r,rr;!"e=";!e
si c>e;ret;fsi
dup.for g
fff=max
dep.for fff
poi=(d+dd)*v
fin
aff.for g,fff
si g_f
!"Voulez-vous changer la distance?\n";?q
    si q_1;sup.for fff; fff=fff-1
    ecr;aff.for [1,fff];goto dist
    fsi
fsi
g=fff

```

```
goto dupli  
end
```