

Application de la notion de cadre aux énoncés de positionnement et de référence

Daniel Schang

► **To cite this version:**

Daniel Schang. Application de la notion de cadre aux énoncés de positionnement et de référence. [Rapport de recherche] RR-2529, INRIA. 1995. inria-00074150

HAL Id: inria-00074150

<https://hal.inria.fr/inria-00074150>

Submitted on 24 May 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

*Application de la notion de cadre aux énoncés
de positionnement et de référence*

D. Schang

N 2529

Avril 1995

PROGRAMME 3



*Rapport
de recherche*



Application de la notion de cadre aux énoncés de positionnement et de référence

D. Schang

Programme 3 — Intelligence artificielle, systèmes cognitifs
et interaction homme-machine

Projet Dialogue

Rapport de recherche n° 2529 — Avril 1995 — 91 pages

Unité de recherche INRIA Lorraine
Technopôle de Nancy-Brabois, Campus scientifique,
615 rue de Jardin Botanique, BP 101, 54600 VILLERS LÈS NANCY (France)
Téléphone : (33) 83 59 30 30 – Télécopie : (33) 83 27 83 19
Antenne de Metz, technopôle de Metz 2000, 4 rue Marconi, 55070 METZ
Téléphone : (33) 87 20 35 00 – Télécopie : (33) 87 76 39 77

Résumé : Ce rapport présente un ensemble de recherches dont l'objectif est de mettre en oeuvre des systèmes de dialogue homme-machine possédant un comportement aussi naturel que possible. Il s'inscrit dans le cadre du traitement automatique des informations linguistiques concernant l'espace, en vue de définir et d'implémenter sur machine des modèles formels simulant le raisonnement spatial exprimé en français.

La première partie de ce rapport démontre la difficulté qu'il y a de modéliser l'espace en langage naturel, les outils comme la géométrie ou la logique pure demeurant insuffisants.

Dans la seconde partie, nous démontrons la nécessité d'adjoindre à ces modèles la notion de cadre, notion qui semble suffisamment générale pour rendre compte de bien des phénomènes liés à la modélisation de l'espace en langage naturel.

La troisième partie s'intéresse au problème de la référence en général puis à la référence spatiale. L'analyse de plusieurs systèmes de traitement de la référence spatiale en IA nous conduit à proposer notre propre modèle qui s'applique à un système graphique de multifenêtrage.

La quatrième partie pose le problème concomitant de la référence : celui du positionnement. Nous proposons un modèle cognitif des énoncés de positionnement que nous validons sur un corpus de dialogue qui a été recueilli dans notre équipe.

Au terme de cette étude, nous esquissons le modèle de CAO qui constitue la première partie de notre maquette informatique visant à la validation de la notion de cadre pour les énoncés de référence et de positionnement.

Mots-clé : dialogue homme-machine, langage naturel, référence, modélisation spatiale.

(Abstract: pto)

Mon adresse e-mail : *schang@loria.fr*, enfin je tiens à adresser ici mes plus vifs remerciements à L. Romary sans qui ce rapport n'aurait jamais vu le jour.

Application of the reference context notion to the treatment of positioning and referring expressions

Abstract: This report presents a research which may be situated in the overall problem of designing man-machine dialogue systems with a behaviour which we want to be as natural as possible. More precisely, our work aims at dealing with linguistic expressions referring to space in order to define and implement formal models which lies at the root of any reference to space.

The first section of this report shows the difficulties inherent to the modelization of space in natural language, since such tools as geometry or logic do not seem to answer the question completely.

Within the second section, we present some arguments supporting the notion of reference context, which appears to be general enough to take into account a great deal among the phenomena bound to the space modeling in natural language.

The third section tackles the problem of reference in a more general way to see how spatial reference hinges with results obtained sofar. An overview of several artificial systems dedicated to the treatment of space leads us to several proposition which we exemplify in the domain of the access to a windowing environment.

We address in the fourth section a specific issue which is closely related to that of spatial reference: the positioning of objects in a graphical environmment. We propose a cognitive representation of such utterances which we validate over a set of examples taken from a Wizard of Oz Corpus made within our team.

Fianlly, we give the main elements of a CAO model which constitutes the first step towards a full implementation of reference contexts for referring and positioning utterances.

Key-words: Man-machine dialogue, natural language, reference, spatial modeling.

Table des matières

1	Introduction	8
2	De la difficulté de modéliser l'espace en langage naturel	10
2.1	Introduction	10
2.2	Vers une solution?	12
2.2.1	Limites de l'approche géométrique	12
2.2.2	Limites d'une approche logique pure	12
2.2.3	Vers une approche fonctionnelle	13
2.3	Conclusion	15
3	La notion de cadre	16
3.1	De la nécessité des cadres	16
3.1.1	Pour l'espace	16
3.1.2	De la nécessité des cadres en général	19
3.2	Définition et caractérisation de la notion de cadre	24
3.2.1	Définition générale	24
3.2.2	Définition appliquée à l'espace	25
3.2.3	Caractérisation spatiale des cadres	25
3.3	Évolution discursive des cadres	27
3.4	Fonctionnement discursif des cadres sur un exemple	28
3.5	Etude de l'existant	30
3.5.1	Points communs avec l'approche de Talmy	30
3.5.2	Points communs avec l'approche de Vandeloise	30
3.5.3	Points communs avec l'approche de Langacker	31
3.5.4	Points communs avec l'approche de Romary	32
3.5.5	Points communs avec l'approche de Pribbenow	32
3.5.6	Points communs avec l'approche de Asher et Sablayrolles	34

3.6	Validation et apports des cadres aux autres approches	40
3.7	Conclusion	46
4	Le problème de la référence	47
4.1	Le problème de la référence en général	47
4.2	La référence dans un système graphique	47
4.2.1	A la recherche d'un modèle	49
4.2.2	La référence spatiale	50
4.3	Une étude détaillée de la préposition "à gauche de"	50
4.3.1	Le modèle de Mukerjee	51
4.3.2	L'approche de Vandeloise	53
4.3.3	Les nuages de Schirra	53
4.3.4	Le modèle de Wazinski	55
4.3.5	Le modèle de Briffault	56
4.3.6	Vers une autre approche	56
4.3.7	Quand les cadres volent à notre secours	58
4.4	Conclusion	60
5	Le problème du positionnement	61
5.1	Motivations de l'étude	61
5.2	Vers un positionnement plus "cognitif", le cas de <i>ici</i>	63
5.3	Un modèle général pour les énoncés de positionnement	65
5.3.1	Le modèle à cinq niveaux reposant sur la notion de cadre	66
5.3.2	Fonctionnement sur un exemple extrait du corpus de dialogue	70
5.4	Conclusion	73
6	La maquette informatique	76
6.1	Définition du système à réaliser	76
6.2	Notre système	76
6.2.1	"Une base solide" : le modèle de CAO	76
6.2.2	Caractéristiques détaillées du modèle	77
6.2.3	Adaptation au système de multifenêtrage	79
6.3	Conclusion	80
7	Conclusion	82

Table des figures

2.1	Un “dans” plus insaisissable qu’il n’y paraît.	11
3.1	De la nécessité des cadres.	17
3.2	Deux feuilles accrochées sur un tableau.	17
3.3	Limitations de la portée de la préposition <i>devant</i>	19
3.4	“le N”.	20
3.5	“De l’intérêt de l’axiologie <i>fenêtre</i> \rightarrow <i>non fenêtre</i> ”.	22
3.6	“ce N”.	22
3.7	“un N”.	23
3.8	Différents points de vue sur une île.	26
3.9	Deux directions pertinentes: la latéralité et la verticalité.	27
3.10	<i>Exemple discursif</i>	29
3.11	L’importance des <i>cadres</i>	31
3.12	Construction du <i>search domain</i> en considérant les objets environnants.	32
3.13	Changements de classe des verbes de mouvement.	35
3.14	Les 7 zones de localisation d’après Asher et Sablayrolles 1994.	36
3.15	<i>À l’opposé de</i>	41
3.16	Pertinence des cadres.	42
3.17	Deux bateaux pris isolément.	42
3.18	Relation entre les cadres et les prépositions.	43
3.19	Les <i>cadres de référence</i>	43
4.1	Avantages d’une optique différentielle.	51
4.2	Un modèle trop grossier.	51
4.3	Acceptabilité des prépositions selon Briffault.	52
4.4	Différentes acceptabilités de <i>à gauche de</i>	53
4.5	Le nuage de probabilité de <i>à gauche de</i>	54
4.6	Imprécision des centres de gravité.	54

4.7	L'approche de Wazinski, ses limites.	56
4.8	Référence absolue et référence dans les coins.	57
4.9	Notre modèle.	58
4.10	De la nécessité des cadres.	59
5.1	"le salon".	62
5.2	Modèle abstrait pour le positionnement.	66
5.3	"près de - loin de."	68
5.4	Orientations et zones interdites.	69
5.5	<i>Mets le fauteuil en face de la cheminée.</i>	70
5.6	Cadre initial et cadre final pour <i>mets le canapé en face de la cheminée.</i>	71
6.1	Les 4 modes graphiques possibles.	78

Chapitre 1

Introduction

Notre travail se situe en continuité du projet ESPRIT II Multiworks [Gaiffe 90] qui vise à la réalisation d'une plateforme de dialogue multimodale capable de piloter un environnement hypertexte (l'environnement *MultiCardTM* développé par Bull) à l'aide du langage naturel en autorisant des opérations de création, modification ou de destruction d'objets à l'écran.

Le rôle d'une telle interface est de faciliter au maximum l'utilisation de nouveaux logiciels à des utilisateurs novices et de permettre à des utilisateurs plus expérimentés des gains de temps appréciables en évitant des manipulations qui deviennent parfois interminables.

Une séquence d'ordres possibles pouvant être du type :

- a) *Ouvre le noeud Mailbox.*
- b) *Mets un bouton nommé Print à droite du bouton Ok.*
- c) *Ouvre le fichier Essai.*
- d) *Mets le dans le script du bouton.*
- e) *Ferme le noeud.*

Ces quelques requêtes, outre le fait qu'elles nécessitent un bon traitement des références définies, indéfinies et démonstratives [Gaiffe 92a], nécessitent également un module de référence et de positionnement spatiaux qui soient performants.

Pour aborder le problème, nous verrons dans un premier temps la difficulté qu'il y a à modéliser l'espace en langage naturel, nous verrons en particulier que les modèles d'inspiration purement géométriques ou purement logiques doivent laisser la

place à des modèles à forte composante fonctionnelle.

Dans le troisième chapitre, nous démontrerons l'utilité d'adjoindre à ces derniers modèles la notion de *cadre*, notion qui semble suffisamment générale pour rendre compte de bien des phénomènes liés à la modélisation de l'espace en langage naturel.

Dans le chapitre suivant, nous nous intéresserons au problème de la référence spatiale dans un système de dialogue homme-machine à forte composante graphique. À la recherche d'un modèle existant qui puisse satisfaire nos requêtes, nous verrons qu'aucun ne semble pouvoir convenir car ces derniers sont tantôt trop précis, tantôt trop grossiers.

Au terme de ce chapitre, nous proposerons notre modèle, modèle qui gagnera à être affiné par la notion de cadre.

Le chapitre 5 sera consacré aux énoncés de positionnement. Après un bref état de l'art, nous proposerons notre modèle à cinq niveaux d'affinages successifs qui est une instanciation de la notion de cadre et validerons ce modèle sur un extrait du corpus qui a été recueilli dans notre équipe.

Finalement, dans la dernière partie de ce rapport, nous détaillerons la plateforme spatiale de type CAO qui nous servira par la suite à valider la notion de cadre pour les énoncés de référence et de positionnement.

Tout au long de ce rapport, nous avons été préoccupés par le point suivant : plutôt que de définir un langage qui serait essentiellement subordonné à ce que l'on estime pouvoir traiter automatiquement, adoptons la démarche inverse : observons la langue telle qu'elle est pour ensuite chercher à obtenir des systèmes informatibles.

Chapitre 2

De la difficulté de modéliser l'espace en langage naturel

2.1 Introduction

Si le traitement automatique du langage naturel pose depuis longtemps de nombreux problèmes, on peut penser qu'en se restreignant uniquement à l'espace, empruntant des outils comme la géométrie ou la logique, on pourrait aboutir alors à une modélisation spatiale relativement aisée. Comme nous allons le démontrer au travers des deux exemples suivants, il n'en est rien, bien au contraire¹ :

- Imaginons que vous discutiez tranquillement à la table d'un café à Nancy avec de nouveaux collègues, l'un d'eux vous demande : *Où habites-tu?*, vous lui répondez par exemple : *À Nancy* et pas *Sur terre!?*

Si cette discussion avait eu lieu à Tahiti avec une belle tahitienne, votre réponse aurait pu être du type : *En France* mais pas : *Au 32, rue du Mouzon!?*

Ces deux exemples démontrent qu'il convient d'apporter suffisamment d'information à son interlocuteur (1er exemple) mais pas trop d'information non plus (2ème exemple); dans ces deux exemples, la réponse dépend pour une bonne part de l'endroit d'énonciation.

Notons au passage que ces deux exemples peuvent s'inscrire dans une optique de raisonnement à profondeur variable tel qu'il est prôné par [Kayser 88] qui lui même

1. Nous rejoignons les conclusions de S. Nirenburg et V. Raskin [Nirenburg 87] où les auteurs démontrent que les phénomènes de représentations et raisonnements spatiaux sont complexes, à facettes multiples et parfois relativement flous.

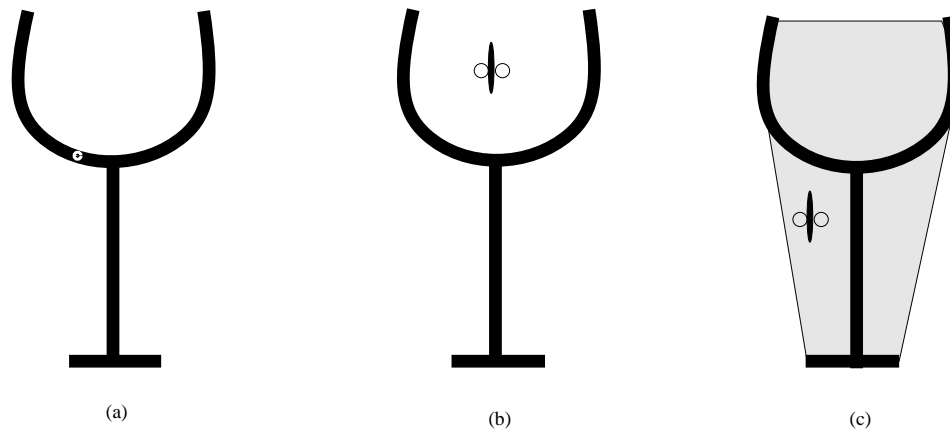


FIG. 2.1 - Un "dans" plus insaisissable qu'il n'y paraît.

peut s'inscrire dans un domaine plus large qui est celui des maximes de [Grice 75]:

Maxime de quantité:

"L'intervention doit apporter suffisamment d'information. Elle ne doit pas apporter plus d'information que ce qui est nécessaire."

- Voyons à présent un exemple qui peut sembler simple de prime abord, si j'énonce:

- (1) *Il y a une bulle dans le verre.*
- (2) *Il y a une mouche dans le verre.*

Sans équivoque, mon interlocuteur comprendra respectivement chacune des illustrations des figures 2.1.a et 2.1.b. On aurait alors deux types de *dans* à formaliser:

- pour l'exemple (1) il suffit d'utiliser "A est *dans* B" si $A^* \subseteq B^*$, (A^* et B^* représentant le volume d'espace effectivement occupé par les objets en question).
- pour l'exemple (2) on peut penser s'en sortir en proposant "A est *dans* B" si A est dans l'enveloppe convexe contenant B mais que dire alors pour la figure 2.1.c?

Ces exemples démontrent la richesse du langage naturel, langage qui ne communique que le strict minimum, toute l'information étant pour ainsi dire encapsulée dans le mot, dans la phrase et surtout dans le contexte du dialogue.

2.2 Vers une solution ?

2.2.1 Limites de l'approche géométrique

Comme mentionné précédemment, la géométrie semble être un bon outil (car très ancienne et très bien formalisée) pour opérer une modélisation de l'espace mais là encore, le langage naturel nous réserve quelques surprises.

Les exemples démontrant l'insuffisance d'une modélisation géométrique sont nombreux, voir en particulier [Vandeloise 86] p12-14. Ainsi pour la préposition *près de*, dans des phrases du type "A est *près de* B", se cantonner à la seule distance séparant A de B est délicat, si je dis : *l'épicerie est près de l'église*, je parle de distances très différentes si j'envisage le chemin à pied ou en voiture. De la même manière deux galaxies séparées par des années lumières peuvent être jugées proches, alors que ma fourchette me paraît loin si elle n'est pas à portée de ma main.

Finalement, si l'utilité de l'approche géométrique est incontournable, elle ne semble pas fournir une description assez fine des mécanismes mis en oeuvre dans les énoncés traitant de l'espace en langage naturel.

2.2.2 Limites d'une approche logique pure

La logique est de plus en plus utilisée dans les systèmes d'I.A, cependant nous allons voir que si elle s'impose comme une composante essentielle à la réalisation d'un système de modélisation spatiale, elle n'en demeure pas moins insuffisante.

On peut structurer l'espace sémantique des prépositions spatiales à l'aide de la propriété de conversité : une préposition et son opposé sont *converses* si elles vérifient la propriété suivante : A est prép B \iff B est (*opposé* prép) A.

Si les prépositions *devant/derrière* semblent être converses dans les phrases (3.a) et (3.b) il n'en est rien pour les phrases (4.a) et (4.b) :

- (3.a) *Le chêne est devant la maison.*
- (3.b) *La maison est derrière le chêne.*
- (4.a) *La pierre est devant la maison.*

(4.b) * *La maison est derrière la pierre.*

L'expérience démontre que cette asymétrie sujet/objet est très difficile à prendre en compte par des outils purement logiques [Vandeloise 86].

De façon similaire, on attend souvent d'une modélisation formelle qu'elle exprime des propriétés simples telles que la transitivité. Cette propriété est néanmoins bien souvent trop forte pour le langage naturel :

de : *Le livre est sur la table.* SUR(L,T).

et : *La table est sur le sol.* SUR(T,S).

on ne peut en déduire par transitivité que : *le livre est sur le sol.* SUR(L,S).

De la même manière, pour la préposition *dans* qui semble transitive par essence : si "A est dans B" et "B est dans C" alors "A est dans C" ne l'est pas toujours, ainsi de l'exemple : *l'abeille est dans la rose et la rose est dans le vase* on ne peut pas en déduire que *l'abeille est dans le vase*.

L. Vieu et M. Aurnague ne partagent pas ce point de vue ([Vieu 91]) et proposent une approche basée sur une logique non monotone (nous y reviendrons au paragraphe suivant).

Toutefois, ces deux auteurs s'accordent à dire qu'en préalable à une formalisation trop brutale de l'espace, il est nécessaire de préciser le fonctionnement précis des opérateurs spatiaux en langage naturel. Une analyse plus fonctionnelle de l'espace semble être en mesure d'apporter quelques éléments de réponse.

2.2.3 Vers une approche fonctionnelle

L'approche fonctionnelle s'appuie sur les connaissances extra-linguistiques de l'espace que partagent les locuteurs d'une même langue, elle donne donc une importance accrue aux facteurs non spatiaux déterminés par exemple par :

- le contexte du dialogue,
- les circonstances dans lesquelles les objets sont localisés ou utilisés,
- les fonctionnalités des objets.

Vandeloise, s'appuyant sur la notion de ressemblance de famille² définit les prépositions *sur/sous* à l'aide de la combinaison des traits caractéristiques suivants :

1. Si A est *sur/sous* B alors A est généralement plus haut/plus bas que B, (le "généralement" est mis pour des exemples du type : *l'affiche est sur le mur* ou pour : *la tache est sur la plafond.*)
2. Si A est *sur* B, il y a généralement un contact (indirect) entre A et B.
3. Si A est *sous* B, A est généralement rendu inaccessible à la perception par B.
4. Dans les relations A est *sur/sous* B, A est généralement plus petit que B.
5. Si A est *sur* B, l'action de B s'oppose à l'action de la pesanteur sur A.

Pour l'exemple précédent de *la mouche est dans le verre*, l'approche fonctionnelle distingue trois types de *dans* [Aurnague 90] :

- le dans partiel : *les fleurs sont dans le vase*,
- le dans total : *le livre est dans la bibliothèque*,
- le dans partie-de : *l'homme est dans la foule*.

Actuellement, parmi les différents types d'approches à forte composante fonctionnelle s'intéressant à la modélisation de l'espace en langage naturel, on peut distinguer les approches *linguistiques*, *logico-computationnelles* et *computationnelles*.

Dans le premier groupe, l'approche de [Vandeloise 86]³ occupe une place de choix. L'auteur se base sur le principe extra-linguistique suivant : "*un objet dont la position est incertaine ne peut être localisé sans référence à une entité dont la position est mieux connue*". Il aboutit alors à la notion de *cible/site*⁴, le *site* ayant une place plus stable et mieux connue que celle de la *cible* et étant plus facilement perceptible que la *cible*...

La formalisation de Vandeloise s'applique donc à des phrases du type "*cible relation-spatiale site*" en mettant l'accent sur le côté fonctionnel des prépositions spatiales.

2. **Ressemblance de famille**: concept développé initialement par [Wittgenstein53] représenté par la combinaison des traits qui le caractérisent sachant que chaque trait n'est pas toujours nécessaire ni suffisant, il s'agit d'une structuration qui permet aux membres d'une catégorie d'être reliés les uns aux autres sans avoir une propriété en commun qui définisse la catégorie. Ce qui les réunit, ce sont des similarités, des ressemblances qui s'entre-croisent, se recouvrent partiellement.

3. On pourrait également citer l'approche de [Herskovitz 86] qui, si elle est sensiblement différente, aboutit néanmoins à des résultats similaires.

4. **Trajector/Landmark** dans la grammaire cognitive de [Langacker 91].
Figure/Ground pour [Talmy 83]

Dans le second groupe, celui des approches *logico-computationnelles*, l'une des plus récente approche est celle de [Aurnague 93a], [Aurnague 93b], [Aurnague 93a] qui prône une modélisation de l'espace à 3 niveaux: le niveau *géométrique* (représentant l'espace objectif décrit par le texte analysé), le niveau *fonctionnel* (prenant en compte les relations entre les objets, en particulier celles qui ne sont pas géométriques) et le niveau *pragmatique* (en s'appuyant pour l'essentiel sur des informations extérieures au texte lui même: le contexte, les intentions du locuteur, etc. Ce niveau modifie les résultats du niveau fonctionnel de manière à adapter sa sémantique à la situation réelle). Les auteurs ont implémenté un système qui autorise certaines inférences spatiales (phrase 5) et en interdit d'autres (phrase 6):

(5) *Paul est dans la maison* \wedge *La maison est dans l'île* \implies *Paul est dans l'île*.

(6) *Paul est dans l'île* \wedge *L'île est dans la mer* \implies *Paul est dans la mer*, (Faux).

Enfin dans le dernier groupe, celui des systèmes utilisant des formalisations à *composante computationnelle*, nous aimerions mettre l'accent sur le projet VITRA⁵ et plus particulièrement sur le système SOCCER [Wahlster 87]⁶. L'un des points novateurs de cette approche étant la modélisation des prépositions spatiales à l'aide de "nuages de probabilité" avec la possibilité de combiner plusieurs nuages entre eux. Nous verrons ce modèle au §4.3.3.

2.3 Conclusion

Actuellement, l'approche fonctionnelle semble être la plus prometteuse [Vandeloise 86], [Vieu 91], même si certains auteurs, au travers de la notion de ressemblance de famille, lui reprochent son côté peu formel ce qui la rend délicate à implémenter, néanmoins de plus en plus de modèles actuels mettant en oeuvre ce type d'approche voient le jour : ([Cuyckens 93], [Hottenroth 93]).

5. **VI**sual **TR**Anslator.

6. Signalons cependant que SOCCER n'est pas uniquement un système computationnel, de manière générale on le rangerait dans les systèmes linguistiques et/ou cognitifs.

Chapitre 3

La notion de cadre

3.1 De la nécessité des cadres

3.1.1 Pour l'espace

Prenons l'exemple de la phrase *Mets un bouton nommé Print à droite du bouton Ok* qui est tiré de l'introduction.

Si l'on peut trouver la position verticale du bouton à l'aide d'informations prototypiques¹ (en utilisant le prototype de *à droite de*) et si la taille du bouton peut être déduite de celui existant, en revanche, trois positions horizontales pour le bouton sont au moins possibles (voir figure 3.1).

Parmi ces trois possibilités, il faut qu'il ne reste plus que la première. La position 2 peut être rejetée d'emblée pour des raisons fonctionnelles basées sur la tâche (i.e. un bouton doit être placé dans une fenêtre et pas directement sur l'écran).

La position 3 semble contradictoire avec la signification de la phrase qui établit une relation directe entre les 2 boutons dans le même contexte, c'est précisément ce type de contraintes qui sera exprimé à l'aide de la notion de cadre.

Vandeloise, en se basant sur la notion de *cible/site* à laquelle il adjoint celle de ressemblance de famille nous propose les caractérisations suivantes de la préposition *au dessus de* :

- caractéristique A1 : A est *au dessus de* B si A est (dans la direction verticale) *au dessus de* B.

1. Pour une étude détaillée du prototype voir [Rosch 78], [Kleiber 90], [Dubois 91].

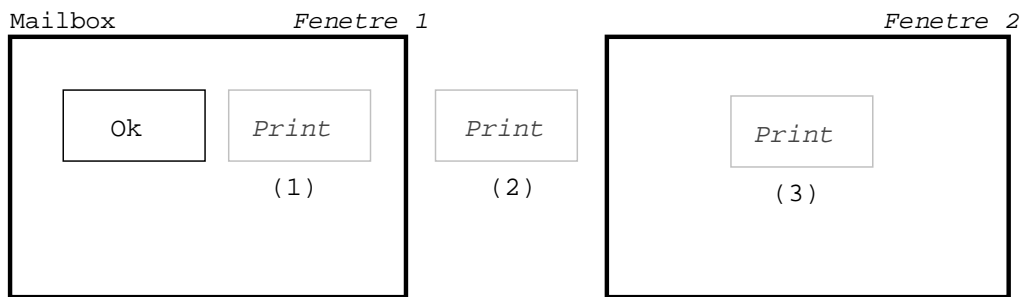


FIG. 3.1 - *De la nécessité des cadres.*

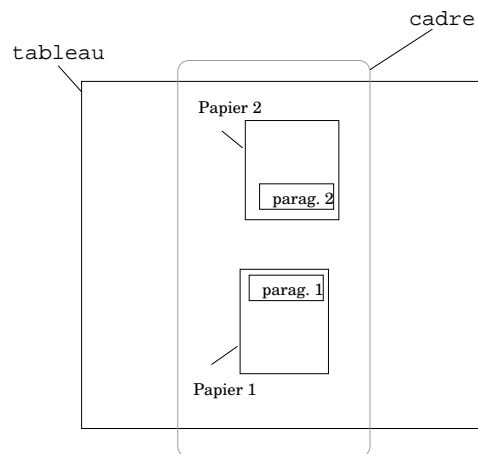


FIG. 3.2 - *Deux feuilles accrochées sur un tableau.*

- caractéristique A2: si A est *au dessus de* B, leur intersection verticale est généralement vide.

Pour illustrer ceci, imaginons deux feuilles accrochées à un tableau², voir figure 3.2. Le fait de dire que *la feuille 2 se trouve au dessus de la feuille 1* ne pose pas de problème. En revanche dire que *le paragraphe 2 se trouve au dessus du paragraphe*

2. Merci à B. Gaiffe pour cet exemple.

l semble tout à fait curieux. Cette dernière phrase nécessite de pouvoir placer les deux paragraphes dans un même contexte : c'est ce contexte que nous nommerons par la suite *cadre*.

Ceci démontre qu'une entité (comme par exemple le paragraphe 1) peut jouer le rôle de site aussi longtemps que les entités sont situées à une certaine distance de ce dernier (nous rejoignons [Pribbenow 93] sur ce point). Le pouvoir localisateur du site est donc nécessairement limité à une certaine zone. Ainsi, le paragraphe 1 aurait pu être utilisé pour situer n'importe qu'elle entité qui se serait trouvée sur la même feuille mais pas au delà. Concernant les deux feuilles de papier, la préposition *au dessus de* peut être utilisée car on peut les placer dans un même cadre, celui du tableau.

De la même manière, Vandeloise définit les prépositions *devant/derrière* comme suit :

- A est *devant/derrière* B si A se trouve du côté positif/négatif déterminé par l'orientation générale³.

Cette définition n'est pas assez restrictive car elle ne limite pas assez le pouvoir localisateur du site. Prenons l'exemple de la phrase (1) où l'interlocuteur regardera devant la voiture mais nécessairement à proximité de cette dernière.

(1) *Oh! regarde, on dirait qu'il y a un billet de 500 Francs là-bas devant la Jaguar.*

Plus généralement la portée d'une préposition dépend de l'objet sur lequel elle s'appuie, ainsi la portée de *devant* dans *devant l'église* couvre une zone plus vaste que dans *devant la voiture* par exemple. Ce fait est indirectement pris en compte dans le système Soccer [Schirra 93a] par les nuages de probabilités qui s'appuient sur le bord des objets dépendant de leur forme et de leur étendue [Schirra 93b].

Le site a donc une portée limitée que ce soit à *droite*, à *gauche*, *devant*, *derrière*, *etc.*, c'est ce type de contrainte que nous souhaiterions traduire/exprimer par la notion de cadre. Notons enfin que cette étendue peut être limitée d'avantage, en particulier par des entités contiguës. Prenons l'exemple d'un joueur qui se trouve sur un terrain de football avec une balle située près de lui, dans le but (voir figure 3.3). On préférera (2.a) à (2.b) sans nul doute :

3. **Orientation générale :** ressemblance de famille dont les principaux traits sont : *la direction frontale, la direction du mouvement, la direction du regard, la direction des autres organes de perception : le goût, l'ouïe, etc.*

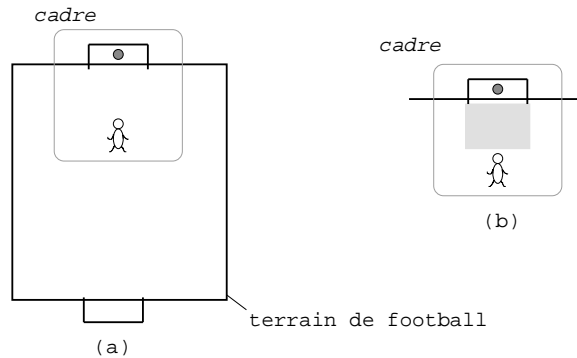


FIG. 3.3 - Limitations de la portée de la préposition devant

(2.a) *La balle est dans le but.*

(2.b) *La balle est devant le joueur.*

La portée de *devant* est limitée par la discontinuité spatio-fonctionnelle du bord du terrain de football, limitation qui est renforcée par l'introduction d'un autre site (le but) qui limite la portée localisatrice du joueur de football.

3.1.2 De la nécessité des cadres en général

Notre équipe mène depuis quelques années une réflexion sur la référence, réflexion qui a abouti à définir le modèle exposé ci-après dont l'évolution peut être retracée dans [Romary 89], [Romary 91], [Gaiffe 91], [Gaiffe 92a], [Gaiffe 92b] et [Gaiffe 93].

Ces travaux se basent sur la notion d'axiologie⁴ qui est un point de vue référentiel adopté par un discours pour exprimer une opposition élémentaire ([Gaiffe 92a]). Gaiffe fait la distinction entre deux types de propriétés: les propriétés nominales (notées N) et les propriétés particulières (notées P). Les propriétés nominales regroupent tout ce qui, du point de vue du sens, est directement exprimé dans le nom commun. Quant aux propriétés particulières, notées P, ce sont elles qui spécialisent un référent en dessous de propriétés N. Ce sont ces propriétés qui permettent d'iso-

4. La notion d'axiologie peut être rapprochée de la sémantique différentielle chère à Rastier ([Rastier 87] et [Rastier 91]).

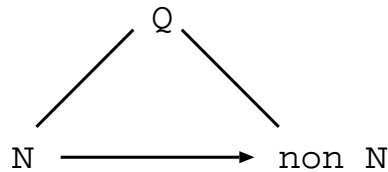


FIG. 3.4 - "le N".

ler un élément de la classe de tous ceux qui sont N. Voyons à présent les schémas référentiels associés aux différents types de groupes nominaux non génériques⁵.

Les groupes nominaux en "le N"

Les groupes nominaux de la forme "le N" sont destinés à sélectionner dans un ensemble d'objets celui qui porte le nom N par opposition aux objets du même ensemble qui ne méritent pas ce nom.

Le fait qu'il est souhaitable qu'il y ait un N et des non N pour les groupes nominaux définis permet d'expliquer des exemples tels que (3.a) et (3.b) :

(3.a) *Déplace les fenêtres vertes.*

(3.b) *Mets la fenêtre la plus à gauche en arrière plan.*

Ainsi, dans cet exemple, les deux ensembles candidats pour justifier le GN défini sont respectivement : l'ensemble de tous les objets visibles sur l'écran, ce qui est l'interprétation non anaphorique, et l'ensemble des fenêtres vertes, ce qui correspond à l'interprétation anaphorique. La présence du "N" "fenêtre" dans le second énoncé serait inutile s'il s'agissait d'une anaphore : dans l'ensemble des fenêtres vertes, tous les éléments sont des fenêtres. En revanche, cette catégorie "fenêtre" se justifie pleinement dans le cas de l'interprétation non anaphorique : parmi les objets présents à l'écran, certains sont des fenêtres, d'autres non.

Dans le cas de l'exemple (4.a) et (4.b) :

5. Comme le démontre B. Gaiffe, les groupes nominaux génériques n'apparaissent que très peu dans les dialogues de commande finalisés.

- (4.a) *Déplace les fenêtres vertes et les icônes bleues.*
- (4.b) *Mets la fenêtre la plus à gauche en arrière plan.*

l'ensemble correspondant à l'interprétation anaphorique devient un aussi bon candidat que l'ensemble correspondant à l'interprétation non anaphorique ; l'ambiguïté est donc possible. Il faudrait alors d'autres critères pour faire le choix.

Finalement, l'auteur propose le schéma de la figure 3.4 pour lequel on a bien un ensemble (noté Q) et une axiologie (N vs non N). Sur cette axiologie, on se situe au point noté : " N ". Ce schéma permet de retarder le choix entre la référence directe et l'anaphore. Dans le premier cas, l'ensemble Q du schéma sera instancié par l'ensemble des objets de l'écran, dans le second, il s'agira d'un ensemble issu des énoncés précédents comme dans :

- (5.a) *Déplace la fenêtre verte et l'icône bleue.*
- (5.b) *Mets la fenêtre en arrière plan.*

où l'ensemble Q est instancié par {fenêtre-verte, icône-bleue}.

Un problème ne semble pas être pris en compte par ce modèle, ainsi si je dis "détruis la fenêtre", il paraît naturel de rechercher une entité qui ait le type fenêtre par rapport à d'autres entités qui n'aient pas ce type (ce qui est bien pris en compte par le schéma de la figure 3.5.a).

Par contre de "détruis la fenêtre verte" il semble curieux de conserver l'axiologie *Fenêtre* \rightarrow *non Fenêtre* (voir à la figure 3.5.b) en imaginant que seules une fenêtre verte et une fenêtre rouge se trouvent affichées sur l'écran (voir figure 3.5.c). En effet, il n'y a pas de *non fenêtres* sur l'écran.

Le fait de parler de *non fenêtres* est cependant légitime car on pourrait aussi bien parler de l'écran, du fond de l'écran, etc. Pour corroborer ce fait, imaginons un dialogue où l'utilisateur n'aurait que la possibilité de manipuler des fenêtres au travers de phrases du type : *Détruis la fenêtre bleue. Mets une fenêtre verte ici, etc.* Ainsi, si il mentionne le nom "fenêtre", l'utilisateur apporterait trop d'information à la machine en ne respectant pas la maxime de Grice de quantité qui conduirait plutôt l'utilisateur à utiliser des requêtes du type : *Détruis la bleue, mets une verte ici, etc.*

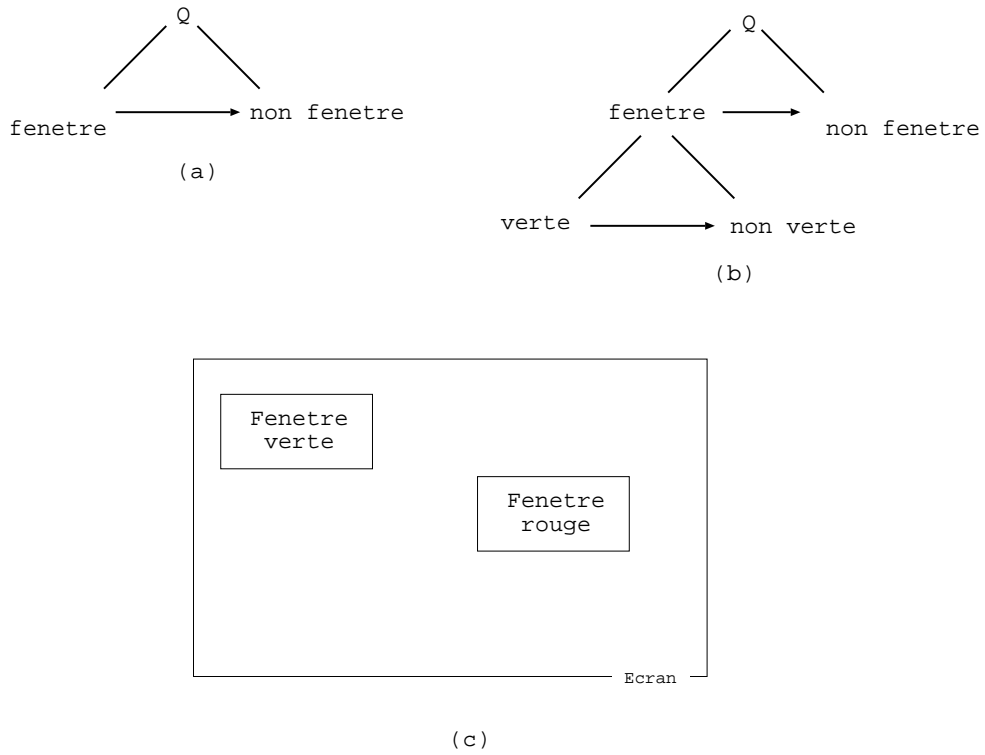


FIG. 3.5 - "De l'intérêt de l'axiologie fenêtre \rightarrow non fenêtre".

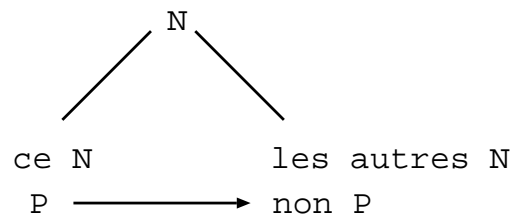


FIG. 3.6 - "ce N".

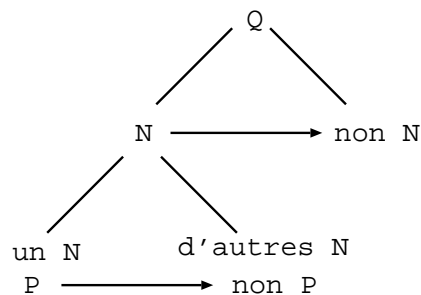


FIG. 3.7 - "un N".

Les groupes nominaux en "ce N"

"Ce N" permet de re-catégoriser un référent ([Gaiffe 92a] p. 49-50). Le référent doit donc apparaître en dessous de la catégorie N. De ce fait, lorsqu'on dit "ce N", on oppose un individu particulier, celui dont on parle aux autres qui sont également N.

Cet individu particulier, pour pouvoir être opposé aux autres N doit avoir des propriétés qui lui soient propres. On a donc le schéma de la figure 3.6. Les propriétés permettant d'individualiser le référent pourront être le fait d'être désigné du doigt ou d'avoir été singularisé dans le discours. Autrement dit, "ce N" ne pourra reprendre qu'un référent saillant.

Pour "ce N" comme pour "le N", les schémas référentiels proposés unifient donc le cas de l'anaphore et celui de la référence directe.

Les groupes nominaux en "un N"

Dans le cas des groupes nominaux indéfinis, l'auteur envisage deux axes d'opposition: "un chat" fait partie de la classe des chats et peut donc s'opposer à ce qui n'est pas chat, c'est également un individu susceptible de se distinguer des autres chats. Ces deux axes d'opposition possibles nous amènent au schéma de la figure 3.7. La différence avec le groupe démonstratif tient au fait que l'indéfini permet d'introduire un nouveau référent dans le discours, l'axe P vs non P qui individualise le référent dans la classe est donc posé a priori dans le cas de l'indéfini. L'auteur justifie l'axe d'opposition N vs non N à l'aide de l'exemple:

(6) *J'ai acheté une Toyota parce qu'elles sont robustes et pas chères.*

Il semble clair qu'à partir de cet exemple, un fort sous-entendu porte sur le caractère moins robuste d'autres voitures que les Toyotas. C'est bien ce qui justifie de disposer à partir de l'antécédent "une Toyota" de la possibilité d'opposer la classe des Toyotas à d'autres classes. Cette opposition est précisément l'axe N vs non N de notre schéma de référence.

Notons qu'on ne peut rassembler dans Q que des éléments qui sont relativement homogènes car une axiologie n'a de sens que dans un certain ensemble qui permette de regrouper les objets que l'on oppose, ainsi, si deux objets ne peuvent pas être rassemblés, les opposer n'a en définitive pas grand sens, il suffit de prendre pour exemple : "Hier j'ai vu Paul et le marteau"?!? Ainsi, si Q peu rassembler différentes catégories il possède nécessairement une certaine consistance.

Dans le paragraphe suivant, nous donnerons une définition de la notion de cadre qui semble pouvoir unifier les structures d'interprétation requises pour le traitement des expressions définies, indéfinies et démonstratives.

3.2 Définition et caractérisation de la notion de cadre

3.2.1 Définition générale

Nous proposons la définition suivante des cadres :

Déf. Cadre : Ensemble d'entités (nécessaires et suffisantes) vues sous un angle particulier permettant l'interprétation des expressions référentielles par différentiation.

Quelques remarques :

La notion de différentiation provient directement de la notion d'axiologie. Nous avons pu voir au paragraphe précédent comment ce mécanisme opérait, nous verrons qu'il permet de prendre finement en compte le contexte.

Le fait de devoir disposer d'entités nécessaires et suffisantes sera illustré pour les expressions référentielles spatiales au paragraphe suivant.

La notion d'angle particulier est très importante car elle permet de prendre en compte l'utilisateur et les visions qu'il a du monde qui l'entoure, en particulier les

perceptions qu'il a de ce monde. Jackendoff illustre très bien ce point lorsqu'il dit : "les choses ne sont rien d'autre que ce que nous en percevons".

3.2.2 Définition appliquée à l'espace

Déf. Cadre spatial : entité qui restreint l'univers à une portion d'espace ou à la connexion de portions d'espaces nécessaires et suffisantes vue(s) sous un angle particulier permettant l'interprétation des expressions référentielles spatiales par différenciation.

Cette définition est une instanciation de la définition précédente appliquée à l'espace.

Le fait d'avoir une connexion de portions d'espace est nécessaire pour des exemples du type *Où est Jean-Marie ?* Une réponse du type : *Il n'est pas dans son bureau* conduira à imaginer l'ensemble des lieux où peut se trouver Jean-Marie sous la forme d'une réunion de lieux possibles comme *le couloir, les autres bureaux du bâtiment, etc.* par opposition à des lieux comme Paris, New-York, etc. Il s'agit en fait de l'ensemble des alternatives possibles où peut se trouver Jean-Marie. Cette notion d'alternative est très importante comme on le verra par la suite.

La notion d'angle particulier permet de prendre en compte les trois orientations que l'on peut rencontrer pour orienter un ou plusieurs objets (l'orientation déictique, l'orientation extrinsèque et l'orientation intrinsèque⁶).

3.2.3 Caractérisation spatiale des cadres

Voici quelques caractéristiques des cadres :

- Le cadre est un concept dynamique, il évolue au fil du discours ; par essence même car c'est la portion d'espace minimale (ou réunions de portions d'espace) qui permettra le calcul des référents spatiaux (on verra au §3.3 l'ensemble de ces évolutions possibles).
- Les frontières du cadre sont relativement floues, même si, bien souvent, elles se précisent au fil du discours. L'étendue du cadre dépend fonctionnellement des objets ou parties d'objets qu'il englobe comme on a pu le voir au §3.1.1.

6. Orientation déictique : orientation basée sur le locuteur.

Orientation extrinsèque : orientation à partir des traits marquants de l'univers comme la gravité.

Orientation intrinsèque : orientation à partir d'un objet.

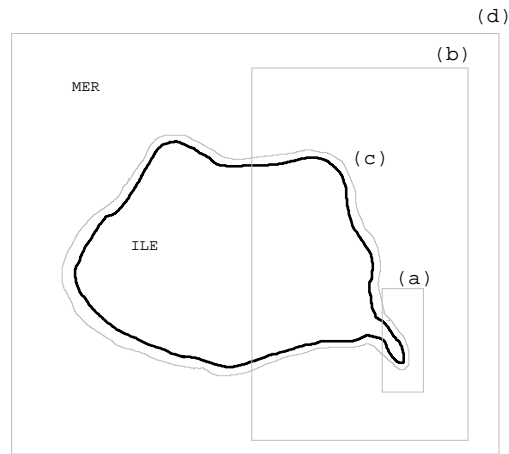
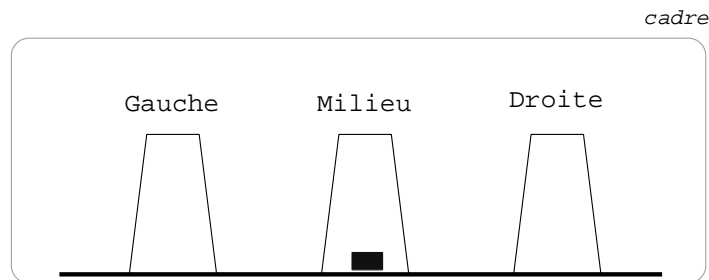


FIG. 3.8 - *Différents points de vue sur une île.*

Si nous savons que l'étendue du cadre ne doit pas être trop importante, il faut cependant qu'elle soit suffisante, sans quoi, les entités qui s'y trouvent ne prennent pas pleinement leur signification. Par analogie, [Langacker 91], démontre qu'un objet n'a pas de signification, s'il est pris isolément mais qu'il faut nécessairement le placer dans un contexte suffisamment large, voir en particulier la figure 3.8 où on ne peut vraiment commencer à parler d'île qu'à partir du "point de vue" 3.8.d.

Ce point justifie donc la notion de "portions d'espaces nécessaires" que l'on trouve dans la définition des cadres spatiaux.

- A la notion de cadre, nous adjoignons celle de *focus*. Le *focus* pouvant être défini comme la zone (à l'intérieur du cadre) où se trouve l'entité (respectivement les entités) dont on parle, par opposition aux zones où l'on n'est sûr qu'elle(s) ne s'y trouve(nt) pas. Comme on le verra, le *focus* peut se déplacer à l'intérieur du cadre au fil du discours, pouvant alors s'intéresser à un élément particulier du cadre, à une zone particulière du cadre (le fond, le bord, etc.)
- Notons enfin que dans un cadre donné on a pas nécessairement toutes les orientations (verticalité, latéralité, perspective). En fait seules les orientations pertinentes sont présentes : dans certains cas une seule direction est pertinente. Ainsi dans l'exemple des trois gobelets (voir figure 3.9) sous lequel se trouve



Le jeton se trouve sous le gobelet du milieu.

FIG. 3.9 - Deux directions pertinentes: la latéralité et la verticalité.

cachée une pièce dont il faut deviner l'emplacement, les seules directions pertinentes sont la verticalité et la latéralité (la perspective est absente ici).

Ces caractérisations des cadres font ressortir le principe sous-jacent qui les gouvernent : la notion de principe d'économie cognitive, ainsi, à un instant donné du dialogue, tout ce qui se trouve en dehors du cadre n'est pas pris en compte, de même que dans le cadre, toutes les directions ne sont pas prises en compte non plus d'où la notion de "portions d'espace suffisantes" de la définition.

3.3 Évolution discursive des cadres

Tout comme le focus de Grosz-Sidner [Sidner 86], les cadres n'évoluent pas de manière aléatoire, et peuvent permettre de restreindre le calcul de la référence en se restreignant à une partie de l'espace mais le parallèle s'arrête là, le modèle de Grosz-Sidner ne pouvant s'appliquer de manière générale à la modélisation de l'espace (pour une étude détaillée voir [GaiFFE 92a]).

Voici les caractérisations des quatre évolutions possibles des cadres au cours du discours :

- conservation : $\text{cadre}' = \text{cadre}$ ⁷ C'est le cas par défaut. Ce cas répond à un principe d'économie cognitive, l'utilisateur ne va pas "sauter du coq à l'âne" et va raisonner de manière privilégiée dans le cadre courant.

⁷. *cadre* désigne l'ancien cadre, *cadre'* désigne le nouveau cadre qu'il faut calculer.

- rétrécissement : le nouveau site appartient à l’ancien cadre et $\text{cadre}' \subset \text{cadre}$.
- agrandissement : $\text{cadre}' \cap \text{cadre} \neq \emptyset \wedge \text{cadre}' \not\subseteq \text{cadre}$. C’est le cas lorsqu’une entité du cadre courant sert à localiser une autre entité (resp. est localisée par rapport à une autre entité) qui n’appartient pas au cadre courant.
- fermeture : $\text{cadre}' \cap \text{cadre} = \emptyset$. C’est le cas lorsqu’on localise une entité par rapport à une autre, les deux entités n’appartenant pas au cadre courant.

Nous sommes conscients de rester à un niveau de description très élevé des cadres, ce choix est délibéré, nous en verrons une instanciation pour les énoncés de positionnement au chapitre 5.

3.4 Fonctionnement discursif des cadres sur un exemple

Analysons l’exemple suivant en ne nous préoccupant que de l’évolution des cadres spatiaux en laissant de côté le calcul des références non spatiales :

- (1) *Tandis que le petit Pierre joue devant la maison,*
- (2) *sa soeur Lisa joue derrière : cela leur évite de se bagarrer.*
- (3) *Pendant ce temps, leur maman repasse le linge dans la cuisine.*
- (4) *Richard, le mari, tapisse les autres pièces de la maison.*
- (5) *Amusant! à Paris, au même instant,*
- (6) *René le frère de Richard retapisse également le salon.*

1. Le cadre est construit à partir de la “maison”, son étendue est de l’ordre de grandeur de la maison. Le *focus* matérialise le “devant de la maison” (figure 3.10.a).
2. *Conservation* : le cadre reste le même, il n’y a que le *focus* qui change (figure 3.10.b).
3. *Rétrécissement* : le cadre devient plus précis, il englobe à présent la “cuisine”. Le *focus* se réduit lui à la cuisine (figure 3.10.c).
4. *Agrandissement* : le cadre s’étend à présent à toutes les pièces de la maison. Notons que les pièces sont représentées sous forme d’un ensemble à la figure 3.10.d car leur configuration spatiale ne nous intéresse pas ici, le fait important étant uniquement que le mari se trouve dans l’une de ces pièces au moment où l’on parle.

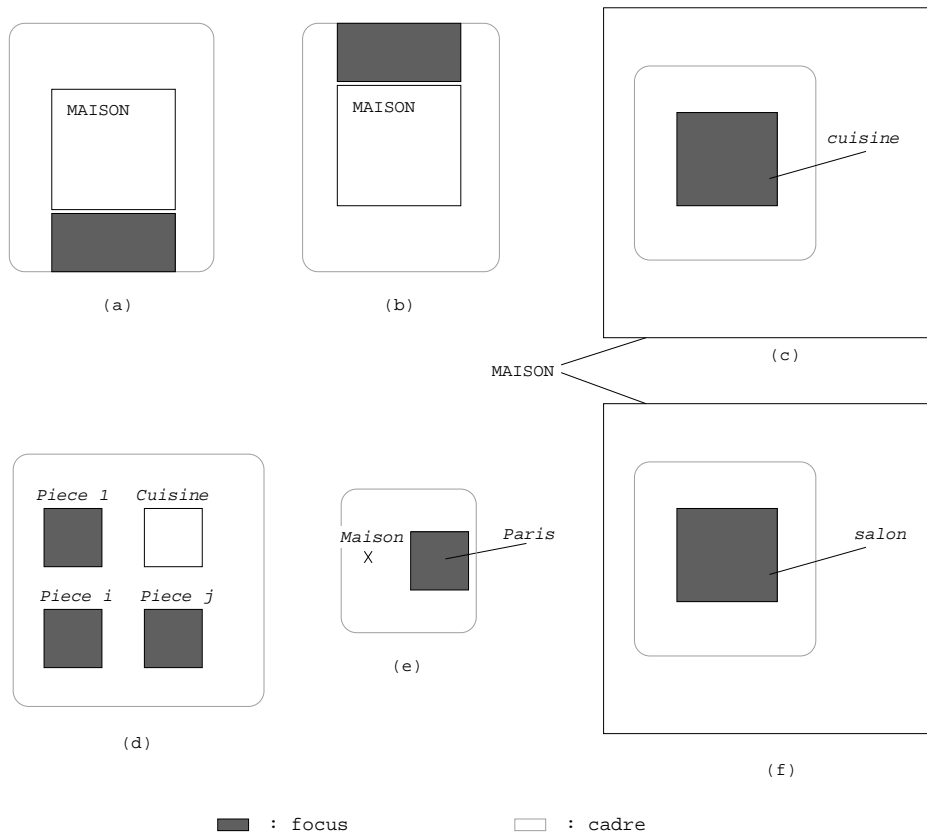


FIG. 3.10 - Exemple discursif.

5. *Fermeture* : Le cadre change complètement, il englobe à présent “Paris” (figure 3.10.e). La maison reste cependant incluse dans le cadre sous forme ponctuelle, il y a en quelque sorte opposition de lieu entre la maison et Paris.
6. *Rétrécissement* : Le cadre se focalise sur le “salon” d’une maison dont on a pas plus d’information pour l’instant (figure 3.10.f), le cadre se précisera sûrement dans la suite du discours.

3.5 Etude de l’existant

Placer une origine absolue aux “cadres” n’est pas aisé car cette notion apparaît déjà en filigrane dans bien des travaux sur l’espace avant de naître concrètement mais sous une forme trop générale ce qui la rend difficilement applicable dans un système informatique.

3.5.1 Points communs avec l’approche de Talmy

[Talmy 83] énonce : “*De façon générale, le choix d’une expression locative comporte d’abord un choix d’échelle et de perspective*”.

A présent, imaginons que vous dites à quelqu’un *Pierre est dans l’église*, en étant vous même et votre interlocuteur dans cette même église. Dans ce cas, la perspective selon laquelle est conçue la scène ne peut pas se réduire au cadre de la figure 3.11.a, il faut au contraire que se trouvent inclus dans le cadre l’église et une portion autour de cette dernière afin de supporter cette opposition potentielle : être dedans ou dehors de l’église (voir figure 3.11.b).

Comme on le verra par la suite, la notion de cadre prend finement en compte ce choix d’échelle et de perspective, mais se veut plus générale, car dans ces deux concepts, ne se trouve pas la notion d’alternative que l’on a dans le cadre.

3.5.2 Points communs avec l’approche de Vandeloise

Dans la même optique, [Vandeloise 86] introduit une notion voisine de celle des cadres : la notion de “scène objective” qu’il ne développe malheureusement pas :

“Principe de transfert” : *Le locuteur a la faculté de se déplacer en tout point utile à la perspective selon laquelle il conçoit la scène objective qu’il décrit* (p. 38).

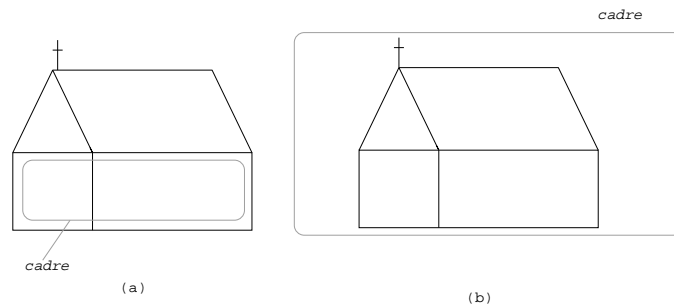


FIG. 3.11 - *L'importance des cadres.*

En première approximation, la notion de cadre regroupe la notion de perspective et de scène objective. Elle est cependant plus générale car elle peut regrouper des espaces distincts comme on a pu le voir au §3.2.2 (Cf. l'exemple : "Où est Jean-Marie? Il n'est pas dans le bureau") en se basant sur la notion d'alternative.

3.5.3 Points communs avec l'approche de Langacker

[Langacker 91], au travers des "cognitive domains", démontre l'importance des "search domains" :

"Each locative expression confines the subject to a specific "search domain" which then constitutes the scope of predication for the locative that follows. Thus, in (7), the locative upstairs confines the quilt to an upper story, and "in the bedroom" is construed relative to this restricted region. Only an upstairs bedroom need be considered."

(7) *The quilt is upstairs in the bedroom in the closet on the top shelf behind the boxes.*

Par cet exemple, Langacker, nous démontre l'importance des "search domain" mais il n'en propose pas de formalisation ; c'est justement ce que nous tenterons de faire au paragraphe suivant au travers de la notion plus générale de cadre.

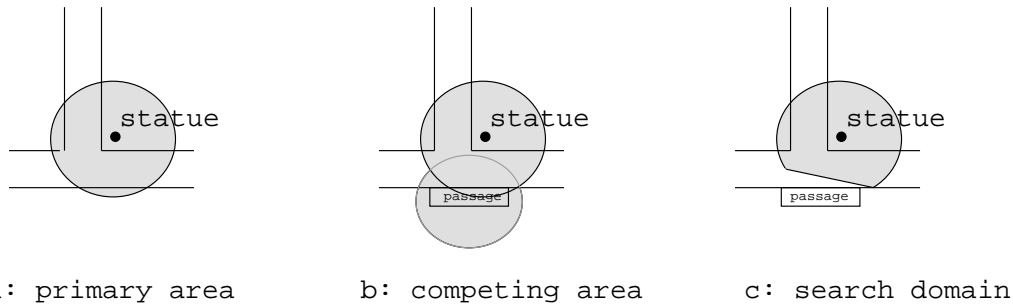


FIG. 3.12 - *Construction du search domain en considérant les objets environnants.*

3.5.4 Points communs avec l’approche de Romary

[Romary 93] démontre que ses “espaces de référence” constituent une notion indispensable pour interpréter le sens de “*ici*”, le déctique ayant pour rôle de discriminer un endroit ou une entité particulière par opposition à d’autres endroits ou entités et ce à l’intérieur d’un sous-espace particulier : “l’espace de référence” (nous verrons au §5.2 l’exposé détaillé de cette étude, au travers des énoncés de positionnement notamment).

Cette notion est celle qui se rapproche le plus de celle des cadres car on peut voir les espaces de référence de L. Romary comme l’instanciation de la notion de cadre pour le cas particulier de la préposition “*ici*”.

3.5.5 Points communs avec l’approche de Pribbenow

Pour mieux préciser la notion de cadre, nous allons la comparer à l’approche de [Pribbenow 93]. Pour rappel, l’approche de Pribbenow réalise l’étude des phrases du type LE REL RO où la LE (*entity to be located*) est localisée par rapport au RO (*Reference Object*) à l’aide d’une préposition REL⁸. Tout comme [Kaufmann 93] l’auteur propose un modèle à deux niveaux :

- 1) le niveau linguistique,
- 2) le niveau conceptuel.

8. Par “entity”, l’auteur entend la possibilité de localiser des situations et des mouvements.

Le premier niveau prend en compte les traits spécifiques du langage tandis que le second se base sur un ensemble de concepts (spatiaux) indépendants du langage. Chaque expression localisatrice est interprétée à l'aide d'une aire (le *search domain*) dans laquelle le LE doit se trouver avec la plus haute probabilité.

Le calcul des expressions de localisation se fait donc en 2 étapes :

- la première étape segmente l'espace pour obtenir la *primary area* (figure 3.12), cette segmentation opérant une différenciation entre la *remoteness area* et la *nearness area*, cette dernière étant découpée suivant l'orientation qui a été choisie parmi les 3 systèmes de référence suivants : intrinsèque, déictique ou géographique. ([Pribbenow 93] p. 454).
- la seconde étape calcule le *search domain* (figure 3.12.c) en utilisant les *priorizations*⁹. Le domaine résultant est l'intersection de la *priorization area* et des présentes aires de localisation, initialement la *primary area*. Les objets qui entrent en compétition avec le site constituent des *priorizations* négatives qui vont limiter le pouvoir localisateur du site.

On relève plusieurs similitudes entre cette approche et la nôtre :

- La règle de Pribbenow selon laquelle : "The bigger the object, the bigger the area", est prise en compte dans notre approche par le fait que, prototypiquement, le cadre englobant une table aura une extension qui sera plus petite que le cadre englobant une maison, cette extension dépendant du RO et du LO¹⁰ [Schang 94b].
- Chez Pribbenow, l'extension de la *primary area* est bien souvent limitée par les objets environnants, de notre côté, le cadre est bien souvent limité par des discontinuités spatio-fonctionnelles [Schang 94b].
- Enfin, on peut faire un parallèle certain entre la *primary area* de Pribbenow et notre niveau purement spatial dans le cas des énoncés de positionnement. De la même manière, l'affinage des *primary area* pour aboutir aux *search domain* est pris en compte dans notre approche par les connaissances sur le monde comme on le verra par la suite.

Parmi, les différences, on notera les 2 différences significatives suivantes :

- L'approche de Pribbenow est statique alors que la nôtre se veut discursive.

9. Par exemple, les localisations prototypiques de la LE et les relations prototypiques entre le RO et le LE induisent des *priorizations* positives.

10. Notre étude porte sur des LO (*Localised Objects*) mais cela ne porte pas à conséquence car, comme le démontre Pribbenow, on peut toujours transformer la LE en un ou plusieurs LO.

- Du point précédent, il découle entre autre qu’une séparation nette entre le niveau conceptuel et linguistique tel qu’il est prôné par Pribbenow n’est pas souhaitable car cette séparation est source de pertes d’indices linguistiques précieux notamment pour la prise en compte de l’évolution discursive des cadres.

3.5.6 Points communs avec l’approche de Asher et Sablayrolles

Nous terminerons ce bref état de l’art en mentionnant l’approche de [Asher 94] qui réalise une étude précise de la sémantique spatio-temporelle des verbes de changement de lieu (intransitifs) et des prépositions spatiales¹¹. Cette description distingue quatre classes de verbes de mouvement :

- *Les verbes de changement de lieu* : ces verbes, comme par exemple *partir*, *rentrer* ou *éloigner*, imposent que la cible change de localisation par rapport au site, dans (8), *Jean* se trouve à l’extérieur de la *maison* au début du procès et à l’intérieur à la fin.

(8) *Jean rentre dans la maison.*

- *Les verbes de changement d’emplacement* : cette classe comprend des verbes comme *déplacer*, *circuler*, *descendre*. Les verbes de changement d’emplacement imposent seulement que la cible change de localisation par rapport à une partie du site. Dans la phrase (9), *Jean* reste à l’intérieur de *la ville* durant tout le procès, même si sa position doit obligatoirement changer par rapport à un point particulier de la *la ville*.

(9) *Jean se déplace dans la ville.*

- *Les verbes de changement potentiel d’emplacement* : ces verbes permettent que la cible change de localisation par rapport à une partie du site mais ils ne l’imposent pas. Cette classe comprend des verbes comme *courir*, *voler*, *danser*. Par exemple, dans (10), *Jean* ne change pas du tout de localisation.

(10) *Jean court sur place.*

11. Un grand merci à N. Hathout pour ses compléments [Hathout 94].

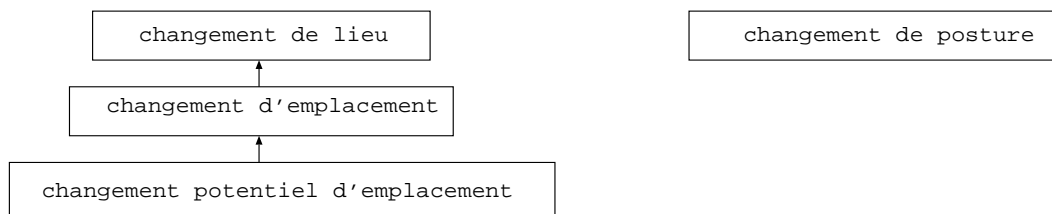


FIG. 3.13 - *Changements de classe des verbes de mouvement.*

Le fait de pouvoir se construire avec l'expression *sur place* est un caractère propre aux verbes de changement potentiel d'emplacement, comme en témoignent les exemples suivants :

- * (11.a) *Jean sort sur place.*
- * (11.b) *Jean se déplace sur place.*

– *Les verbes de changement de posture* : ces verbes comme *se pencher*, *se baisser*, *s'agenouiller*, interdisent à la cible de changer d'emplacement ou de lieu.

Changements de classe des verbes de mouvement : un syntagme verbal construit à partir d'un verbe de mouvement et d'un syntagme prépositionnel peut ne peut pas appartenir à la même classe que ce verbe.

- (12.a) *Jean descend dans la l'arène.*
- (12.b) *Jean court partout dans sa cuisine.*
- (12.c) *Jean court jusqu'à la cuisine.*

Un verbe de changement potentiel d'emplacement peut être tête d'un VP qui décrit un changement d'emplacement (12.a) ou un changement de lieu (12.b). De même, un verbe de changement d'emplacement peut donner un VP de changement de lieu (12.c). Les verbes de changement de lieu et de changement de posture ne peuvent, en revanche, être tête que d'un VP de leur classe. La figure 3.13 résume schématiquement les changements de classe des VP de mouvement.

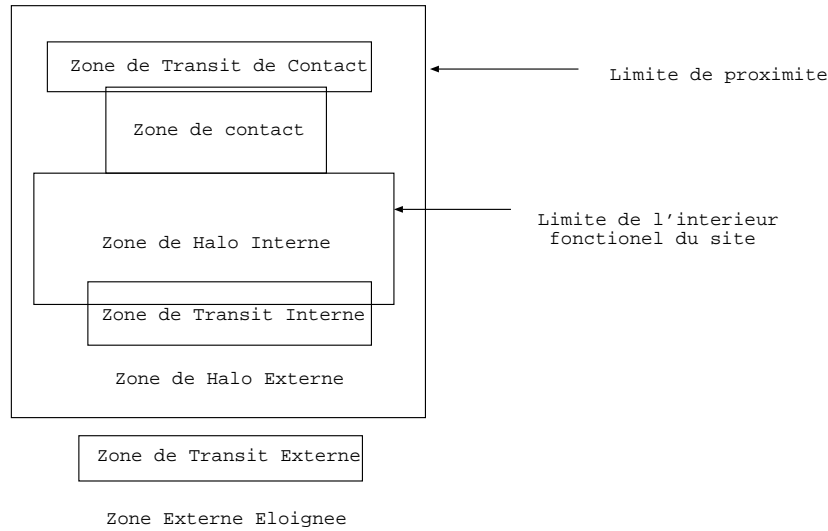


FIG. 3.14 - Les 7 zones de localisation d'après Asher et Sablayrolles 1994.

Dans le reste de ce paragraphe, je ne considère uniquement les verbes de changement de lieu car ce sont les seuls pour lesquels Asher et Sablayrolles ont proposé une définition de la sémantique spatio-temporelle.

La description de Asher et Sablayrolles repose sur un découpage de l'espace en sept zones (figure 3.14) définies relativement au site et qui sont des fonctions du site, de la cible et du contexte [Aurnague 93b]. On peut définir à l'aide de ces zones la sémantique spatio-temporelle d'un verbe comme *partir* de la manière suivante :

$$(13) \quad \textit{partir}(e) \rightarrow \\
P(\textit{source}(e), ZHI(\textit{cible}(e), Lref(e))) \wedge \\
P(\textit{chemin}(e), ZHE(\textit{cible}(e), Lref(e))) \wedge \\
P(\textit{but}(e), ZEE(\textit{cible}(e), Lref(e))) \wedge \\
\textit{Initial}(e, Lref(e))^{12}$$

Cette formule décrit que e est un événement (le départ) dont la localisation se fait par rapport à $Lref(e)$, à savoir le site considéré dans la relation spatiale introduite par le verbe. Les trois premiers termes dans la conséquence établissent que le début,

12. ZHI: Zone de Halo Interne, ZHE: Zone de Halo Externe, etc.

le milieu et la fin de e doivent se dérouler respectivement à l'intérieur du site, à proximité (à l'extérieur) du site et loin du site. Le quatrième terme conjoint indique que e est de polarité initiale¹³ par rapport au site. Dans le cas du verbe *rentrer*, l'événement est une entrée; son début, son milieu et sa fin ont lieu respectivement à proximité (à l'extérieur) du site, à la limite du site et à l'intérieur du site; sa polarité est finale par rapport au site :

$$(14) \quad \textit{rentrer}(e) \rightarrow \\ P(\textit{source}(e), ZHE(\textit{cible}(e), Lref(e))) \wedge \\ P(\textit{chemin}(e), ZTI(\textit{cible}(e), Lref(e))) \wedge \\ P(\textit{but}(e), ZHI(\textit{cible}(e), Lref(e))) \wedge \\ \textit{Final}(e, Lref(e))$$

Finalement, on obtient 10 groupes de verbes suivant les zones qu'ils mettent en jeu, ce qui permet de donner une sémantique spatio-temporelle à chacun des verbes de changement de lieu (intransitifs).

Le cadre formel utilisé pour décrire la sémantique spatio-temporelle des verbes de changement de lieu permet de décrire celle des prépositions spatiales à l'aide des mêmes notions que pour les verbes. Par exemple, (15) définit la sémantique spatio-temporelle de *pour* et (16) définit celle de la préposition *de* dans le cas où elle a le sens de *l'intérieur de*, *de dedans*.

$$(15) \quad \textit{pour}(\textit{site}, e) \rightarrow \\ P(\textit{but}(e), ZEE(\textit{cible}(e), \textit{site})) \wedge \\ \textit{Final}(e, \textit{site})$$

$$(16) \quad \textit{de}(\textit{site}, e) \rightarrow \\ P(\textit{source}(e), ZHI(\textit{cible}(e), \textit{site})) \wedge \\ \textit{Initial}(e, \textit{site})$$

13. Les auteurs s'appuient ici sur le formalisme développé par [Laur 93] qui est très similaire à celui de [Boers 87] en répartissant les verbes de changement de localisation en verbes de polarité de type initiale, médiane et finale :

- polarité initiale : s'éloigner, partir, décoller, etc.
- polarité médiane : passer, etc.
- polarité finale : s'approcher, arriver, entrer, etc.

Ces formules donnent, pour les groupes prépositionnels *de Paris* et *pour Nancy* les représentations sémantiques (17) et (18).

$$(17) \quad \exists x, Paris(x) \wedge \\ P(source(e), ZHI(cible(e), STref(x))) \wedge \\ Initial(e, STref(x))$$

$$(18) \quad \exists y, Nancy(y) \wedge \\ P(but(e), ZEE(cible(e), STref(y))) \wedge \\ Final(e, STref(y))$$

La sémantique proposée par Asher et Sablayrolles permet de construire, compositionnellement des représentations pour ce qu'ils appellent des *complexes verbaux*, à savoir des syntagmes composés d'un verbe de déplacement et d'un ou plusieurs syntagmes prépositionnels. Le calcul de ces représentations est fait à l'aide de huit règles ; ces règles ont comme principaux paramètres la nature directionnelle *vs.* positionnelle de la préposition et l'identité de la polarité éventuelle de la préposition avec celle du verbe¹⁴. Je ne présente ici que les règles qui concernent les prépositions directionnelles (20) à (23) ; les règles relatives aux propositions positionnelles mettent en oeuvre des formes de composition identiques à celles qui sont discutées ci-dessus.

$$(20) \quad directionnelles(P) \wedge polarite(P) = polarite(v) \rightarrow \\ Lref(e) = ZonePrep(cible(e), STref(N))$$

$$(21) \quad directionnelle(P) \wedge polarite(P) \neq polarite(v) \wedge polarite(v) = I \rightarrow \\ P(source(e), ZonePrep(cible(e), STref(N))) \wedge Lref(e) = ?$$

$$(22) \quad directionnelle(P) \wedge polarite(P) \neq polarite(v) \wedge polarite(v) = M \rightarrow \\ P(chemin(e), ZonePrep(cible(e), STref(N))) \wedge Lref(e) = ?$$

$$(23) \quad directionnelle(P) \wedge polarite(P) \neq polarite(v) \wedge polarite(v) = F \rightarrow \\ P(but(e), ZonePrep(cible(e), STref(N))) \wedge Lref(e) = ?$$

14. Les prépositions positionnelles n'ont pas de polarité. Dans le cas de la composition de ces prépositions avec un verbe de déplacement, seule la polarité du verbe est prise en considération pour le calcul de la représentation sémantique du VP.

Dans ces formules, $ZonePrep(cible(e), STref(N))$ est la zone définie par le groupe prépositionnel et N est le référent de discours associé au DP complément de la préposition. L'utilisation de ces règles peut être illustrée en détaillant le calcul de la représentation sémantique du VP *partir de Paris pour Nancy*. Ce calcul se déroule en deux étapes : on commence par composer la représentation de *partir* donnée en (13) avec celle de *de Paris*, donnée en (17)). La préposition *de* a dans ce cas une utilisation directionnelle et elle a la même polarité que *partir* ; on doit donc utiliser la règle (20) pour composer les représentations de ces deux constituants. On obtient ainsi l'équation :

$$\exists x, Paris(x) \wedge Lref(e) = ZHI(cible(e), STref(x))$$

qui reportée dans (13), donne :

$$(24) \quad \exists x, Paris(x) \wedge \\ P(source(e), ZHI(cible(e), ZHI(cible(e), STref(x)))) \wedge \\ P(chemin(e), ZHE(cible(e), ZHI(cible(e), STref(x)))) \wedge \\ P(but(e), ZEE(cible(e), ZHI(cible(e), STref(x)))) \wedge \\ Initial(e, ZHI(cible(e), STref(x)))$$

Notons que, dans tous les cas, la composition préserve les caractéristiques du verbe comme par exemple sa polarité ; en d'autres termes, elles est compatible avec la définition syntaxique de ce dernier comme tête du syntagme verbal. La représentation du syntagme verbal complet (26) est calculée lors de la seconde étape en composant (24) avec (18). Le complexe verbal (24) étant de polarité initial et le PP étant de polarité finale, on utilise (21) qui produit l'ajout de la condition :

$$(25) \quad \exists y, Nancy(y) \wedge P(but(e), ZEE(cible(e), STref(y)))$$

à la représentation de *partir de Paris*. On obtient donc :

$$(26) \quad \exists x, \exists y, Paris(x) \wedge Nancy(y) \wedge \\ P(source(e), ZHI(cible(e), ZHI(cible(e), STref(x)))) \wedge \\ P(chemin(e), ZHE(cible(e), ZHI(cible(e), STref(x)))) \wedge \\ P(but(e), ZEE(cible(e), ZHI(cible(e), STref(x)))) \wedge \\ P(but(e), ZEE(cible(e), STref(y))) \wedge$$

$$\text{Initial}(e, ZHI(\text{cible}(e), STref(x)))$$

Cet exemple illustre le fait que les PP spatiaux ne remplissent pas tous le même rôle (sémantique) par rapport au verbe dont ils dépendent. On a ainsi deux modes de composition sémantique des verbes et des complexes verbaux avec les syntagmes prépositionnels. Le premier, illustré par (24), est l’instanciation du lieu de référence : le PP définit le site de la relation spatiale introduite par le verbe. En d’autres termes, le PP complète la représentation du verbe lui-même ; les représentations du verbe et du PP sont dans ce cas mutuellement dépendantes : le verbe apporte la cible et le PP le site. Le second mode de composition, illustré par (26), est l’ajout à la représentation du verbe ou du complexe verbal de conditions supplémentaires sur la localisation de la cible par rapport à un site introduit par le PP. La représentation du verbe est alors indépendante de celle du PP puisque la localisation a lieu par rapport à un site propre au PP ; le PP dépend du verbe pour la cible.

Par rapport au modèle de Asher et Sablayrolles, la notion de cadre se place à un niveau de généralité supérieur¹⁵ ainsi, la notion de cadre semble pouvoir unifier les sept zones de localisations mises en oeuvre par les auteurs car le cadre décrit justement la portion d’espace minimale nécessaire à l’achèvement du raisonnement spatial courant. Soulignons cependant la qualité de cette étude qui semble être la plus précise jamais réalisée concernant les verbes de changement de lieu.

3.6 Validation et apports des cadres aux autres approches

Comme nous allons le voir, le cadre est une notion essentielle ; dire “A est *prép* B” sans cadre, c’est un peu comme parler d’une hypothénuse sans triangle rectangle. Voici donc quelques arguments dont la liste n’est pas exhaustive :

- La préposition *à l’opposé de* ne peut s’interpréter sans la notion de cadre dont nous proposons le script suivant : “à l’opposé de” réalise un mouvement du *focus* de sa position courante à la position diamétralement opposée et ce à l’intérieur du cadre (voir figure 3.15). Il en va de même de la préposition *de l’autre côté de*. À terme, nous pensons définir une image mentale abstraite pour chaque préposition spatiale, ce schéma mental abstrait est à rapprocher des archétypes cognitifs ou schèmes sémantico-cognitifs de [Abraham 91].

15. Nous entrons ici dans l’opposition modèles généraux/spécifiques, tandis que la première classe de modèles manque de précision dans son comportement ce qui nécessite une méthodologie d’instanciation, la seconde classe est difficile à étendre.

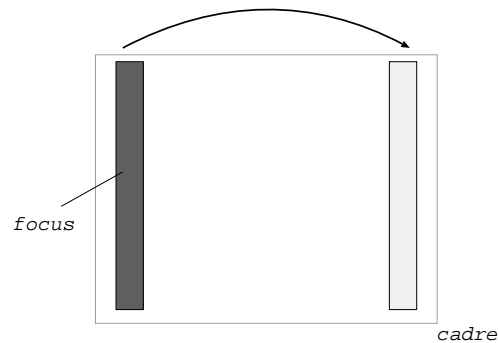


FIG. 3.15 - À l'opposé de.

- Pour pouvoir parler de *site* et de *cible*, il faut se restreindre à un sous-espace, sinon on aurait une multitude de sites et de cibles. Ce point apparaît en filigrane dans la notion de “scène objective” du principe de transfert de Vandeloise, mais l’auteur ne développe pas cette notion or cela nous paraît capital. De même comment expliquer le fait qu’une *cible* peut devenir *site* (entre 27.a et 27.b) et qu’un *site* peut devenir *cible* (entre 27.c et 27.d)?

(27.a) Devant la *maison*_{site} se trouve notre *voiture*_{cible}.

(27.b) Il y a des *taches*_{cible} d’huile sous la *voiture*_{site}.

(27.c) Il y a plein de *taches*_{cible} sur ma *voiture*_{site}.

(27.d) A mon avis ces taches proviennent du fait que tu gares toujours ta *voiture*_{cible} sous cet *arbre*_{site}.

Dans notre optique, ces passages *cible-site* et *site-cible* s’expliquent facilement à l’aide du cadre courant : le cadre ayant changé entre (3.16.a) et (3.16.b)¹⁶ ainsi qu’entre (3.16.c) et (3.16.d), le *site* (défini comme l’entité la plus saillante dans le cadre change lui aussi, ainsi que la *cible*, voir figure 3.16).

16. Si l’approche de Vandeloise ne se place pas dans une optique discursive, elle peut néanmoins prendre en compte (27.a et 27.b) : selon l’auteur un *site* présente généralement la caractéristique d’être apparu plus tôt dans le discours par rapport à la *cible*.

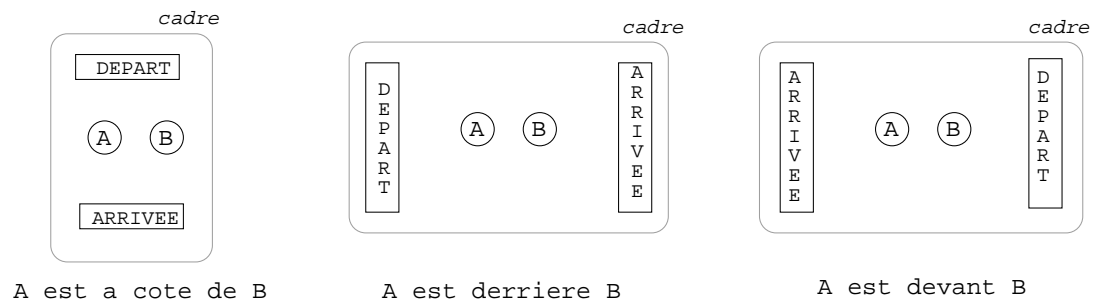


FIG. 3.18 - *Relation entre les cadres et les prépositions.*

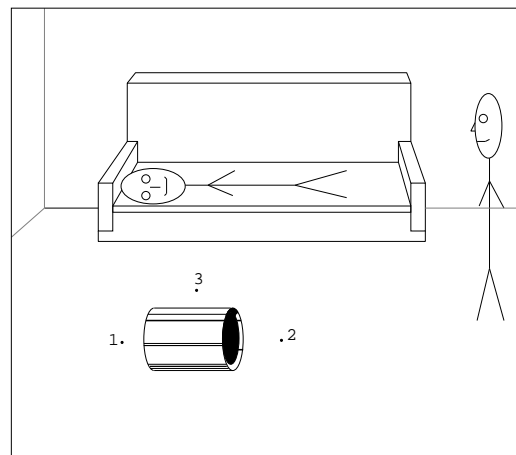


FIG. 3.19 - *Les cadres de référence.*

détails d'une scène. Le calcul d'un référent par exemple se fera en priorité dans le cadre courant. Notre approche va donc dans le sens du "raisonnement à profondeur variable" tel qu'il a été prôné par [Kayser 88]. Un parallèle est possible entre notre approche et celle de Allen, tout comme nous le faisons pour l'espace, Allen se restreint à des intervalles de temps ce qui lui permet de ne travailler que sur une partie du temps.

- [Carlson-Radvansky 93] démontre que les prépositions comme *au dessus de* doivent être utilisées et interprétées relativement à ce qu'elle dénomme les *frames of reference* (cadres de référence ci-après). On aboutit alors à trois types possibles de cadres de référence lorsqu'il s'agit de déterminer quel objet se situe *au dessus* de la poubelle (figure 3.19). Ainsi, suivant la perspective de la personne qui est couchée sur le canapé, l'objet 1 se situe au dessus de la poubelle (*viewer-centered reference frame*), suivant le cadre de référence de type *object-centered reference frame* l'objet 2 est situé au dessus de la poubelle, tandis que l'objet 3 se situe *au dessus* de la poubelle suivant un cadre de référence du type *environment-centered reference frame*.

Si l'idée des trois types de cadres de référence semble unanimement adoptée : ([Retz-Schmidt 88], [Hernández 94] pour ne citer qu'eux), nous doutons de la pertinence de l'exemple choisi par Carlson-Radvansky. Ainsi, l'usage de la préposition *au dessus de* qui est fait dans cet exemple ne semble pas très naturel, le cadre de référence ne parvenant pas à occulter le caractère déterminant de cette préposition, à savoir le critère de verticalité gravitationnelle.

Comme nous l'avons vu, la notion de cadre s'appuie sur les trois types de cadres de référence sus-cités, nous retiendrons donc la validation psychologique des "reference frames" pour la notion de cadre.

- La notion de cadre semble être en mesure d'expliquer les phénomènes d'ambiguïté qui peuvent surgir entre deux locuteurs. Ainsi, nous posons le postulat que les ambiguïtés surgissent quand les deux locuteurs ne partagent pas le même cadre.

Pour illustrer ce postulat, imaginons la conversation téléphonique suivante qu'un locuteur (appelons le Pierre) aurait avec l'un des ses amis :

(28.a) Pierre : *Sais-tu où se trouve Paul en ce moment ?*

(28.b) Ami : *Oui il est sur terre.?!?*¹⁷

Alors qu'en (28.a), le locuteur Pierre s'est construit un cadre potentiel qui fait l'union de plusieurs villes où Pierre croit que Paul se trouve, la réponse de son ami l'oblige à reconsidérer complètement ce cadre dans une optique du type *sur terre* → non *sur terre*. Ainsi la surprise de Pierre entre (28.a) et (28.b)

17. En imaginant que cet ami n'est pas fou!

découle du fait que ce dernier ne partageait pas du tout le même cadre que son interlocuteur.

- Les cadres semblent également pouvoir expliquer les phénomènes de transitivité, reprenons l'exemple donné par L. Vieu du second chapitre :

(29) *Paul est dans la maison* \wedge *La maison est dans l'île* \implies *Paul est dans l'île*.

(30) *Paul est dans l'île* \wedge *L'île est dans la mer* \implies *Paul est dans la mer*, (Faux).

De notre point de vue, il n'est pas si sûr de pouvoir déduire que *Paul est dans l'île*, ainsi *Paul est dans la maison* prend sa signification en fonction des alternatives, si Paul est soit dans la maison (il se repose) soit dans l'île (et non pas sur le continent), Paul n'est surtout pas dans l'île (sauf si on change de cadre à nouveau pour se demander si Paul est dans l'île ou le continent).

De même pour (30), lorsqu'on dit que *Paul est dans l'île* par opposition à *Paul est sur le continent*, on peut très bien imaginer qu'un typhon se serait abattu sur cette île emportant Paul au large, rendant alors envisageable le fait que Paul se trouve dans la mer!

- Terminons cette liste par un exemple qui prouvera que là encore notre notion, par sa généralité peut simplifier les modèles existants.

L'axe vertical est caractérisé par sa direction et son sens. Il est indépendant de la position du locuteur aussi longtemps que le discours n'implique qu'une portion limitée du globe terrestre. [Vandeloise 86] p. 89-91.

L'auteur illustre son propos en mentionnant le fait que la verticale française n'est pas la même que la verticale africaine, prenant pour exemple : *L'Afrique est en dessous de la France mais le sorcier est au-dessus de la montagne*¹⁸. Là encore, la notion de verticalité s'interprète tout naturellement en se restreignant tour à tour à telle ou telle portion du globe terrestre : le cadre.

18. Cet exemple ne semble pas tout à fait naturel pour la bonne et simple raison que la comparaison se fait dans deux cadres différents.

3.7 Conclusion

Ce chapitre présente et définit la notion de cadre, notion qui nous semble inévitable si l'on souhaite étendre les résultats actuels concernant les domaines de la représentation et du raisonnement spatial.

Qui plus est, cette notion intègre les précédents travaux que nous avons menés sur la référence : limitation de l'ensemble de référence et contraintes fonctionnelles entre les entités comme nous allons le voir dès à présent au travers des énoncés de référence.

Chapitre 4

Le problème de la référence

4.1 Le problème de la référence en général

L'une des facilités essentielles du langage est de permettre la désignation variée d'objets et d'entités du monde. Ceci donne cohésion et continuité aux textes et discours en évitant de répéter des informations déjà introduites. Un problème non trivial consiste alors à retrouver la ou les entités que désignent ces reprises (auxquelles elles se réfèrent), d'où la notion de référence qui est définie par [Kleiber 81] comme *la fonction qui permet aux signes linguistiques de renvoyer à la réalité extralinguistique*. Il s'agit d'un vaste problème qui a mobilisé bien des esprits en linguistique [Charolles 94], [Milner 82], [Kleiber 81], [Gaiffe 93] pour ne citer qu'eux. Côté informatique, comme l'a montré [Gaiffe 92a], le problème de la référence limite de beaucoup les performances des systèmes de dialogue actuels.

4.2 La référence dans un système graphique

Dans le domaine du dialogue homme-machine, et plus particulièrement quand il s'agit de dialogues multimodaux, le problème de la résolution des références peut être abordé de deux manières différentes. Dans les cas où les utilisateurs du système sont des professionnels prêts à accepter une période de formation ou lorsque l'application est suffisamment simple, l'approche la plus efficace est de définir un langage artificiel tel que celui que nous avons mis en oeuvre pour la commande d'une console sonar [Souvay 92]. Dans de telles situations, le langage est défini de manière à simplifier la résolution des références exprimées par un utilisateur. Dans d'autres situations cependant, il n'est pas possible de procéder de cette façon.

Nous savons en effet que le dialogue est finalisé par la tâche et que donc le langage à traiter correspond à un sous-langage de ce qu'on nomme communément le langage naturel. Malgré tout, il s'avère que quelle que soit la tâche, il n'est pas possible d'obtenir de la part d'un utilisateur une réduction de la palette des expressions référentielles dont il dispose de façon générale, à savoir les expressions définies, indéfinies, démonstratives ou pronominales. Les restrictions observées se situent beaucoup plus au niveau du nombre de concepts pertinents pour la tâche et se traduisent par une réduction du lexique et des valeurs sémantiques que peuvent prendre les entrées dans celui-ci. L'option que nous avons choisie est donc de considérer les expressions référentielles telles qu'elles sont utilisées dans le langage courant (alias le langage naturel) et parier qu'elles seront employées d'une façon similaire en situation de dialogue finalisé, plutôt que de préjuger des réactions d'un utilisateur devant un système de dialogue et limiter nos implémentations à ces simples projections (qui bien souvent s'apparentent plus à des absurdités¹).

Si l'importance et l'engouement pour le développement de systèmes de dialogue homme-machine multimodal est de plus en plus manifeste ([Carré 91], [Bourguet 92], [Poirier 93] et [Gaildrat 93] pour ne citer qu'eux), nous pensons que les systèmes actuels gagneraient en efficacité à améliorer les modèles d'interprétation de chacune de ces modalités (même indépendamment) les unes des autres [Schang 94a].

Dans cette optique, nous souhaitons étendre notre maquette MultiWorks par un module de référence spatial qui fonctionne dans un premier temps indépendamment des autres modalités tels que la désignation gestuelle.

Ayant recherché des solutions dans les systèmes actuels, et empruntant l'une des réflexions de [Romary 93], nous verrons que ces systèmes, laissant de côté les capacités de l'homme pour ne prendre en compte que les capacités de calcul de la machine aboutissent bien souvent à des modélisations tantôt trop grossières, tantôt trop précises.

Il nous est alors apparu nécessaire de reprendre la réflexion que nous avons menée dans notre équipe sur la référence, en particulier la notion d'axiologie, pour l'appliquer à la référence spatiale.

Comme nous le verrons tout au long de ce chapitre, le problème de la référence spatiale est essentiel dans un système de visualisation graphique car il concerne les trois opérations principales que l'on peut y faire :

1. C'est notamment le cas des démonstratifs qui sont presque systématiquement traités de manière coréférentielle alors que la reprise anaphorique existe :

Déplace la fenêtre verte.

Mets cette fenêtre en arrière plan.

- la création d'objets (ex : *crée une icône à droite de la fenêtre rouge*, etc.),
- la modification d'objets (ce qui correspond dans certains cas à une navigation dans le système, ex : *mets la fenêtre verte à droite*).
- la destruction d'objets (ex : *détruis l'icône de droite*).

Rejoignant les réflexions de [Klein 93], nous verrons que le problème de la référence est un processus cognitif complexe, car il met en jeu l'interaction d'une multitude de facteurs, ainsi, toujours selon cet auteur, le traitement d'une requête aussi simple que l'énoncé (1) requiert l'étude de bien des principes syntaxiques, sémantiques et pragmatiques.

(1) *Ce livre doit être remis sur l'étagère à gauche de la porte.*

C'est dans la dernière partie de ce chapitre que nous détaillerons le module de gestion de la référence, au travers de la préposition à *gauche de*, module qui gagnera à être affiné par la notion de cadre.

4.2.1 A la recherche d'un modèle

Dans un système hypertexte, la référence, peut être vue comme la possibilité qu'a l'utilisateur de désigner une ou plusieurs entités pouvant ou non se trouver affichées sur son écran au moment où il parle. Cependant, dans la suite de notre propos, nous ne nous préoccuperons que de références à des objets se trouvant effectivement affichés à l'écran ne prenant pas en compte le cas des références à des objets ayant disparu de l'écran, ex : *iconifie la fenêtre verte, euh... non réaffiche la.*

Comme le mentionne [Gaiffe 92a], dans les dialogues de commandes simples les référents sont peu susceptibles d'être négociés, par conséquent, on peut se permettre un mapping direct du nom d'un objet ou type correspondant. Reste un problème : si on donne un type a priori aux objets, l'utilisateur peut pourtant utiliser un nom plus générique que celui associé au type de l'objet. Par exemple, si on étiquette un objet de l'application comme "lettre", l'utilisateur peut quand même l'appeler un texte. Autrement dit, il faut pouvoir détecter que "une lettre" est un "texte". Etant donné qu'on a fait le choix d'associer directement un nom à un type, on peut indifféremment représenter cet arbre (ou plutôt ce graphe sans circuit) sur les noms ou sur les types d'objets.

Le lecteur pourra se reporter au §3.1.2 où le modèle de traitement des références non spatiales a été exposé.

4.2.2 La référence spatiale

L'opération de référence spatiale est la faculté pour un être humain de retrouver la(les) bonne(s) entité(s) référenciée(s) en faisant appel à tout type de processus cognitif d'origine spatiale. Ces derniers pouvant regrouper des opérations mentales aussi diverses que la perception, la catégorisation, etc. On peut opposer un système opérant des références sur une scène d'objets à un système générant une description de ces objets : si pour un système descriptif il est nécessaire de faire une description précise des objets considérés (voir Soccer²), en revanche, un système de référence peut se contenter du "strict minimum"³, en tenant compte au maximum du contexte lui permettant de discriminer un (resp. des) objet(s) par rapport à d'autres suivant l'une de ses (resp. de leur) propriétés, l'axe d'opposition étant spatial. Cette opposition se fait à l'intérieur d'un cadre (dans l'approche de Gaiffe, l'opposition se faisait dans un ensemble).

Pour illustrer ceci, il suffit de prendre pour exemple la figure 4.1, pour laquelle la phrase 1 serait amplement suffisante en référence alors que la phrase 2 pourrait être nécessaire pour faire un descriptif de ce qui se trouve affiché à l'écran :

- 1) *L'icône qui se trouve au dessus de la fenêtre.*
- 2) *L'icône est au dessus de la fenêtre, légèrement à gauche et près du bord de l'écran.*

Il convient à présent de préciser le statut des axiologies à *gauche de* \rightarrow *non à gauche de*, à *droite de* \rightarrow *non à droite de*, à *dessus de* \rightarrow *non au dessus de*, etc, pour ce faire, il nous faut un modèle qui puisse calculer la partie gauche de ces axiologies, l'autre partie se déduisant de la première.

Plutôt que d'aborder globalement les modèles existants nous les aborderons de manière spécifique au travers de l'étude de la préposition à *gauche de*, étude qui pourrait être généralisée facilement à d'autres prépositions.

4.3 Une étude détaillée de la préposition "à gauche de"

Reprenant l'un après l'autre les modèles existants, nous allons démontrer qu'ils sont inadaptés à référencier efficacement les objets dans un espace de visualisation graphique car, tantôt trop grossiers, tantôt trop précis.

2. Le système Soccer [Schirra 93a] génère un commentaire d'un match de football à partir d'une séquence d'images.

3. A rapprocher de la notion de pertinence de [Sperber 89].

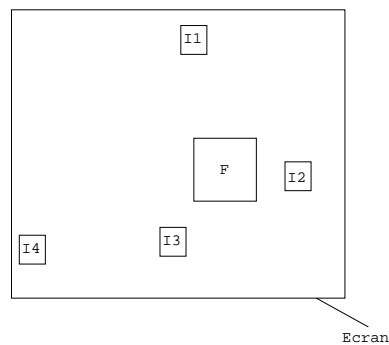


FIG. 4.1 - *Avantages d'une optique différentielle.*

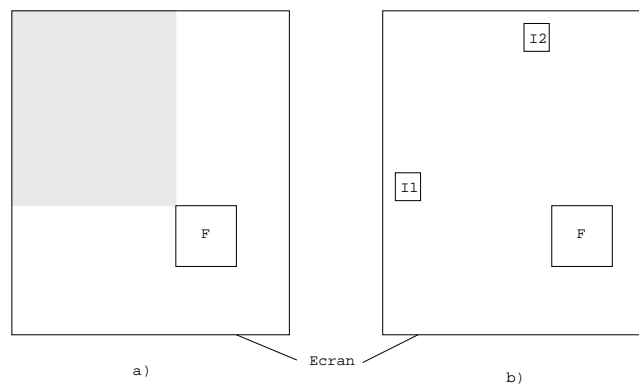


FIG. 4.2 - *Un modèle trop grossier.*

4.3.1 Le modèle de Mukerjee

Assimilant nos fenêtres et icônes à des rectangles nous pouvons reprendre l'étude de [Mukerjee 89]⁴. L'auteur adapte la logique de Allen⁵ pour des dimensions supérieures à 1, projetant chaque solide sur chacun des trois axes, il calcule alors, selon

4. L'étude de [Mukerjee 90] plus récente ne peut pas s'appliquer ici car elle ne fonctionne qu'avec des parallépipèdes ayant des directions frontales non parallèles.

5. Pour un exposé de la logique de Allen voir [Romary 89] p.110-111.

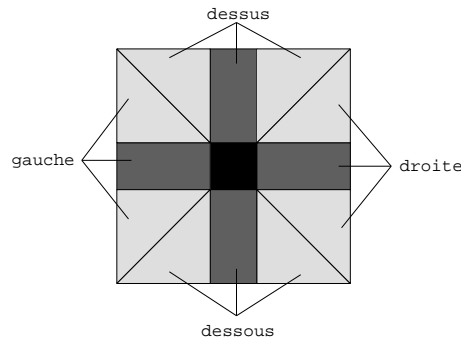


FIG. 4.3 - *Acceptabilité des prépositions selon Briffault.*

Allen, les positions, respectives des segments (ou intervalles) résultants. Ainsi dans le modèle de Mukerjee, tout objet qui se trouve dans la zone grisée de la figure 4.2.a se trouve exactement à la même position par rapport à la fenêtre F . Cependant, reprenant le diagramme de [Briffault 92] p. 70 (voir figure 4.3⁶) nous pensons au contraire que l'icône I1 de la figure 4.2.b pourrait être référencée sans problème comme *l'icône à gauche de la fenêtre F* alors que l'icône I2 serait référencée comme étant *l'icône au dessus de F* . Ce petit exemple illustre le caractère quelque peu trop grossier du modèle de Mukerjee pour le problème qui nous intéresse car la discrétisation de l'espace en termes d'intervalles ne semble pas être assez fine. De plus, le principal reproche que l'on peut faire à ce modèle est qu'il a perdu la plausibilité cognitive qu'avait l'approche de Allen dans le domaine temporel. En effet, nous ne passons pas notre temps à décomposer le monde suivant 2 ou 3 axes pour déterminer ensuite l'ensemble des relations qui sont mises en jeu entre les entités du monde.

6. Les zones hachurées en traits épais correspondent aux zones d'acceptabilité maximale de la préposition, tandis que les secteurs hachurés en clair correspondent aux zones où l'acceptabilité de cette préposition est supérieure à celle des autres.

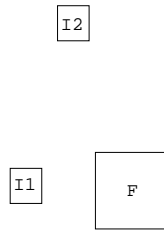


FIG. 4.4 - Différentes acceptabilités de à gauche de.

4.3.2 L'approche de Vandeloise

Vandeloise propose pour les prépositions *à droite de/à gauche de* la règle suivante :

G/D : A est *à droite/à gauche de* B si A se trouve du côté *droit/gauche* de l'orientation latérale⁷ de B.

Ainsi, suivant cette approche, à la figure 4.4, l'icône I2 est autant à gauche de la fenêtre F que l'icône I1 ce qui bien entendu est trop grossier et pratiquement faux dans une optique différentielle. Tout comme au paragraphe précédent, I2 serait vue comme *au dessus de* F plutôt que *à gauche de* F. Vandeloise sent cependant qu'il y a un problème, ainsi, pour la préposition *au dessus de*, l'auteur indique que la relation entre les deux objets doit être saillante (p. 104).

4.3.3 Les nuages de Schirra

Le système SOCCER ([Wahlster 87], [Schirra 93a]) est capable de générer un compte rendu simultané d'événements à partir d'une séquence d'images en cours de traitement. L'application étant celle de commentaires en direct de matchs de football. Pour réaliser ceci Schirra s'appuie entre autres sur des nuages de probabilité⁸

7. **Orientation latérale** : il s'agit d'une ressemblance de famille dont les principaux traits sont la *direction latérale* et la perpendiculaire à l'*orientation générale*.

Orientation générale : ressemblance de famille qui compte pour principaux traits : la *direction frontale*, la *direction du mouvement*, la *ligne du regard*, la *direction dans laquelle sont dirigés les autres organes* : la *perception*, l'*odorat*, l'*ouïe*...

8. Nuages de probabilité : matérialisent le degré d'applicabilité (compris dans l'intervalle [0.0;1.0]) de la préposition concernée).

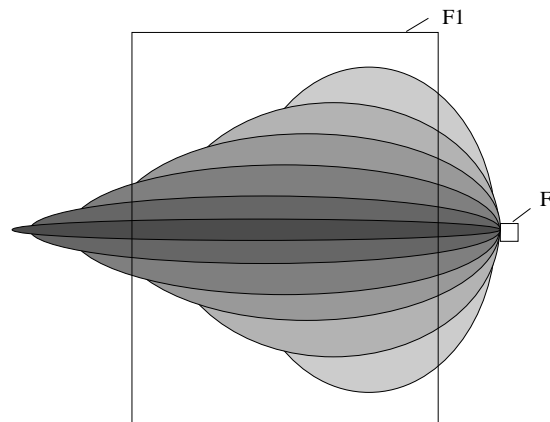


FIG. 4.5 - *Le nuage de probabilité de à gauche de.*

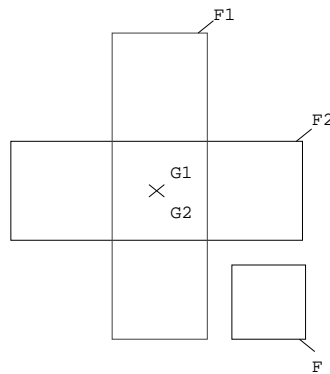


FIG. 4.6 - *Imprécision des centres de gravité.*

dont la forme est définie pour chaque préposition (voir le nuage de *à gauche de* à la figure 4.5).

Si l'on tente d'appliquer tel quel le système de Schirra à nos fenêtres et icônes (en les restreignant à leur centre de gravité comme l'auteur le fait pour les joueurs de football) cela ne fonctionne pas (voir figure 4.6) où l'on a envie de dire : *F1 est à*

gauche de F et *F2 est au dessus de F* et pas quelque chose du type: *F1 est en haut à gauche de F?! et F2 est en haut à gauche de F?!* du fait de la position de leurs centres de gravité G1 et G2.

La solution de prendre les 4 coins n'est pas meilleure (voir figure 4.5) car ces derniers peuvent se trouver hors du nuage alors que la fenêtre F1 est bien à gauche de la petite fenêtre F.

La solution que nous préconisons dans ce cas est la prise en compte simultanée des 4 coins et du centre de gravité de l'objet référencié. Il suffit alors de chercher quel est le nuage (*à gauche de, au dessus de, etc.*) qui maximilise la somme des densités des quatre coins et du centre de gravité⁹.

Si cette méthode semble parfaitement adaptée à un système de génération de texte (car très précise) elle nous semble cependant trop compliquée pour être utilisée dans un système de calcul de la référence car elle s'écarte trop des principes cognitifs simples que mettraient en oeuvre un humain pour effectuer de telles références, nous proposerons au §4.3.6 une autre méthode qui semble mieux refléter ces principes.

4.3.4 Le modèle de Wazinski

Wazinski, travaillant en 2D étudie précisément la position relative de deux rectangles ([Wazinski 91] et [Wazinski 93]). Voulant référencier un rectangle R2 par rapport à un rectangle R1, Wazinski découpe l'espace 2D en 9 sous-espaces définis à partir de R1 (voir figure 4.7.a) il calcule ensuite la proportion de surface de R2 qui se trouve dans tel ou tel secteur.

L'idée semble séduisante car comment approcher plus finement des rectangles qu'en prenant en compte leur surface? Mais là encore, cette étude, trop formelle, s'écarte des principes cognitifs simples que nous utilisons. Elle peut même devenir fausse, il suffit de prendre pour exemple la figure 4.7.b où le modèle de Wazinski concluerai que R2 se situe en haut à droite de R1 car la surface qui se trouve dans le secteur en haut à droite de R1 est bien plus importante que celle se trouvant dans le secteur directement à droite de R1?!

Comme nous le verrons au §4.3.6, le fait de couper l'une des bandes principales "centre horizontal" ou "centre vertical" est déterminant.

9. Il se peut là encore qu'en cas de grande différence de taille entre F et F1 que le centre de gravité de F1 ne fasse lui non plus pas partie du nuage, la seule solution consiste alors à calculer la position de F par rapport à F1 et de prendre la préposition converse, si conversité il y a entre F et F1.

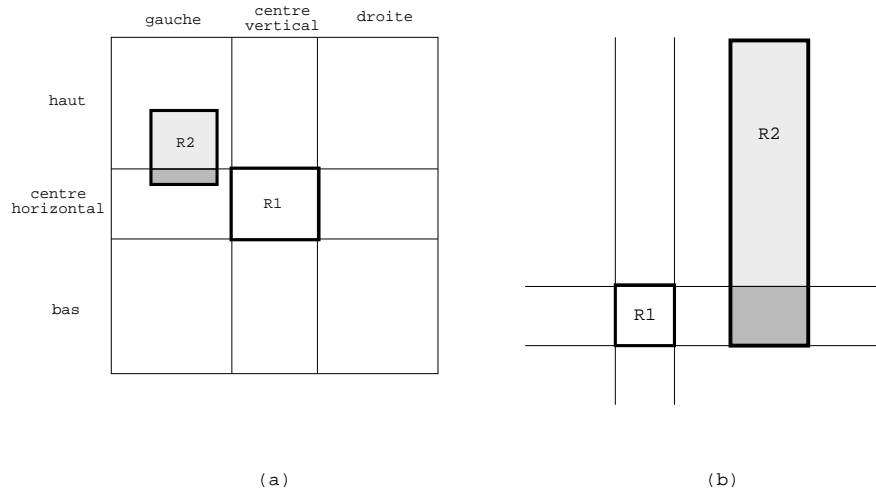


FIG. 4.7 - *L'approche de Wazinski, ses limites.*

4.3.5 Le modèle de Briffault

Briffault esquisse dans [Briffault 92], [Briffault 93] ce que l'on pourrait appeler une vision cognitive des prépositions *à droite de*, *à gauche de*, *au dessus de*, *en dessous de* (voir figure 4.3). Adoptant ce système à nos fenêtres et icônes, l'algorithme qui consisterait à tester si le centre de gravité appartient à tel ou tel secteur ne fonctionne bien entendu pas pour la même raison que celle explicitée à la figure 4.6. Néanmoins, le système fonctionne bien si l'entité à localiser appartient totalement à l'un des secteurs, cependant l'auteur ne dit rien lorsqu'un rectangle chevauche plusieurs secteurs. Il convient donc d'étendre ce système à des configurations quelconques de rectangles.

4.3.6 Vers une autre approche

En nous basant sur le modèle de X. Briffault, nous proposons l'algorithme suivant permettant de déterminer la position d'un rectangle R2 par rapport à un rectangle R1. Cet algorithme fonctionne en 2D pour des configurations de rectangles non inclinés (c.à.d ayant leur base horizontale) ce qui correspond aux positions classiques des fenêtres et icônes dans un système hypertexte. Nous laissons de côté les problèmes

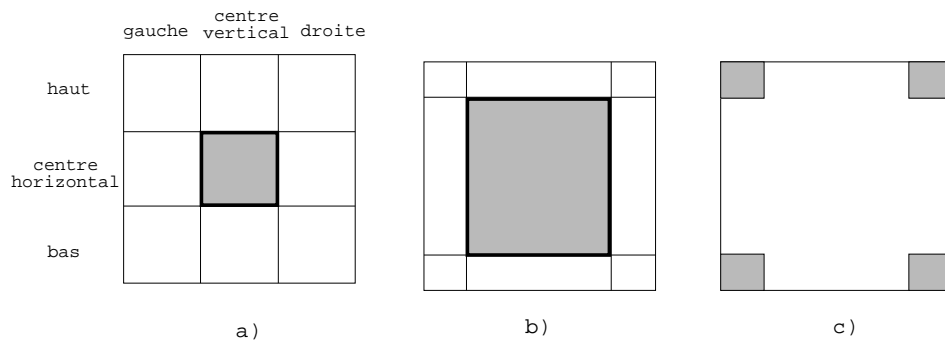


FIG. 4.8 - Référence absolue et référence dans les coins.

de chevauchement du type *devant/derrière*, ex : *la fenêtre verte sur l'icône, détruis les icônes derrière la fenêtre verte*. Enfin, dans ce système, du fait de la position canonique de l'utilisateur face à son écran et de la nature des entités qu'il manipule, nous ne rencontrons pas les conflits classiques entre les orientations intrinsèque, déictique et extrinsèque¹⁰. Pour un exposé détaillé de ces conflits, voir [Sondheimer 76], [Retz-Schmidt 88], [Garnham 89] et [Carlson-Radvansky 93].

Avant d'exposer l'algorithme, reprenant les travaux de [Wazinski 91], notons que notre système fonctionne également pour les localisations absolues : *la fenêtre de droite* et pour les localisations dans les coins : *la fenêtre dans le coin supérieur droit de l'écran*.

Dans le premier cas, on référence par rapport au centre de l'écran qui joue alors le rôle du rectangle de référence (voir figure 4.8.a) dans le second, on référence toujours par rapport au centre de l'écran que l'on agrandit ce qui a pour effet de restreindre les 4 coins (voir figure 4.8.b et 4.8.c).

Voici donc l'algorithme :

1. si l'objet appartient totalement à l'un des quartiers : ok car pris en compte par l'approche de Briffault (ex : R0 peut être référencié (pas localisé!) comme étant à *gauche de* R1 sur la figure 4.9.a).

¹⁰. Cependant, une fois le conflit résolu, on peut appliquer notre algorithme qui va être exposé.

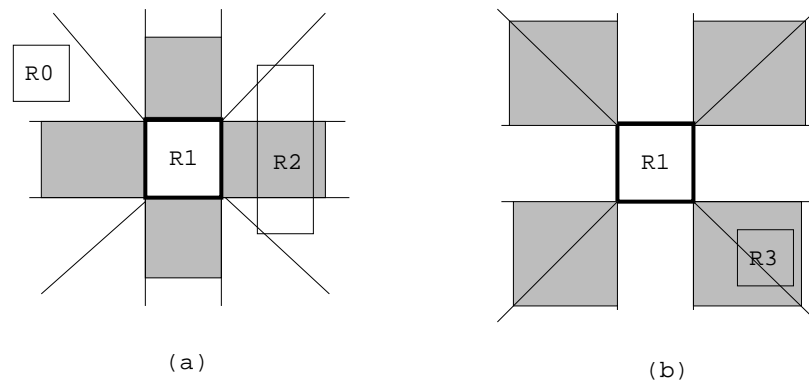


FIG. 4.9 - Notre modèle.

2. **si** l'objet coupe l'une des bandes principales (en grisé sur la figure 4.9.a), et il ne peut en couper qu'une (car nos objets sont perpendiculaires et disjoints), il est, suivant la bande coupée, *à droite de*, *à gauche de*, *au dessus de*, *en dessous de* R1 (ex : R2 peut être référencié comme étant à droite de R1 sur la figure 4.9.a).
3. **sinon** il est nécessairement dans un des coins et coupe 2 bandes (voir figure 4.9.b), il est alors référencié par rapport à ces deux secteurs (ex : R3 peut être référencié comme étant *en bas à droite de* R1 sur la figure 4.9.b).

4.3.7 Quand les cadres volent à notre secours

Comme nous l'avons exposé dans ce rapport la notion de cadre joue un rôle important pour tout ce qui concerne la modélisation de l'espace en langage naturel.

La notion de cadre s'applique tout naturellement au système hypertexte, ainsi, par défaut la recherche du (des) entité(s) référenciée(s) se fait sur la totalité de l'écran mais bien souvent elle pourrait se faire avec profit seulement sur une partie de ce dernier. Il suffit de prendre pour exemple la figure 4.10 sur laquelle on fait la requête suivante :

(3) *Détruis le bouton qui se trouve à droite du bouton Mailbox.*

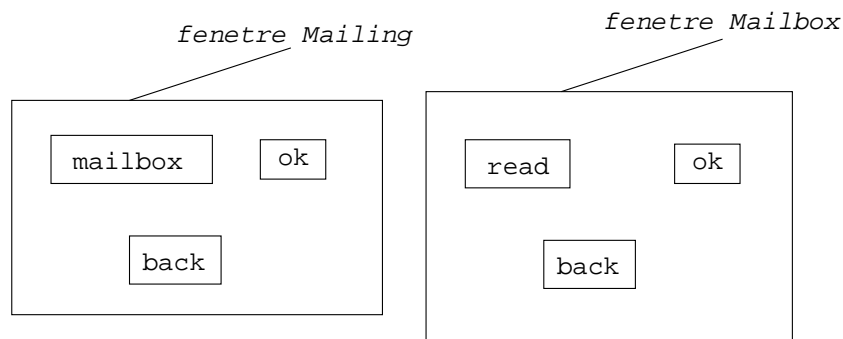


FIG. 4.10 - *De la nécessité des cadres.*

Alors qu'un système classique hésiterait entre les 3 boutons "Ok", "Read" et "Ok", nous proposerions le bouton "Ok" qui se trouve dans la fenêtre "Mailing" car tout comme dans l'exemple des *paragraphes* du §3.1.1, la portée localisatrice de "à droite du bouton Mailbox" se limite à la fenêtre Mailing.

À présent, imaginons la suite de requêtes suivantes :

- (4) *Mets le bouton back de la fenêtre Mailing à gauche.*
- (5) *Rapproche le bouton ok du bouton back.*

Un système classique hésiterait en (5) entre les deux boutons "back", de plus, quand bien même le système aurait trouvé le bon bouton "back" il lui faudrait encore résoudre l'ambiguïté entre les deux boutons "Ok".

Pour nous, la référence au bouton "back" n'est pas ambiguë car nous calculons en priorité les référents dans le cadre courant. Enfin l'ambiguïté entre les deux boutons "ok" peut être levée par le fait que "rapprocher A de B" joue à cadre constant comme on le verra pour les énoncés de positionnement : ici, le cadre englobe le bouton back et ne peut sortir de la fenêtre mailing d'où la référence non ambiguë au bouton ok. En conséquence, il suffit de modifier l'algorithme qui précède en opérant d'abord une recherche dans le cadre englobant le site, ce cadre pouvant éventuellement être limité par des discontinuités spatio-fonctionnelles ou par des objets environnants.

4.4 Conclusion

Nous avons proposé dans ce chapitre un modèle de calcul des référents spatiaux dans un système graphique qui prend finement en compte le contexte, se rapprochant réellement du fonctionnement cognitif humain, modèle qui gagne à être affiné par la notion de cadre.

Une partie du système proposé dans ce chapitre est doré et déjà implémentée en Common Lisp ; pour l'instant, l'interface hypertexte est simulée à l'aide d'un modèle de CAO 3D à facettes planes (chapitre 6).

Les apports de ce modèle doublés de celle des cadres nous conduisent à penser que ces réflexions peuvent affiner des systèmes comme IcpDraw [Caelen 91], ou encore Compèrobot [Andrès 93].

Certains points restent ouverts, en particulier le passage du 2D au 3D qui devrait se faire assez facilement. Par la suite nous intégrerons cette réflexion dans la maquette multimodale que nous développons actuellement qui permettra une désignation d'objets à l'écran au moyen d'un gant de désignation [Bellalem 93], [Romary 94a], [Romary 94b].

Chapitre 5

Le problème du positionnement

5.1 Motivations de l'étude

Un problème important dans les systèmes de dialogue à composante graphique concerne le problème du positionnement d'objets à l'écran qui requiert une bonne modélisation de l'espace et des actions [Schang 94c].

À notre connaissance, il n'existe pas de système qui réalise ce type d'opération tout en effectuant une compréhension fine des mécanismes sémantiques sous-jacents. Des systèmes comme Compèrobot [Andrès 93] ou l'étude de [Arnold 92] n'étudient pas le positionnement en temps que tel mais s'en servent comme outil d'étude des mécanismes plus généraux relatifs à la modélisation de l'espace en langage naturel.

Trop souvent, on pense que des énoncés de positionnement se réduiraient au prototype: *mets ça ici*. Il n'en est rien comme le démontre l'expérience de type Magicien d'Oz mise en oeuvre dans notre équipe en collaboration avec le Cermat [Dauchy 93] et [Mignot 93] dans le but d'étudier le comportement de sujets dans une situation de dialogue multimodal (associant voix et geste). L'espace de la tâche tel qu'il pouvait être vu dans cette expérience est représenté à la figure 5.1. Il comprend d'une part une portion d'appartement (un salon...), avec une fenêtre, une porte et une cheminée, et d'autre part un ensemble de meubles disposés à l'extérieur de la pièce qu'il s'agit de positionner (en partie) dans celle-ci. On peut imaginer que ces meubles sont en quelque sorte en vrac dans un camion de déménagement et que l'utilisateur, dans la peau d'un déménageur, les met en place un à un¹. Deux résultats importants ressortent de cette expérience: tout d'abord, les énoncés qui sont produits spontanément par

1. On trouvera à la fin de ce chapitre un extrait du dialogue spontané d'un utilisateur dans une telle situation de commande.

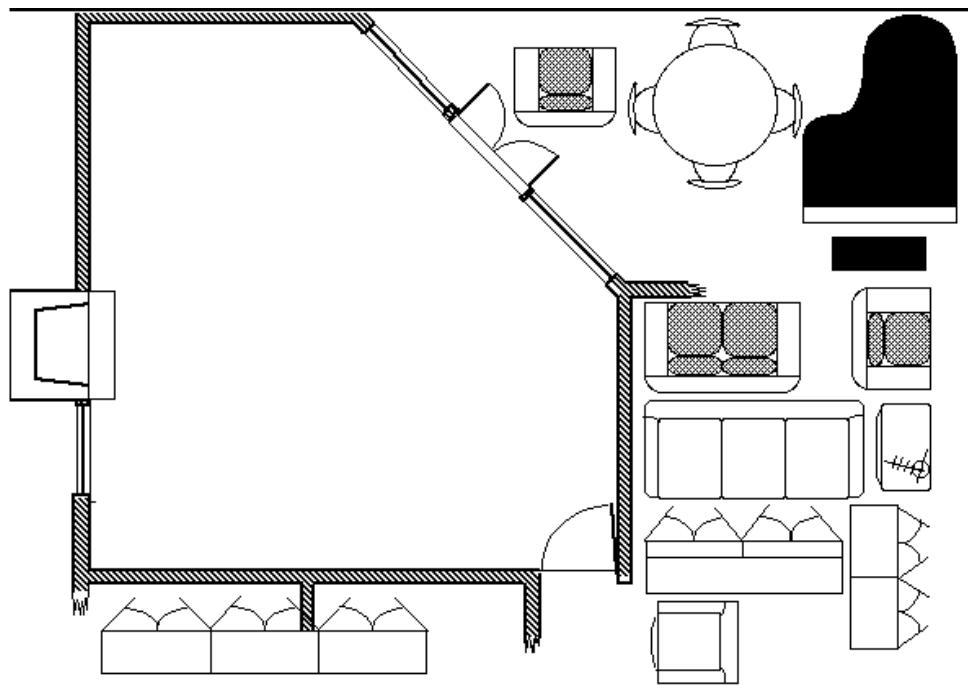


FIG. 5.1 - "le salon".

un utilisateur en situation de commande sont bien loin de l'archétype *mets ça ici*, enfin 38 % des commandes énoncées sont purement vocales et sont utilisées prioritairement pour spécifier des actions abstraites [Mignot 93]. Ces points justifient notre étude concernant les énoncés de positionnement purement vocaux².

Le problème du positionnement semble très important car il repose, pour un bonne part sur une compréhension fine des mécanismes de référence et plus généralement des intentions du locuteur. Bien des problèmes liés à la référence sont sous-jacents aux énoncés de positionnement, prenons la phrase : *rapproche le canapé de la cheminée* où il faut que la référence au canapé ne soit pas ambiguë même si un autre canapé se trouve à l'extérieur de la pièce. Ainsi, *rapprocher A de B* joue dans un espace où se trouvent déjà A et B et cela contraint la référence à A dans

2. Notre réflexion semble suffisamment générale pour lui adjoindre par la suite une composante gestuelle.

la cas où la référence à B n'est pas ambiguë. Dans ce cas précis, la notion de site "la cheminée" devient plus forte que ce qu'on considère habituellement puisqu'un espace incluant ce site devient pertinent pour l'identification de la cible.

On peut envisager l'hypothèse inverse : *mets le canapé dans le salon* où au contraire, seuls les meubles à l'extérieur de la pièce devraient être concernés pour en extraire le canapé.

Ces 2 exemples démontrent que les énoncés du type *mets A prép B* ne fonctionnent pas de manière purement compositionnelle suivant les 3 étapes suivantes :

- recherche de A,
- recherche de B,
- placement de A par rapport à B tel que A se trouve "prép B" (ex : A est à côté de B, A est derrière B, A est devant B).

Voyons à présent un cas particulier des énoncés de positionnement, celui des énoncés en "ici".

5.2 Vers un positionnement plus "cognitif", le cas de *ici*

[Romary 93] démontre que bien trop souvent on a associé *ici* à une portion élémentaire de l'espace, en l'occurrence un point M de coordonnées (x,y) dans le plan de désignation³.

Réduire le lieu à un point a une conséquence immédiate sur les énoncés de positionnement : un objet ne pourra être placé sur l'écran qu'à la condition qu'un point de celui-ci soit associé directement au lieu de désignation.

Cependant, dès que la tâche se complexifie un peu, un certain nombre de problèmes apparaissent. Le lieu correspondant au geste ne peut plus être simplement recherché comme un point précis de l'espace, au risque de conduire à des interprétations pour le moins comiques. Ainsi, si on considère une tâche d'aménagement intérieur, l'énoncé "mets de la moquette ici" associé à un geste ne peut s'interpréter comme une demande limitant le positionnement de la moquette au point précis où il y a eu geste⁴. Il est donc nécessaire d'envisager une certaine extension à ce fameux

3. Cette analyse est partagée par de nombreux auteurs comme en témoignent un certain nombre d'articles des actes des journées IHM 92. On peut remarquer qu'il y a alors un petit côté arbitraire à subordonner la description de l'espace à la définition (en pixel/pouces) de l'écran.

4. Si on imagine de plus qu'il s'agit de carrés de moquette, il n'est peut être pas judicieux de les empiler en un point précis de la pièce.

lieu associé à *ici*.

Parfois même, ce lieu semble être disjoint de la désignation alors que l'interprétation procède des mêmes principes. Ainsi, dans une circonstance similaire au cas précédent, l'interprétation de l'énoncé "mets trois prises de courant ici" doit conduire à ce que trois prises soient globalement associées à la pièce considérée, pour ultérieurement les voir apparaître sur les murs de celle-ci.

L'auteur aboutit donc aux deux conclusions suivantes :

- *ici*, en association à un geste, active un lieu qui semble être autre chose qu'un simple point. Ce lieu possède une certaine étendue et même une certaine structure (ce peut être une pièce dans un appartement). Dans des tâches de positionnement, on associera globalement l'objet au lieu (exemple : les prises à la pièce désignée),
- dans un deuxième temps, le positionnement final de l'objet par rapport au lieu va dépendre de contraintes inhérentes à la tâche et ne paraît pas être du ressort de la sémantique de l'énoncé. Il faudra donc posséder une description de la manière dont on place de la moquette ou des prises de courant dans une pièce (i.e. les prises de courant sur les murs).

Romary démontre ensuite la nécessité d'introduire la notion de pavage qui consiste en fait à découper notre notion de cadre en entités élémentaires. Ce pavage ne correspond bien sûr pas à une représentation unique de l'espace, mais à un point de vue possible que le locuteur cherchera à faire partager à l'auditeur dans le discours. Ainsi, un appartement pourra être vu comme un ensemble de pièces, un jeu d'échec comme un ensemble de cases, etc.⁵

La notion de point de vue associée à un pavage est particulièrement importante puisqu'elle permet de définir différentes configurations spatiales pour une même entité. Supposons ainsi une table, celle-ci peut-être vue dans le cadre d'un pavage représentant une dimension verticale pour exprimer des relations du type "la lampe se trouve au dessus de la table" ou en opposition avec d'autres surfaces planes ("ne mets pas le livre sur la table, pose le sur l'étagère"). La définition d'un pavage de l'espace sur lequel va s'opérer l'interprétation de *ici* permet de donner une certaine granularité à la précision de la référence spatiale associée.

Finalement, l'auteur énonce la règle suivante pour *ici* :

Ici réalise un filtrage sur une structure de pavage de l'espace pour isoler l'un des éléments de cette structure qui se trouve directement mis en évidence par

5. Entre deux unités, on peut avoir besoin d'un espace de séparation (pour éviter les cas limites).

le locuteur (i.e. intentionnellement), soit par la simple situation d'énonciation (position du locuteur ou du message), soit par un geste lorsque l'unité de pavage est relativement fine.

Dans le paragraphe suivant, nous proposerons un modèle qui s'appuie sur la notion de cadre (notion qui semble pouvoir subsumer celle d'espaces de référence). Il s'agit d'un modèle à cinq niveaux d'affinage permettant la prise en compte d'énoncés de positionnement généraux. Ce modèle sera validé sur une partie du corpus de dialogue de type Magicien d'Oz qui a été recueilli dans notre équipe.

5.3 Un modèle général pour les énoncés de positionnement

Nous proposons d'étudier spécifiquement les énoncés du type *mettre*⁶ *A prép B* (où A et B peuvent être n'importe quels syntagmes du moment qu'ils ne contiennent pas à leur tour des informations spatiales); ainsi les phrases *installez le canapé face à la cheminée*, *placer le fauteuil perpendiculairement au canapé*, etc. sont valides. Le schéma abstrait du positionnement pour les phases du type *mettre A prép B* est représenté à la figure 5.2.

Alors que le cadre final est à composante essentiellement spatiale, le cadre initial n'est pas spatial, il est purement esembliste au sens où il ne sert qu'à piocher un ou plusieurs éléments dans un ensemble d'objets rendus saillants. Le calcul du (des) référent(s) du cadre initial est effectué par le modèle de B. Gaiffe [Gaiffe 93]. Notre attention portera essentiellement sur le calcul du cadre final.

Pour instancier effectivement ce schéma, nous proposons un modèle à cinq niveaux, chaque niveau ayant pour rôle d'affiner le schéma de la figure 5.2 :

1. la contribution de la notion de cadre,
2. la contribution des indices linguistiques,
3. la contribution purement spatiale,
4. la contribution de la tâche et du contexte,

6. Notons que *mettre* peut être remplacé par tout verbe possédant une sémantique voisine comme *placer*, *installer*, etc.

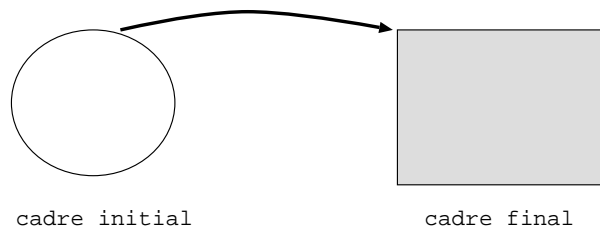


FIG. 5.2 - *Modèle abstrait pour le positionnement.*

5. la contribution des connaissances sur le monde (ou connaissances encyclopédiques chez [Klein 93]).

Au terme de ces cinq affinages ou contributions, on doit déboucher sur un cadre initial et un cadre final suffisamment précis pour qu'il n'y ait pas d'ambiguïtés dans le choix des éléments à positionner et de la (des) position(s) qui sera (seront) occupée(s) par cette (ces) entité(s). Dans le cas contraire, l'ordinateur sera amené à demander plus d'information. C'est le cas pour la requête 1) du corpus où il y a ambiguïté au niveau du cadre initial car deux choix de canapés sont possibles⁷. Nous allons à présent voir en détail ces différentes contributions.

5.3.1 Le modèle à cinq niveaux reposant sur la notion de cadre

1) Contribution des cadres :

Les caractérisations des quatre évolutions possibles des cadres ont été exposées au §3.3, nous les rappelons ici :

- conservation : $\text{cadre-f}' = \text{cadre-f}$ ⁸ C'est le cas par défaut. Ce cas répond à un principe d'économie cognitive, l'utilisateur ne va pas "sauter du coq à l'âne" et va raisonner de manière privilégiée dans le cadre final courant.
- rétrécissement : le nouveau site appartient à l'ancien cadre final et $\text{cadre-f}' \subset \text{cadre-f}$.

7. On pourrait croire de prime abord que l'utilisateur ne respecte pas les maximes de Grice (celle d'être suffisamment précis en particulier), en fait il n'en est rien, le locuteur dans sa perspective n'a même pas prêté attention au canapé clair qui se trouve en dessous du canapé grisé, pour lui, seul ce dernier a été pris en compte.

8. *cadre* désigne l'ancien cadre, *cadre-f'* désigne le nouveau cadre qu'il faut calculer.

- agrandissement : $\text{cadre-f}' \cap \text{cadre-f} \neq \emptyset \wedge \text{cadre-f}' \not\subseteq \text{cadre-f}$. C'est le cas lorsqu'une entité du cadre final courant sert à localiser une autre entité (resp. est localisée par rapport à une autre entité) qui n'appartient pas au cadre final courant.
- fermeture : $\text{cadre-f}' \cap \text{cadre-f} = \emptyset$. C'est le cas lorsqu'on localise une entité par rapport à une autre, les deux entités n'appartenant pas au cadre final courant.

2) Contribution des indices linguistiques :

Dans un premier temps nous souhaitons mettre l'accent sur les marqueurs linguistiques comme à *présent*, *maintenant*, *etc.* qui sont bien souvent précédés d'un temps d'attente (noté [...]), voir entre (6) et (7) et entre (13) et (14). Ces marqueurs nous indiquent que l'utilisateur effectue une sorte de réinitialisation dans la conception qu'il a de la scène, cette réinitialisation s'accompagne généralement d'un temps de réflexion qui apparaît dans le dialogue. Ce temps de réflexion permet alors de concevoir la scène sous un autre angle, typiquement, un *agrandissement* ou une *fermeture*. Les opérations comme la *conservation* ou la *focalisation* requièrent quant à elles des temps de réflexion bien moindres car elles se font dans la continuité du raisonnement⁹.

Actuellement, nous travaillons à la mise en oeuvre d'autres indices linguistiques qui contribueraient encore à affiner le schéma abstrait de la figure 5.2. Ainsi pour les énoncés de positionnement du type *mettre A prép B*, si A est un défini (ex : *mets le fauteuil à droite du canapé*) ou un démonstratif (*mets ce fauteuil*¹⁰ à droite du canapé) le cadre initial appartient nécessairement à l'écran, si c'est un indéfini il y a de fortes chances pour que le cadre initial n'appartienne pas à l'écran, le ou les élément(s) étant pioché(s) dans un ensemble abstrait.

La préposition *ici* quant à elle marque toujours une focalisation dans le raisonnement, même si cette focalisation n'est que temporaire¹¹.

3) Contribution purement spatiale :

Cette contribution va préciser l'allure du cadre final, on trouvera un exemple de l'allure de ce dernier pour *près de* et *loin de* à la figure 5.3 (la zone grisée désigne la

9. Il conviendrait bien sûr de valider ce résultat par une étude psycho-linguistique.

10. *Ce fauteuil* peut permettre une reprise anaphorique ou au contraire être coréférent avec un geste de désignation.

11. Pour une étude détaillée de la préposition *ici* voir [Romary 93].

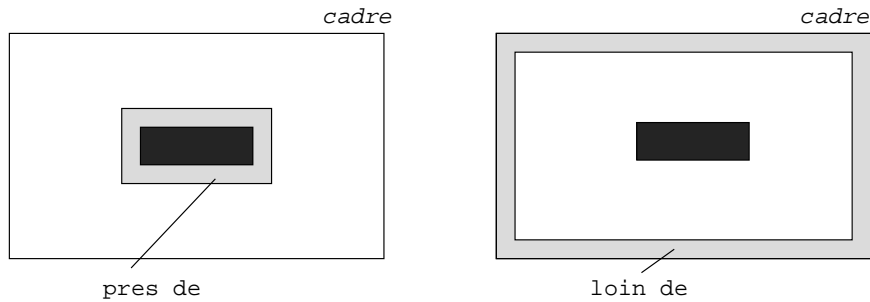


FIG. 5.3 - "près de - loin de."

zone concernée, la zone foncée désigne le site). L'image mentale¹² choisie est en 2D car l'être humain raisonne de manière privilégiée en 2D.

4) Contribution de la tâche et du contexte :

Le contexte va surtout servir à affiner l'étendue du cadre final, ainsi, rejoignant [Pribbenow 93] et [Schang 94b], il faut que le cadre final ait une extension suffisante, extension qui peut cependant être réduite par des objets environnants.

Le contexte va également préciser si le cadre final est de type 1D, 2D, 2D 1/2¹³ ou 3D, notons qu'il ne faut pas voir ces types 1D, 2D, etc, au sens géométrique strict du terme mais plutôt dans un sens métaphorique à savoir que l'on peut s'intéresser à une ligne (correspondant par exemple au cas d'une file de voyageurs attendant à un guichet), à un plan (comme le dessus d'une table), à un plan avec une notion d'épaisseur en plus (c'est le cas dans notre corpus) et enfin à l'espace de sens commun de tous les jours (par exemple dans le cas de la conception d'une scène architecturale).

La tâche influe principalement sur l'appartenance du cadre initial et du cadre final, typiquement pour le corpus de dialogue, le cadre final sera pris dans le salon de manière privilégiée alors que le cadre initial sera l'ensemble des objets qui se

12. Au sens de [Denis 89b] et [Denis 89a]. Dans cette image mentale, les parties de l'objet conservent en somme leurs positions relatives. Les distances entre les différentes parties paraissent également préservées : si nous imaginons visuellement la Tour Eiffel, la distance séparant le premier et le dernier étage est perçue comme plus petite que la distance séparant le deuxième et le troisième étage.

13. 2D plus la notion d'épaisseur.

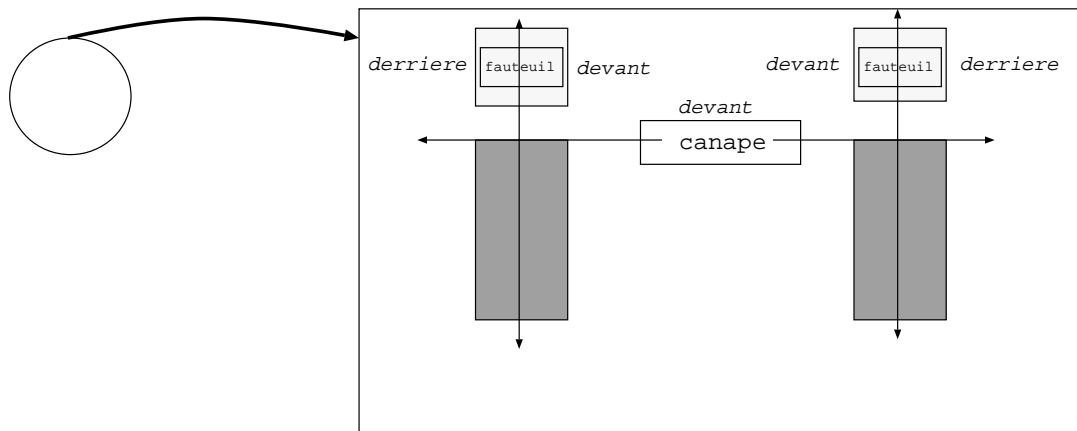


FIG. 5.4 - Orientations et zones interdites.

trouvent à l'extérieur du salon.

5) Contribution des connaissances sur le monde :

Ce niveau sert essentiellement à préciser si possible la position exacte de la (des) entités à positionner. Ainsi pour la requête (1) du corpus où il faut placer le canapé en face de la cheminée, il faut qu'il soit placé suffisamment près de cette dernière afin de profiter de la chaleur, mais pas trop près non plus afin de ne pas le brûler. Il faut également laisser assez de place autour du canapé pour pouvoir bien circuler, etc.

Ce niveau permet d'affiner la posture de l'entité à placer. Ainsi, en termes de rotation, une fois la position trouvée, il convient d'orienter convenablement l'entité considérée, c'est le cas pour les requêtes (3) et (4) qui peuvent être traduites par *installez ce fauteuil perpendiculairement au canapé*. Les connaissances sur le monde vont nous permettre de ne pas placer le fauteuil dans la zone hachurée, car les fauteuils se placent rarement dos au canapé (voir figure 5.4).

De la même manière, les connaissances sur le monde vont nous indiquer comment orienter les fauteuils relativement au canapé, le devant du fauteuil étant orienté vers

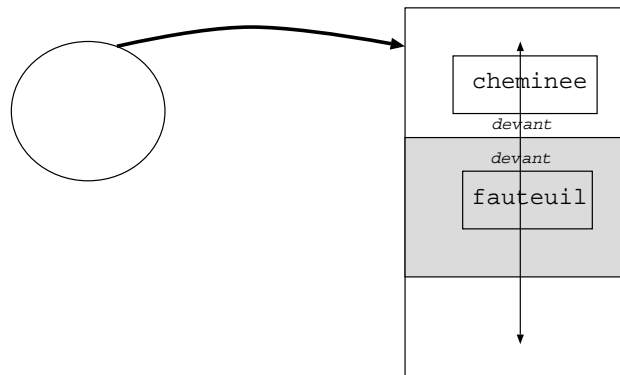


FIG. 5.5 - Mets le fauteuil en face de la cheminée.

le devant du canapé (voir figure 5.4).

Nous sommes conscients que ces connaissances sur le monde sont relativement difficiles à mettre en oeuvre mais elles contribuent à rendre le système vraiment efficace. Nous travaillons actuellement à la définition d'une méthodologie de représentation et de minimalisation (par le biais d'un choix d'un faisceau de points de vue) dans le cadre de la tâche spécifique d'aménagement du salon et dans la tâche spécifique du système de visualisation graphique.

Voyons à présent comment cet affinage fonctionne sur le corpus sus-cité.

5.3.2 Fonctionnement sur un exemple extrait du corpus de dialogue

1) *installez le canapé face à la cheminée.*

- contribution des cadres : au départ il n'y a pas de cadre initial et final¹⁴.
- contribution des indices linguistiques : on sait que le cadre initial appartient à l'écran car *le canapé* désigne un groupe nominal défini.

¹⁴. Tout au plus pourrait-on parler de cadres finaux potentiels : le cadre final qui serait défini relativement à la cheminée, celui relatif à la porte et enfin celui relatif à la fenêtre, ce point est actuellement en cours d'étude [Romary 95].

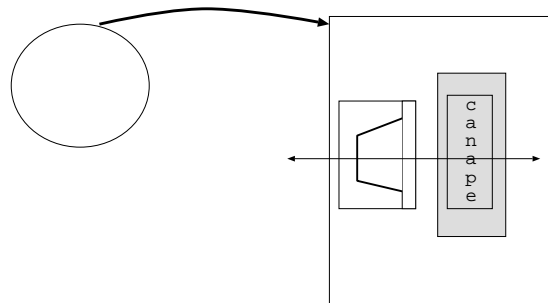


FIG. 5.6 - *Cadre initial et cadre final pour mets le canapé en face de la cheminée.*

- contribution purement spatiale: *face à* impose des contraintes quand au placement du fauteuil par rapport au canapé (voir figure 5.5).
- contribution de la tâche et du contexte:
 - le contexte nous indique qu'il faut nous retrouver avec une image mentale en 2D 1/2, il nous donne également l'étendue approximative du cadre en se basant sur la cheminée, de même que son étendue maximum : il ne faut pas qu'il sorte du salon.
 - la tâche nous indique que de manière privilégiée le cadre final appartient au salon et le cadre initial fait partie de l'extérieur de ce dernier.
- contribution des connaissances sur le monde : voir le §3.1 point 5).

Finalement on obtient le schéma de la figure 5.6 qui est ambigu pour le cadre initial car on hésite entre les deux canapés.

2) le plus petit.

- contribution des cadres : par défaut, on conserve le précédent cadre.

Cette requête va servir à désambiguïser le cadre initial, parmi les 2 canapés, on choisit le plus petit. Finalement, il ne reste plus qu'à placer le canapé en position et place que nous avons définis à la précédente requête.

3) 4) et 5) On peut résumer ces 3 phrases par *installez ce fauteuil perpendiculairement au canapé*.

- contribution des cadres : on a un canapé dans notre cadre final, nous sommes donc en présence d’une focalisation.
- contribution des indices linguistiques : on sait que le cadre initial appartient à l’écran car *ce fauteuil* désigne un groupe nominal démonstratif.
- contribution purement spatiale : le fauteuil sera placé perpendiculairement au canapé (la perpendiculaire étant prise par rapport à la plus grande longueur du canapé), voir figure 5.4.
- contribution de la tâche et du contexte : idem que pour la phrase 1).
- contribution des connaissances sur le monde : comme mentionné à la figure 5.4 la zone grisée est interdite et l’orientation du fauteuil est déduite à partir de celle du canapé, le fauteuil étant placé à une certaine distance prototypique du canapé dans la direction perpendiculaire.
A ce stade, le cadre initial n’est plus ambigu, il convient cependant de préciser dans le cadre final si le canapé doit être placé dans la zone de droite ou de gauche. Notre système poserait la même question que le magicien d’Oz (g).

6) *oui, ici exactement* (geste).

Ici réalise une focalisation temporaire pour affiner la position du fauteuil.

[...] Ce temps d’attente nous informe qu’il risque d’y avoir une rupture du cadre final par une fermeture ou un élargissement.

7) *placez à présent l’autre fauteuil perpendiculairement au canapé*.

- contribution des cadres et contribution des indices linguistiques : le marqueur à *présent* accentue encore le risque de rupture de type *agrandissement* ou *fermeture*.
- contribution purement spatiale : le fauteuil sera placé perpendiculairement au canapé (la perpendiculaire étant prise par rapport à la plus grande longueur du canapé), voir figure 5.4.
- contribution de la tâche et du contexte : idem requêtes 3) 4) et 5).

– contribution des connaissances sur le monde : idem requêtes 3) 4) et 5).

Finalement, on place le fauteuil comme en 3) 4) et 5) en sachant qu'il y a déjà une place de prise (le bon fauteuil étant retrouvé à l'aide du modèle de [GaiFFE 93]).
8) *ici* (geste)

Idem 6) avec la focalisation temporaire.

[...] Risque de rupture, etc.

5.4 Conclusion

Ce chapitre démontre l'importance du problème du positionnement dans les systèmes à forte composante graphique. Il illustre également le fait que la notion de cadre est aussi adaptée à la prise en compte des phénomènes généraux de modélisation de l'espace en langage naturel que ceux relevant du positionnement dans les systèmes de dialogue homme-machine finalisés. Il convient à présent d'effectuer une implémentation grandeur nature de ce modèle.

Voici l'extrait du corpus de dialogue de [Mignot 93] :

<i>Locuteur</i>	<i>Magicien</i>
1) installez le canapé face à la cheminée	a) lequel?
2) le plus petit	b) compris
3) installez un fauteuil	c) (action) ici?
4) celui-ci (geste)	d) lequel?
5) perpendiculairement au canapé	e) compris
6) oui, ici exactement (geste)	f) où exactement?
[...]	g) (action) ici?
7) placez à présent l'autre fauteuil perpendiculairement au canapé	h) compris
8) ici (geste)	i) où exactement?
[...]	j) compris
9) installez le piano dans l'angle près de la porte	k) (action) êtes-vous satisfait?
10) rapprochez le de la cheminée	l) compris (action)
[...]	m) où exactement?
11) installation de la table	n) (action) ici?
12) dans l'emplacement où il y a suffisamment de place	o) compris (action)
13) ici (geste)	p) lequel?
[...]	
14) à présent placez le placard	
15) le plus petit	
[etc...]	

Chapitre 6

La maquette informatique

6.1 Définition du système à réaliser

Nous souhaitons valider la notion de cadre au travers des énoncés de référence (chapitre 3) et de positionnement (chapitre 4). Pour ce faire, il nous faut une plateforme spatiale qui soit la plus évolutive possible d'où le recours à un modèle de type CAO.

6.2 Notre système

6.2.1 "Une base solide" : le modèle de CAO

Notre maquette se base sur un modèle de CAO-3D à facettes planes orientées. Pour le situer, notons qu'il présente sensiblement les mêmes possibilités que le système KALIPSOS dont fait état [Arnold 92]. Le système présenté a été réalisé en Common-Lisp. Nous effectuons une construction incrémentale des objets à l'aide d'opérateurs tels que :

- CRÉER/DÉTRUIRE-POINT,
- CRÉER/DÉTRUIRE-ARÊTE,
- CRÉER/DÉTRUIRE-FACE.

Empruntant les principes sous-jacents aux L.G.B¹, nous n'accédons à nos objets qu'au travers d'opérateurs de haut niveau comme :

- CRÉER-PARALLÉPIPÈDE,
- CRÉER-CYLINDRE,
- CRÉER-PYRAMIDE,
- CRÉER-SPHÈRE,
- DÉTRUIRE-ENTITÉ.

Les opérations d'affichage et de modifications d'entités à l'écran sont réalisées par les quatre opérateurs suivants :

- AFFICHE-ENTITÉ,
- EFFACE-ENTITÉ,
- TRANSLATE-ENTITÉ dans l'espace 3D,
- ROTATION-ENTITÉ autorise des rotations autour des 3 axes.

6.2.2 Caractéristiques détaillées du modèle

Tout comme le préconise [Gardan 87], nous manipulons quatre types de coordonnées dans le modèle :

- les coordonnées réelles (coordonnées 3D dans le monde réel),
- les coordonnées observateur (coordonnées 3D réelles transférées dans le repère centré sur l'utilisateur),
- les coordonnées dans la fenêtre (coordonnées 2D projetées),
- les coordonnées dans la cloture (coordonnées 2D sur l'écran).

Pour un exposé détaillé des formules de transformation d'une coordonnée à l'autre voir [Gardan 87] ou [Dony 85].

Les entités peuvent être affichées dans quatre modes :

- fil de fer (voir figure 6.1.a),

1. **L.G.B**: Langages graphiques de Base, il s'agit de modèles réalisant des interfaces de conception graphiques "propres" en se basant sur une construction modulaire des entités de la tâche, voir [Gardan 87] pour un exposé détaillé.

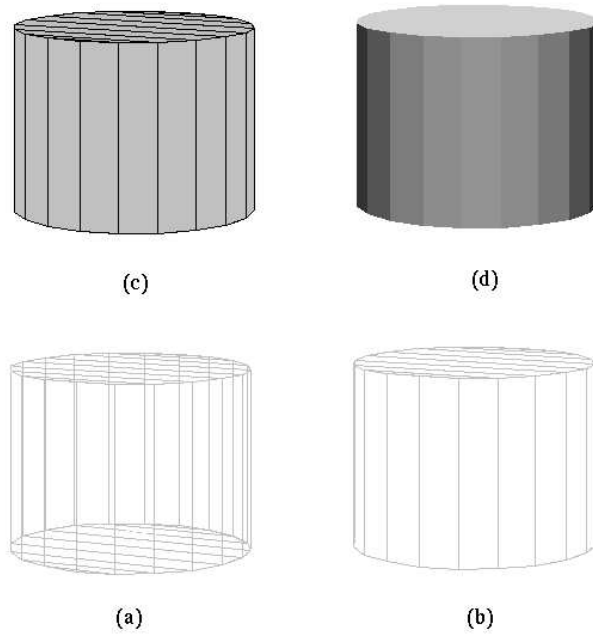


FIG. 6.1 - *Les 4 modes graphiques possibles.*

- fil de fer sans arêtes arrières (voir figure 6.1.b),
- facettes planes + éclairage constant + élimination des facettes arrières + tracé des arêtes (6.1.c),
- rendu réaliste : les facettes sont éclairées par une source lumineuse placée à l’infini, les faces arrières étant éliminées (6.1.d). Pour ce faire, le modèle de Lambert a été utilisé².

Nous nous sommes limités à ce type de rendu, laissant de côté des modèles comme celui de Gouraud ou Phong³ qui auraient considérablement ralenti les performances de notre système.

6.2.3 Adaptation au système de multifenêtrage

Le modèle 3D s’adapte naturellement à la simulation du système de multifenêtrage (de type 2D 1/2 : 2D plus la notion d’épaisseur).

Pour ce faire, les fenêtres et icônes sont construites dans le plan YoZ. Tandis qu’une icône est représentée par une facette plane, une fenêtre se trouve en revanche représentée par 2 à 4 facettes (une pour le corps de la fenêtre, une pour l’entête et éventuellement une à deux facettes pour l’ascenseur horizontal et vertical).

L’oeil étant placé sur l’axe Ox, la visualisation de la scène donne l’impression de se retrouver en présence d’un système de multifenêtrage.

Pour chaque objet, nous avons :

- Des informations de bas niveau :
 - le type de l’objet (ex : *Window*),
 - le nom de l’objet (ex : *Emacs*, *Loria*, *etc.*),
 - l’enveloppe 3D qui contient l’objet dans le monde réel,
 - la liste des facettes le composant,

2. Dans ce modèle, l’intensité de l’éclairage de chaque face est calculé par : $I_d = k * \frac{\mathbf{N} \cdot \mathbf{L}}{\|\mathbf{N}\| \cdot \|\mathbf{L}\|}$ où \mathbf{N} est la normale à la face, \mathbf{L} le vecteur directeur de la source de lumière, k une constante. Cette méthode suppose que la source de lumière se trouve à l’infini, (le produit scalaire $\mathbf{N} \cdot \mathbf{L}$ est alors constant sur toute la surface de la face).

3. Pour un exposé détaillé de ces modèles voir [Peroche 90].

- l’enveloppe 2D de l’objet dans l’écran (qui sert pour le calcul de la référence ainsi que pour les déplacements des objets : si on déplace un objet, on ne réaffiche que les objets éventuels qui ont pu être affectés par ce déplacement ce qui accroît les performances du système).
- Des informations de haut niveau :
 - le champ *perceive* qui nous indique comment est perçue notre entité par l’utilisateur. Ainsi, une fenêtre et une icône sont du type *Window* pour l’ordinateur mais sont perçues différemment par l’utilisateur⁴.
 - le couleur globale de l’objet (si elle existe) sachant que chaque facette peut avoir sa propre couleur.

Le fait de disposer d’un écran fait que l’ordinateur et le locuteur partagent ensemble un certain nombre d’objets (pas comme dans un roman où le locuteur visualise peu à peu la scène par les descriptions qui lui en sont faites, ici il a la scène sous les yeux).

L’utilisateur perçoit nécessairement la scène affichée par l’ordinateur d’une certaine manière et de façon structurée en réalisant une appropriation perceptive que doit aussi réaliser l’ordinateur suivant le schéma :

monde réel \longrightarrow *monde perçu* \longrightarrow *monde conceptualisé*.

Actuellement, nous travaillons à la formulation de ces différentes étapes d’appropriation perceptive [Romary 95].

6.3 Conclusion

Il nous reste à présent à mettre en oeuvre le module de référence et de positionnement pour lesquels nous espérons proposer rapidement une architecture. Tout comme Hernández, nous préconisons un modèle qualitatif :

[Hernández 94] p. 8 - *Qualitative representation provides mechanisms for representing those features that are unique or essential, whereas a quantitative representation allows to represent all those values that can be expressed with respect to a predefined unit.*

4. Plus généralement, en présence d’un objet x , un individu construit une représentation perceptuelle de cet objet x : y ; c’est pourquoi on dit que y est un percept, les propriétés s’appliquent donc sur des percepts c.à.d sur des représentations internes issues des mécanismes de perception et non sur les objets eux-mêmes.

Ce point semble très important, car comme le mentionne [Habel 87], l'un des points sensibles de l'IA réside dans le choix des représentations adéquates pour les connaissances manipulées. Ces représentations constituent non seulement le point de départ et le résultat des processus de calcul mais elles constituent également le type de déductions dont est capable le système.

Le système réalisé est donc un premier pas vers l'ensemble des propriétés que devrait posséder la maquette finale. Comme l'énonce [Arnold 92]: *un environnement de conception informatisé doit, pour être convivial, comporter une interface permettant de passer aisément d'énoncés verbaux à des images (et vice versa).*

Chapitre 7

Conclusion

Nous avons vu que les tentatives de formalisation de la sémantique des expressions spatiales basées exclusivement sur les propriétés géométriques des entités étaient inadaptées car elles ne couvrent qu'un faible nombre des usages des éléments lexicaux. Les travaux de Vandeloise et Herskovitz ont clairement montré qu'il est nécessaire de tenir compte des propriétés fonctionnelles des objets et des constituants pragmatiques des relations spatiales pour décrire de façon satisfaisante la multiplicité des significations des prépositions spatiales.

Nous avons ensuite démontré la nécessité d'introduire la notion de cadre en général et plus spécifiquement pour l'espace dont nous proposons une caractérisation détaillée. Par la suite, nous décrivons les différentes évolutions discursives de ces derniers : la *conservation*, le *rétrécissement*, l'*agrandissement* et la *fermeture*. À travers l'élaboration de la notion de cadre, nous pensons avoir contribué à l'entreprise de formalisation de l'espace de sens commun.

Dans la suite nous revenons au problème qui nous préoccupait initialement à savoir la réalisation d'une application capable de piloter une interface graphique hypertexte. Pour ce faire, nous nous sommes intéressés au problème de la référence en général puis à la référence spatiale en particulier. Après avoir "épluché" les modèles existants nous proposons notre modèle que nous affinons par la notion de cadre.

L'intérêt que nous avons porté au problème de la référence fait immédiatement ressurgir le problème concomitant à savoir celui du positionnement. Nous avons proposé un modèle cognitif basé sur la notion de cadre qui s'applique aux énoncés de

positionnement pour interagir avec l'environnement graphique hypertexte.

Au terme de cette étude, nous esquissons le modèle de CAO qui constitue la première partie de notre maquette informatique visant à la validation du concept de cadre.

La suite logique de ce rapport consiste en premier lieu à approfondir encore la notion de cadre ce qui nous permettra d'en donner une formalisation et d'illustrer cette notion au travers d'une implémentation traitant des énoncés de positionnement. Ainsi, tout comme [Vieu 91], nous pensons qu'au terme de cette première étape, il apparaît clairement que les études de nature empirique sur la signification des marqueurs spatiaux sont un préalable obligé à toute construction d'un système visant à modéliser l'espace de sens commun.

Bibliographie

- [Abraham 91] M. Abraham, R. Afzali and C. Jouis. Les archétypes cognitifs sont générateurs de représentation des connaissances. *T.A informations, Bulletin semestriel de l'ATALA*, 32(1):47–64, 1991.
- [Andrès 93] M. Andrès. Une représentation des figures et de leurs déplacements dans Compèrobot. In *Les cahiers du LAIAC*, volume 13, pages 1–14. 1993. Caen - France.
- [Arnold 92] M. Arnold and C. Lebrun. Utilisation d'une langue pour la création de scènes architecturales en image de synthèse expérience et réflexions. In *Intellectica*, volume 99, pages 151–186. 1992.
- [Asher 94] N. Asher and P. Sablayrolles. A typology and discourse semantics for motion verbs and spatial prepositions in french. 1994. (Sous presse).
- [Aurnague 90] M. Aurnague, M. Borillo and L. Vieu. A cognitive approach to the semantics of space. In *Actes COGNITIVA AFCET*, pages 321–328. Madrid, 1990.
- [Aurnague 93a] M. Aurnague and L. Vieu. *A three-level approach to the semantics of space*, pages 393–439. Mouton de Gruyter, 1993. Cornelia Zelinsky-Wibbelt (Editor).
- [Aurnague 93b] M. Aurnague, L. Vieu and A. Borillo. Vers une représentation formelle des concepts spatiaux dans la langue. In *Images et Langages. Multimodalité et Modélisation Cogni-*

- tive*, pages 109–120. Colloque interdisciplinaire du CNRS, Paris, 1er et 2 Avril 1993.
- [Bellalem 93] N. Bellalem and L. Romary. Le dialogue homme-machine multimodal : vers la compréhension du geste de désignation. In *L'Interface des mondes réels et virtuels*, pages 217–227, Montpellier, 22-26 Mars 1993. 2^{ème} Conférence Internationale.
- [Boers 87] F. Boers. Behind, beyond, under, underneath, beneath, below: a descriptive and explanatory study of spatial and non-spatial senses. In *Antwerp papers in linguistics*, volume 53. 1987.
- [Bourguet 92] M. L. Bourguet. IcpPlan : Dialogue multimodal pour la conception de plans architecturaux. pages 369–374, 1992. 19^{èmes} JEP Bruxelles.
- [Briffault 92] X. Briffault. Modélisation informatique de l'expression de la localisation en langage naturel, 1992. Thèse d'Université de Paris VI.
- [Briffault 93] X. Briffault. Modélisation informatique de l'expression de la localisation spatiale en langage naturel. In *Images et Langages. Multimodalité et Modélisation Cognitive*. Colloque interdisciplinaire du CNRS, Paris, 1er et 2 Avril 1993.
- [Caelen 91] J. Caelen. Interaction multimodale dans ICP-Draw. expérience et perspectives. *Greco PRC*, pages 1–27, Avril 1991.
- [Carlson-Radvansky 93] L. A. Carlson-Radvansky and D.E. Irwin. Frames of reference in vision and language: Where is above? In *Cognition*, volume 46, pages 223–244. 1993.
- [Carré 91] R. Carré, J.F. Dégremont, M. Gross, J.M. Pierrel and G. Sabah. *Langage Humain et Machine*. Presses du CNRS, 1991.
- [Charolles 94] M. Charolles. Sous presse, 1994.

- [Cuyckens 93] H. Cuyckens. *The Dutch spatial preposition in : A cognitive-semantic analysis*, pages 27–71. Mouton de Gruyter, 1993. Cornelia Zelinsky-Wibbelt (Editor).
- [Dauchy 93] P. Dauchy, C. Mignot and C. Valot. Joint speech and gesture analysis: some experimental results on multimodal interface. *EUROSPEECH, Berlin*, pages 1315–1318, 1993.
- [Denis 89a] M. Denis. Approches cognitives de l’image mentale. In *Intellectica*, volume 99, pages 85–107. 1989.
- [Denis 89b] M. Denis. *Image et Cognition*. Presses Universitaires de France, 1989.
- [Dony 85] R. Dony. *Graphisme Scientifique sur micro-ordinateur de la 2ème à la 3ème dimension*. Masson, 1985.
- [Dubois 91] D. Dubois. *Sémantique et cognition, Catégories, Prototypes et Typicalité*. Editions du CNRS, 1991.
- [Gaiffe 90] B. Gaiffe and L. Romary. Managing multimodal information and references in the multiworks project. Rapport Multiworks ESP 2105 et rapport CRIN 90R168, 1990.
- [Gaiffe 91] B. Gaiffe, J.M. Pierrel and L. Romary. Informations lexicales dans un dialogue homme-machine multimodal: application au traitement de la coréférence. In *Lexique et inférence(s) VII ème Colloque International de Linguistique*, pages 203–226. J.E Tyvaert, Metz, 1991.
- [Gaiffe 92a] B. Gaiffe. Référence et dialogue homme-machine: vers un modèle adapté au multi-modal, 1992. Thèse d’Université Henri Poincaré - Nancy I.
- [Gaiffe 92b] B. Gaiffe, L. Romary and J.M. Pierrel. De la référence à la modélisation de la tâche: vers une certaine idée du dialogue homme-machine, 1992. *Séminaire Dialogue des pôles parole et langage naturel du GDR-PRC Communication Homme-Machine*, Dourdan, 15-16 Avril 1992.

- [Gaiffe 93] B. Gaiffe and L. Romary. Ce matin, il pleuvait, ou la référence à l'épreuve du temps. In *Temps, référence et inférence*, pages 203–226. Revue Langages, Larousse, 1993.
- [Gaildrat 93] V. Gaildrat, N. Vigouroux, R. Caubet and G. Pérennou. Conception d'une interface multimodale pour un modèle déclaratif de scènes tridimensionnelles pour la synthèse d'images. In *L'Interface des mondes réels et virtuels*, pages 415–424, Montpellier, 22-26 Mars 1993. 2 ème Conférence Internationale.
- [Gardan 87] Y. Gardan. *Eléments de C.A.O*, volume 1. Hermès, 1987.
- [Garnham 89] A. Garnham. A unified theory of the meaning of some spatial relational terms. *Cognition* 31, pages 45–60, 1989.
- [Grice 75] H. P. Grice. Logic and conversation. *Syntax and Semantics, Speech acts, Cole and Morgan, Academic Press, New-York*, (3):41–58, 1975.
- [Habel 87] C. Habel. Cognitive linguistics: the processing of spatial concepts. *T.A informations, Bulletin semestriel de l'ATALA*, (1):21–56, 1987.
- [Hathout 94] N. Hathout. De la sémantique des verbes de déplacement et des prépositions spatiales vers leur syntaxe: cadres formels et théoriques. Manuscrit, 1994.
- [Hernández 94] D. Hernández. Qualitative representation of spatial knowledge. In *Lecture Notes in Artificial Intelligence 804*. Springer-Verlag, 1994.
- [Hottenroth 93] P.M. Hottenroth. *Prepositions and object concepts: A contribution to cognitive semantics*, pages 179–219. Mouton de Gruyter, 1993. Cornelia Zelinsky-Wibbelt (Editor).
- [Kaufmann 93] I. Kaufmann. *Semantic and conceptual aspects of the preposition durch*, pages 221–247. Mouton de Gruyter, 1993. Cornelia Zelinsky-Wibbelt (Editor).

-
- [Kayser 88] D. Kayser. Le raisonnement à profondeur variable. *Actes des journées du PRC IA Toulouse, Teknea*, pages 109–136, 1988.
- [Kleiber 81] G. Kleiber. *Problèmes de référence: descriptions définies et noms propres*. Klincksiek, 1981.
- [Kleiber 90] G. Kleiber. *La sémantique du prototype, catégories et sens lexical*. P.U.F, 1990.
- [Klein 93] W. Klein. L'expression de la spatialité dans le langage humain. In *Images et Langages. Multimodalité et Modélisation Cognitive*, pages 73–85. Colloque interdisciplinaire du CNRS, Paris, 1er et 2 Avril 1993.
- [Langacker 91] R. W. Langacker. *Concept, Image and Symbol; The Cognitive Basis of Grammar*. Mouton de Gruyter, 1991.
- [Laur 93] D. Laur. La relation entre le verbe et la préposition dans la sémantique du déplacement. In *La couleur des prépositions*, number 110. Larousse, Juin 1993.
- [Mignot 93] C. Mignot, C. Valot and N. Carbonell. An experimental study of future natural multimodal human-computer interaction. *INTERCHI, Amsterdam*, pages 67–68, 1993.
- [Milner 82] J. C. Milner. *Ordres et raisons de la langue*. Editions du Seuil, Paris, 1982.
- [Mukerjee 89] A. Mukerjee. A representation for modelling functional knowledge in geometric structures. In S. Ramani, R. Chandrasekar and K.S.R Anjaneyulu, editors, *Lecture Notes in Artificial Intelligence 444*, pages 192–202. Springer-Verlag, 1989.
- [Mukerjee 90] A. Mukerjee and G. Joe. A qualitative model for space. *AAAI 90 MIT Press Cambridge MA*, pages 721–727, 1990.
- [Nirenburg 87] S. Nirenburg and V. Raskin. Dealing with space in natural language processing, 1987. In proceedings of the Workshop on Spatial Reasoning and Multi-Sensor Fusion, eds A. Kak

- and S-S. Chen, 361-370, Los Altos, Calif. : Morgan Kaufmann.
- [Peroche 90] B. Peroche, J. Argence, D. Ghazanfarpour and D. Michelucci. *La synthèse d'images*. Hermès, 1990.
- [Poirier 93] F. Poirier and P. Lefebvre. Une interface multimodale de dialogue homme-machine sous unix. In *L'Interface des mondes réels et virtuels*, pages 161-170, Montpellier, 22-26 Mars 1993. 2^{ème} Conférence Internationale.
- [Pribbenow 93] S. Pribbenow. *Computing the meaning of localization expressions involving prepositions: the role of concepts and spatial context*, pages 441-470. Mouton de Gruyter, 1993. Cornelia Zelinsky-Wibbelt (Editor).
- [Rastier 87] F. Rastier. *Sémantique interprétative*. Formes sémiotiques P.U.F, 1987.
- [Rastier 91] F. Rastier. *Sémantique et recherches cognitives*. Formes sémiotiques P.U.F, 1991.
- [Retz-Schmidt 88] G. Retz-Schmidt. Various views on spatial prepositions. *A.I Magazine*, 9(2):95-105, 1988.
- [Romary 89] L. Romary. Vers la définition d'un modèle cognitif pour la représentation du temps dans un système de dialogue homme-machine, 1989. Thèse d'Université Henri Poincaré - Nancy I.
- [Romary 91] L. Romary. Integration of spatial and temporal information produced by a natural language discourse. In *Conference Knowledge Modelling and Expertise Transfert*, Sophia Antipolis, April 1991.
- [Romary 93] L. Romary. *Mets ça "ici" où quand "ici" dépend de ça. L'interprétation de "ici" dans les énoncés de positionnement*. In *Workshop : Le Dialogue Homme-Robot en Langage Naturel : Problèmes Psychologiques, Caen - France*, 1993.

- [Romary 94a] L. Romary, N. Bellalem and B. Gaiffe. Gestion de la référence dans un dialogue homme-machine. *Colloque du CNRS Images et Langages: Multimodalité et modélisation cognitive*, 1994.
- [Romary 94b] L. Romary, N. Bellalem and D. Schang. Positioning objects in a graphical environment: reference and gesture. *International Colloquium on Deixis*, 1994. Kentucky (USA).
- [Romary 95] L. Romary, N. Bellalem and D. Schang. A unified object representation for the interpretation of linguistic and gestual reference. *Workshop ESCA*, 1995. Danemark (soumis).
- [Rosch 78] E. Rosch. Principles of categorization. In E. Rosch and B. Loyd, editors, *Cognition and categorization*, pages 27–47. Hillsdale, (N.J.): L. Erlbaum, 1978.
- [Schang 94a] D. Schang. Modélisation de la référence et du positionnement spatial dans un système de dialogue homme-machine. *II èmes Rencontres Nationales des Jeunes Chercheurs en Intelligence Artificielle (RJCIA-94)*, 1994. Marseille.
- [Schang 94b] D. Schang and L. Romary. Frames, a unified model for the representation of reference and space in a man-machine dialogue. *3rd International Conference on Spoken Language Processing (ICSLP-94)*, 1994. Yokohama (Japon).
- [Schang 94c] D. Schang and L. Romary. Framing the world, towards a localised spatial reasoning. *3rd International Conference on the Cognitive Science of Natural Language Processing (CSNLP-94)*, 1994. Dublin (Irlande).
- [Schirra 93a] J. Schirra. *A Contribution to Reference Semantics of Spatial Prepositions: The Visualization Problem and its Solution in VITRA*, pages 471–515. Mouton de Gruyter, 1993. Cornelia Zelinsky-Wibbelt (Editor).
- [Schirra 93b] J. Schirra and E. Stopp. Antlima - a listener model with mental images. *IJCAI*, pages 175–180, 1993.

- [Sidner 86] C. L. Sidner. *Focussing in the comprehension of definite anaphora*. Readings in natural language processing. Morgan Kauffman Publisher, 1986. Edité par B. Grosz, K. S. Jones, B. L. Webber.
- [Sondheimer 76] N. K. Sondheimer. Spatial reference and natural-language machine control. *International Journal of Man-Machine Studies*, (8):329–336, 1976.
- [Souvay 92] G. Souvay. Diapason : un environnement de développement pour l'intégration d'une entrée vocale dans des applications de type commande de machine, 1992. Thèse d'Université Henri Poincaré - Nancy I.
- [Sperber 89] D. Sperber and D. Wilson. *La pertinence. Communication et Cognition*. Editions de Minuit, 1989.
- [Talmy 83] L. Talmy. *How Language Structures Space*. H. Pick & L. Acredolo editeurs, Plenum, New York, 1983.
- [Vandeloise 86] C. Vandeloise. *L'espace en français*. Editions du Seuil, Paris, 1986.
- [Vieu 91] L. Vieu. Sémantique des relations spatiales et inférences spatio-temporelles : une contribution à l'étude des structures formelles de l'espace en langage naturel, 1991. Thèse d'Université Paul Sabatier Toulouse.
- [Wahlster 87] W. Wahlster. One word says more than a thousand pictures, on the automatic verbalization of the results of image sequence analysis system. *T.A informations, Bulletin semestriel de l'ATALA*, (2):57–70, 1987.
- [Wazinski 91] P. Wazinski. Generating spatial descriptions for cross-modal references. Technical Report TM-91-11, DFKI, Saarbrücken, 1991.
- [Wazinski 93] P. Wazinski. Graduated topological relations. Memo 54, SFB 314, Universität des Saarlandes - Saarbrücken, 1993.



Unité de recherche INRIA Lorraine, Technopôle de Nancy-Brabois, Campus scientifique,
615 rue du Jardin Botanique, BP 101, 54600 VILLERS LÈS NANCY
Unité de recherche INRIA Rennes, Irisa, Campus universitaire de Beaulieu, 35042 RENNES Cedex
Unité de recherche INRIA Rhône-Alpes, 46 avenue Félix Viallet, 38031 GRENOBLE Cedex 1
Unité de recherche INRIA Rocquencourt, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex
Unité de recherche INRIA Sophia-Antipolis, 2004 route des Lucioles, BP 93, 06902 SOPHIA-ANTIPOLIS Cedex

Éditeur
INRIA, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex (France)
ISSN 0249-6399