



Université Joseph Fourier

U.F.R. Informatique &  
Mathématiques Appliquées



Institut National Polytechnique de  
Grenoble

ENSIMAG

**I.M.A.G.**

**DEA D'INFORMATIQUE :  
SYSTEMES ET COMMUNICATIONS**

Projet présenté par :

Frédéric VERNIER

*Usage de la multimodalité pour rendre perceptible  
un espace d'informations de grande taille*

Effectué au laboratoire :

CLIPS-IMAG,  
Communication Langagière et Interaction Personne-Système  
Equipe Ingénierie de l'Interaction Homme-Machine (IIHM).

Date : 10/06/97

Jury : Didier Bert,  
Sacha Krakowiak  
Laurence Nigay,  
Nathalie Portolan



## Remerciements

Je tiens à remercier tous les membres de l'équipe IHM du laboratoire CLIPS. Je remercie tout particulièrement Joëlle Coutaz et Laurence Nigay qui m'ont témoigné leur confiance tout au long de mon stage de DEA. Je les remercie notamment pour le matériel de haut niveau qu'elles ont mis à ma disposition et pour ma participation à la conférence CHI.

Je remercie également toutes les personnes qui m'ont aidé dans la programmation de VITESSE : Fabrice Barial, Francis Jambon, François Bérard et Laurent Aublet-Cuvelier.

VITESSE n'est pas qu'un morceau de code. C'est avant tout des idées émises dans les réunions de l'équipe, sur un coin de bureau, devant un comptoir de cafétéria ou sur une table de restaurant. Un grand nombre de personnes m'ont apporté leurs soutiens et m'ont accordé de leur temps pour confronter leurs idées aux miennes : Jim Crowley, Daniel Salber, Nathalie Portolan, Philippe Mulhem et d'autres encore.

Je remercie sincèrement toutes ces personnes ainsi que la communauté « IHM » qui m'ont donné le goût dévorant de la recherche.



# Table des matières

<b>INTRODUCTION .....</b>	<b>3</b>
<b>1. Sujet .....</b>	<b>4</b>
<b>2. Motivations : en quête d'information... ..</b>	<b>4</b>
<b>3. Contexte de recherche .....</b>	<b>5</b>
<b>4. Plan du rapport.....</b>	<b>5</b>
<b>MULTIMODALITÉ .....</b>	<b>7</b>
<b>1. Définition d'une modalité.....</b>	<b>8</b>
<b>2. Utilité des classifications.....</b>	<b>9</b>
2.1. Caractérisation des modalités .....	9
2.2. Aide à la conception .....	10
2.3. Evaluation des interfaces existantes .....	10
2.4. Conclusion .....	10
<b>3. Orientation de la multimodalité .....</b>	<b>10</b>
3.1. Modalité en entrée.....	11
3.2. Modalité en sortie .....	11
3.3. Taxonomie M2LD .....	12
<b>4. Taxonomies existantes des systèmes multimodaux.....</b>	<b>12</b>
4.1. Classification des tâches graphiques de J. Foley.....	13
4.2. Espace de conception de frohlich .....	13
4.3. Taxonomie de Bernsen.....	16
4.4. Propriétés CARE .....	17
4.4.1. Assignation.....	18
4.4.2. Equivalence .....	18
4.4.3. Redondance .....	18
4.4.4. Complémentarité .....	18
4.4.5. Conclusion sur les propriétés CARE .....	19
<b>5. Critères relatifs à la multimodalité en sortie.....</b>	<b>19</b>
<b>6. Conclusion .....</b>	<b>21</b>
<b>GRANDE QUANTITÉ D'INFORMATIONS ET MULTIMODALITÉ .....</b>	<b>23</b>
<b>1. Grandes quantités d'information.....</b>	<b>24</b>
<b>2. Propriétés d'ergonomie .....</b>	<b>25</b>
2.1. Propriétés de souplesse .....	26
2.2. Propriétés de robustesse .....	27
<b>3. Propriétés de sélection relatifs aux grands espaces d'informations .....</b>	<b>27</b>
3.1. Propriétés d'ergonomie relatifs aux grands espaces d'informations.....	27
3.2. Caractérisation des modalités de sortie.....	30

3.2.1.	Caractérisation des dispositifs de sortie .....	30
3.2.2.	Caractérisation des systèmes représentationnels .....	31
3.2.3.	Combinaisons de modalités .....	32
<b>4.</b>	<b>Exemples d'application de nos caractérisations .....</b>	<b>33</b>
4.1.	Lentilles magiques .....	34
4.2.	Système « Drag Mag » .....	35
4.3.	Mur en perspective .....	37
4.4.	Vue en œil de poisson .....	38
4.5.	Technique de « zoom » .....	39
4.6.	Arbre conique .....	40
<b>5.</b>	<b>Représentation multiple .....</b>	<b>41</b>
5.1.	Principes de la représentation multiple .....	41
5.2.	Règles de conception .....	42
<b>6.</b>	<b>Conclusion .....</b>	<b>43</b>
<b>SYSTÈME VITESSE .....</b>		<b>45</b>
<b>1.</b>	<b>Conception de l'interface .....</b>	<b>46</b>
1.1.	Systèmes de représentation .....	46
1.1.1.	Technique de référence .....	47
1.1.2.	Techniques déformantes .....	48
1.1.3.	Passage d'une représentation à une autre .....	52
1.2.	Outils de navigation .....	53
1.2.1.	Manipulation directe et indirecte de l'espace d'information .....	53
1.2.2.	Objets d'interaction supplémentaires .....	54
1.2.3.	Objets pour paramétrer la technique de visualisation .....	55
1.3.	Propriétés vérifiées .....	56
<b>2.</b>	<b>Réalisation logicielle .....</b>	<b>56</b>
2.1.	Architecture logicielle globale .....	56
2.1.1.	Architecture globale .....	57
2.1.2.	Réalisation des Techniques d'Interaction : le contrôleur de dialogue .....	58
2.2.	JAVA et modèle d'architecture .....	59
2.2.1.	Architecture générale et JAVA .....	60
2.2.2.	Architecture du contrôleur de dialogue et AWT .....	61
<b>3.</b>	<b>Evaluation et améliorations futures .....</b>	<b>62</b>
<b>CONCLUSION : BILAN ET PERSPECTIVES .....</b>		<b>65</b>
<b>1.</b>	<b>Bilan et contribution .....</b>	<b>66</b>
<b>2.</b>	<b>Perspectives .....</b>	<b>67</b>
2.1.	Perspectives à court terme .....	67
2.2.	Perspectives à long terme .....	67

# **Chapitre I**

## *Introduction*

## 1. Sujet

Notre étude de D.E.A. s'intitule «Usage de la multimodalité pour rendre perceptible un espace d'informations de grande taille». Nous n'expliquons pas dans le chapitre d'introduction ce qu'est une modalité car ceci fait l'objet du chapitre II. Nous expliquons ici le sujet de l'étude et ses motivations.

L'objectif central à l'étude est le rendu et la manipulation d'un grand espace d'informations qui soient adaptés aux utilisateurs. Ces mines d'informations peuvent être des résultats à une requête (base de données) ou bien une arborescence d'un disque dur ou encore le World Wide Web (WWW), l'hypertexte multimédia d'Internet. Ces grandes quantités d'informations posent de nombreux problèmes en terme d'interface homme-machine. Il est en effet impossible d'afficher la totalité de l'information sur un écran, ce dernier étant de taille limitée. Pour rendre un tel espace perceptible à l'utilisateur, il convient de mettre en œuvre des techniques d'interaction (souvent graphiques) adaptées au type de l'information et bien sûr à l'utilisateur et à sa tâche. Ces techniques de visualisation et plus généralement d'interaction mettent en œuvre des fonctions de déformation, des vues multiples ou encore des filtres.

Indépendamment de cet objectif, nous avons vu émerger puis se multiplier de plus en plus de nouveaux dispositifs physiques permettant un dialogue riche entre l'utilisateur et la machine (caméra, microphone, écran haute-résolution, haut-parleur, etc.). L'interaction entre l'utilisateur et l'ordinateur a évolué vers de nouvelles formes appelées multimodales. Le but de notre travail de recherche consiste à étudier dans quelle mesure ces nouvelles formes d'interaction permettent de mieux rendre perceptible et manipulable une grande quantité d'informations à l'utilisateur.

## 2. Motivations : en quête d'information...

Les récentes avancées des technologies de la communication ouvrent de nouvelles perspectives. En particulier, les autoroutes de l'information permettent une diffusion instantanée d'une grande quantité d'informations. Mais cette abondance à l'échelle planétaire nourrit en même temps le risque de détruire l'information : cette dernière existe mais elle devient inaccessible. Le problème n'est plus tant la disponibilité de l'information mais la capacité de sélection d'une information qui réponde au besoin. Les travaux récents sur la télévision interactive illustrent parfaitement le problème : une grande quantité d'informations doit être accessible à tout public et tenir compte de la variété des conditions d'usage à domicile. A l'évidence, il convient d'assister les utilisateurs dans leur quête d'information pour des situations chaque fois particulières.

Pour assister l'utilisateur dans sa quête d'information, plusieurs approches sont envisageables : améliorer les moteurs de recherche, déléguer certaines tâches de recherche à des agents logiciels, permettre l'expression fine des requêtes en langue naturelle augmentée de dialogues de clarification, améliorer la qualité perceptuelle et



interactionnelle des réponses des moteurs de recherche, faire appel au savoir collectif. A l'évidence, toutes ces pistes sont valides et fonctionnent de manière complémentaire. Dans notre travail, nous adoptons une approche centrée sur l'interaction et nous visons à concevoir de nouvelles formes d'interaction multimodale qui aident l'utilisateur à appréhender un grand espace d'informations et à retrouver une information particulière.

### **3. Contexte de recherche**

La réalisation logicielle repose sur le travail de magistère seconde année [Vernier 96] que j'ai effectué pendant l'été 1996 dans l'équipe d'Ingénierie de l'Interaction Homme-Machine (IIHM) du laboratoire C.L.I.P.S. Ce stage m'a permis de programmer les bases du système VITESSE. Cet année, j'ai étendu VITESSE pour mettre en pratique les résultats présentés dans ce rapport, en le rendant multimodal.

Notre travail a fait l'objet d'une publication [Vernier 97] et le système VITESSE de deux publications [Nigay 97a][Nigay 97b]. L'ensemble de ce travail s'insère dans un contrat de recherche entre le CNET et notre équipe de recherche.

### **4. Plan du rapport**

Après avoir introduit la définition d'une modalité au chapitre II, nous présentons les principales taxonomies existantes des systèmes multimodaux. Une synthèse permet de mettre en évidence l'intérêt de ces classifications dans le cadre des interfaces en sortie (i.e. du système vers l'utilisateur). Le chapitre III présente notre application de la multimodalité au rendu d'un grand espace d'informations : la représentation multiple de l'espace selon des modalités différentes. Le chapitre IV est consacré au système VITESSE, ma réalisation logicielle, qui constitue une application pratique de mes travaux de recherche. Nous présentons la conception de l'interface et nos règles de conception puis sa réalisation logicielle selon un modèle d'architecture. Le chapitre V constitue une conclusion à nos travaux de D.E.A. où nous soulignons les points contributifs et nous développons les perspectives.



## **Chapitre II**

### *Multimodalité*

---

Ce chapitre est consacré à la multimodalité. Nous commençons ce chapitre par une définition d'une modalité afin de fixer la terminologie employée dans la suite. Avant de présenter les taxonomies dédiées à la multimodalité, nous expliquons l'intérêt d'une taxonomie pour la conception et l'évaluation des interfaces au paragraphe 3. A partir des taxonomies présentées au paragraphe 4, nous identifions au paragraphe 5 des caractéristiques des modalités de sortie. Première étape de travail dans notre approche, ces caractéristiques sont utiles pour pouvoir choisir les modalités de sortie qui sont les plus adaptées à la tâche de l'utilisateur et à son contexte.

## 1. Définition d'une modalité

La multimodalité représente le plus souvent, pour les informaticiens, la capacité d'un système interactif à utiliser plusieurs canaux de communication lors de l'interaction entre l'utilisateur et le système. Malheureusement, la communauté n'est pas d'accord sur la définition d'un canal de communication ou d'une modalité. Le flou qui règne autour de la notion de modalité m'incite à définir le terme modalité, afin de fixer la terminologie employée dans ce rapport.

Puisqu'il n'existe pas de définition consensuelle d'une modalité, il n'existe pas non plus de définition de l'interaction multimodale puisque ce dernier exprime le fait d'avoir plusieurs modalités mises en œuvre dans un même processus d'interaction entre l'homme et l'ordinateur. Toutefois on comprend intuitivement qu'une modalité est un concept qui vient se situer entre celui de sens et celui de média. Le terme « sens » est utilisé ici comme un organe de perception (l'ouïe, le toucher, la vue, l'odorat ou le goût). Le sens de perception est un bon critère pour l'homme mais peu adapté à l'ordinateur. Un média est un support d'informations (ce rapport de D.E.A., un CD-ROM ou encore une disquette). C'est un terme utile pour mesurer l'étendue des périphériques d'un système informatique mais il est assez mal adapté à l'homme qui lui interprète ce qu'il perçoit. Par exemple, deux dispositifs dédiés à des médias différents peuvent apparaître comme identiques à certains utilisateurs néophytes. De plus, un média est mal adapté à l'interaction homme-machine car c'est un terme descriptif qui ne rend pas compte des échanges d'information.

Nous choisissons de définir une modalité à un niveau d'abstraction intermédiaire entre le sens et le média. Néanmoins nous voulons garder un point de vue proche du système. Pour définir une modalité, nous nous basons sur le concept de dispositif physique que nous raffinons avec le langage d'interaction qu'il véhicule : la parole est un son caractérisé par le langage naturel alors que la tonalité du téléphone est un son selon un autre langage d'interaction. La notion de langage est très discutable lorsqu'elle désigne qu'un ensemble de signes. C'est pourquoi nous utiliserons le terme plus général de système représentationnel. Ainsi le couple <d,s> où d est le dispositif physique et s le système représentationnel, désigne une modalité.

Le positionnement d'une modalité ainsi définie dans le contexte de l'interaction homme-machine permet de distinguer deux types de modalité : celles en entrée et

---

celles en sortie. Les modalités d'entrée concernent les échanges d'information de l'utilisateur vers le système tandis qu'une modalité de sortie est utilisée par le système pour des retours d'information à destination de l'utilisateur. La multimodalité en entrée et en sortie soulèvent des problèmes différents. En effet le système met à disposition de l'utilisateur un ensemble de modalités d'entrée qu'il peut utiliser. Le choix des modalités d'entrée est laissé à l'utilisateur qui selon son contexte et sa tâche choisira celles qui sont les plus adaptées à la situation. Au contraire dans le cas de la multimodalité en sortie, le choix des modalités est laissé au système. Deux types de systèmes multimodaux en sortie sont alors à distinguer :

- ? les systèmes multimodaux où l'usage des modalités en sortie est programmé et préétabli par le concepteur,
- ? les systèmes multimodaux qui ont la capacité de dynamiquement choisir les modalités de sortie les plus adéquates au contexte d'interaction courant.

Dans notre étude et quel que soit le système multimodal visé, la première étape de travail consiste à identifier des caractéristiques des modalités de sortie afin de pouvoir ultérieurement les distinguer et donc les choisir. Dans cet objectif, nous étudions les taxonomies existantes. Avant de présenter ces taxonomies aux paragraphes 3 et 4, nous rappelons au paragraphe suivant l'utilité de telles taxonomies.

## **2. Utilité des classifications**

Les classifications des systèmes multimodaux sont de deux types : celles qui étudient chaque modalité séparément et celles qui visent à organiser et classer les compositions de modalités. D'une part, une classification des modalités permet de clarifier l'espace de conception d'une interface et de guider le choix des modalités à utiliser. D'autre part, l'étude des différentes manières de composer les modalités permet d'identifier la compatibilité entre les modalités.

Mais avant de présenter les classifications existantes, nous tenons à rapporter l'utilité de telles classifications. En effet le cas des grands espaces d'informations n'est qu'un cas particulier que nous considérons dans l'étude. Il nous semble intéressant de bien comprendre la portée des classifications avant de les exploiter dans le cas des grandes quantités d'informations. Les trois principaux intérêts des taxonomies sont :

- ? la caractérisation des modalités et de leurs compositions,
- ? l'aide à la conception,
- ? l'évaluation des interfaces existantes.

### **2.1. Caractérisation des modalités**

Etablir une taxonomie des interfaces multimodales, consiste à concevoir des espaces multidimensionnels dans lesquels interviennent des propriétés inhérentes à la

multimodalité. Il est alors possible d'en dégager des propriétés et de les inclure dans le cahier des charges d'un logiciel, par exemple sous la forme de critères de qualité à vérifier. Un autre aspect des classifications d'interfaces multimodales est la possibilité de pondérer les caractéristiques multimodales d'une interface [Nigay 93]. En effet le tout ou rien n'est pas de mise dans l'élaboration d'une interface et il convient la plupart du temps de nuancer les caractéristiques d'une interface multimodale. Donner un poids aux caractéristiques multimodales permet de mesurer dans quelle proportion une interaction a telle ou telle propriété multimodale.

## **2.2. Aide à la conception**

Les classifications identifient des classes de systèmes qui peuvent être utiles à la conception. Ces classes sont souvent positionnées par rapport à des propriétés ou des tâches. Sachant les propriétés à atteindre et les tâches à réaliser, le concepteur peut alors s'orienter vers la conception d'une interface d'un type particulier. Plus généralement les classifications permettent au concepteur de se poser les bonnes questions de conception en connaissant l'étendue des possibilités offertes par la multimodalité.

Dans le cas de la conception d'interfaces multimodales en sortie, le choix des modalités à utiliser revient aux concepteurs. Ce choix doit prendre en compte les préférences de l'utilisateur, la tâche à réaliser et les possibilités matérielles [Gargan 88]. Les axes identifiés par une taxonomie peuvent donc aider le concepteur à choisir les modalités les plus adaptées.

## **2.3. Evaluation des interfaces existantes**

L'évaluation ergonomique d'une interface est une étape incontournable dans le cycle de vie d'un système interactif. Les taxonomies définissent des critères pour l'évaluation. De plus l'évaluation permet la comparaison notamment entre deux produits concurrents. Lorsqu'une interface obtient un large succès auprès du public, il s'avère intéressant de savoir ce qui a plu. Il peut s'agir d'une nouvelle modalité jamais exploitée ou d'un couple de modalités très efficace pour une tâche précise. L'enjeu ici est de pouvoir distinguer les systèmes et donc de les classer.

## **2.4. Conclusion**

Nous constatons que les taxonomies sont utiles tout au long d'un cycle de vie d'un système. La classification permet de mieux spécifier l'interface en amont de la réalisation. De plus lors de la conception, les axes des taxonomies constituent un support à la décision. Enfin lors de l'évaluation, les taxonomies servent de guide pour fixer les critères d'évaluation mais aussi permettent de distinguer les systèmes et donc de les comparer.

## **3. Orientation de la multimodalité**

---

Avant de présenter les taxonomies, il est important de mieux expliquer le sens d'une modalité. Comme nous l'avons introduit au paragraphe 1, nous dissocions la multimodalité en entrée de la multimodalité en sortie pour plusieurs raisons :

- ? Une modalité est une technique d'interaction qui est habituellement orientée en sortie ou en entrée.
- ? Les sens de perception humains diffèrent beaucoup de ceux des ordinateurs. De même, les actions que nous entreprenons pour modifier le monde qui nous entoure sont très différentes de celles que peuvent réaliser les ordinateurs.

Le concepteur n'a donc pas la même démarche pour définir les modalités d'entrée qu'il propose à l'utilisateur et les modalités de sortie pour lesquelles il doit faire un choix. De plus les préférences du concepteur ne sont pas toujours celles des utilisateurs finaux. Il est donc légitime pour le concepteur de penser différemment pour l'interaction en entrée et en sortie. Il est cependant important de comprendre cette distinction à des fins de conception. En effet il est clair que les modalités utilisées en entrée peuvent influencer celles en sortie et vice-versa.

### **3.1. Modalité en entrée**

De nombreuses possibilités sont offertes pour la multimodalité en entrée. Chaque jour de nouvelles sortes de capteurs sont inventées tels que les capteurs infrarouges, magnétiques ou encore de température. Cet état de fait laisse un grand éventail de choix. La communication de l'homme vers l'ordinateur se compose essentiellement d'actions que l'ordinateur doit effectuer. L'enjeu principal est de traiter le plus rapidement l'expression d'entrée afin de garantir un temps de réponse acceptable.

### **3.2. Modalité en sortie**

Les modalités de sortie offrent un choix plus restreint car c'est l'homme qui joue le rôle de récepteur (vision, ouïe, toucher). L'odorat et le goût sont, à l'heure actuelle, trop difficiles à mettre en œuvre dans un système informatique. Contrairement à la communication en entrée, l'ordinateur a souvent besoin de communiquer de grandes quantités d'informations à l'utilisateur. La vision est pour cette raison le sens le plus utilisé (texte écrit, icônes symboliques, graphiques, etc.). L'ouïe est un sens en pleine expansion en informatique depuis l'apparition des ordinateurs familiaux dits « multimédias » qui sont équipés de matériel d'acquisition et de restitution de sons. L'audition offre des possibilités différentes de la vision : par exemple prévenir l'utilisateur de l'arrivée d'un message quand ce dernier ne regarde plus son écran. Le troisième sens, et le moins employé, est le toucher qui regroupe un certain nombre de dispositifs novateurs à retour d'effort [Cadoz 94] : les manches à balai de certains simulateurs d'avion reproduisent les secousses dues aux turbulences ou encore les bras manipulateurs peuvent être découplés de la réalité pour servir à manipuler des molécules fictives et permettre de « sentir » le contour des atomes qui la constituent.

### 3.3. Taxonomie M2LD

Après avoir présenté l'orientation des modalités (modalités en entrée et en sortie) comme un critère simple pour classer les modalités, nous présentons un premier espace de conception qui présente des critères élémentaires de classification : M2LD. En utilisant la définition de la modalité, basée sur le couple <dispositif, système représentationnel>, Laurence Nigay [Nigay 94] propose trois différentes classes de systèmes multimodaux en différenciant ceux qui utilisent plusieurs langages, ceux qui utilisent plusieurs dispositifs et enfin ceux qui utilisent plusieurs langages et plusieurs dispositifs. L'auteur utilise ces classes pour caractériser les interfaces en entrée et en sortie.

<i>Pour un système donné, nous différencions :</i>	si le système représentationnel est le même pour chaque modalité	si le système représentationnel est différent pour au moins deux modalités
si les dispositifs physiques sont identiques	<b>Mono Lang, Mono Disp</b> Ce n'est pas une interaction multimodale	<b>Multi Lang, Mono Disp</b> C'est une multimodalité partielle polymorphe
si les dispositifs physiques sont différents	<b>Mono Lang, Multi Disp</b> C'est une multimodalité partielle répartie	<b>Multi Lang, Multi Disp</b> C'est une multimodalité totale

Outre un classement simple des systèmes multimodaux, M2LD identifie des classes de système qui n'offrent pas les mêmes possibilités et avantages à l'utilisateur.

Les systèmes « Mono Disp, Multi Lang » sont intéressants pour communiquer une information compliquée. L'utilisation d'un seul dispositif permet à l'utilisateur de mieux appréhender les liens entre les modalités (complémentarité ou redondance éventuelles).

Les systèmes « Multi Disp Mono Lang » permettent entre autre de prévoir la panne d'un dispositif. De plus ce type de système peut s'adapter aux utilisateurs en choisissant le bon dispositif. Nous pensons en particulier aux personnes handicapés.

Les systèmes « Multi Disp Multi Lang » possèdent les avantages de ceux des deux classes précédentes à condition que certaines modalités soient équivalentes et partagent le même dispositif. De plus de tels systèmes définissent une plus riche interaction entre l'utilisateur et le système qui est adaptée à la tâche et au contexte courant.

## 4. Taxonomies existantes des systèmes multimodaux



---

Les classifications des systèmes multimodaux en entrée sont assez nombreuses. Il est clair que la communauté a beaucoup plus travaillé sur la multimodalité en entrée. Nous trouvons tout d'abord les taxonomies des dispositifs physiques comme celle de Buxton [Buxton 83] ou encore celle de Card [Card 90]. La taxonomie de Foley [Foley 84] classe aussi les dispositifs non plus en terme d'attributs physiques mais selon les tâches qu'ils peuvent réaliser. Nous la présentons au paragraphe 4.1. Nous exposons ensuite la taxonomie de Frohlich [Frohlich 91] car elle intègre les interfaces en entrée et en sortie. Nous présentons au paragraphe suivant la seule taxonomie dédiée aux interfaces en sortie, celle de Bernsen [Bernsen 93]. Enfin nous terminons ce paragraphe en expliquant les propriétés CARE [Martin 94] qui à notre connaissance constitue le seul espace qui caractérise les relations entre modalités.

#### **4.1. Classification des tâches graphiques de J. Foley**

La classification de Foley [Foley 84] se limite aux dispositifs d'entrée qui permettent d'accomplir une tâche graphique. Elle consiste à classer les dispositifs selon les tâches qu'ils permettent de réaliser. Grâce à cette taxonomie, le concepteur peut choisir les dispositifs les plus adaptés à la tâche graphique identifiée. L'intérêt de la taxonomie réside aussi dans la liste des tâches générales identifiées. Cinq tâches graphiques sont présentées : les tâches de sélection, de localisation, d'orientation, de quantification et de texte.

Bien que cette classification vise à classer les dispositifs, elle peut s'appliquer aux modalités en considérant notre définition d'une modalité comme le couple  $\langle d, s \rangle$  où  $d$  est le dispositif physique et  $s$  le système représentationnel.

Pour étendre la taxonomie au cas de la multimodalité en sortie, il conviendrait d'identifier un ensemble de tâches générales que le système doit effectuer. Nous limiterons le problème à l'ensemble des tâches dans le cas du rendu d'une grande quantité d'informations.

#### **4.2. Espace de conception de frohlich**

L'espace de conception de Frohlich [Frohlich 91] englobe aussi bien les modalités d'entrée que de sortie. En effet, cette taxonomie s'intéresse aux deux orientations possibles des modalités. L'orientation fait l'objet du premier critère de différenciation entre les modalités. Cela amène D. Frohlich à proposer deux espaces de conception, présentés aux Figures 1 et 2, reposant sur les mêmes concepts (modes, canaux, médias et style).

Les deux espaces de conception différencient les interfaces langagières des interfaces à manipulation directe respectivement sous les termes de « mode langagier » et de « mode action ». Les interfaces à manipulation directe permettent à l'utilisateur de manipuler les objets informatiques comme ceux du monde réel (déplacer ou toucher). Comme dans le cas de la célèbre métaphore du bureau électronique inventée

par Rank Xerox et reprise par Apple, les objets sont directement manipulables au travers de dispositifs tels que la souris ou un manche à balai (« joystick » en anglais). A l'opposé, le mode langagier permet à l'utilisateur de communiquer avec le système au travers d'un langage. L'utilisateur fournit alors au système des commandes dans un langage donné.

D. Frohlich définit ensuite les canaux de communication de l'utilisateur et du système. Il observe ainsi que l'homme dispose de la voix et de mouvements pour établir une communication en entrée et de l'ouïe, la vue et le toucher pour une communication en sortie (du système vers l'utilisateur). Les systèmes informatiques actuels utilisent en entrée comme en sortie trois types de canaux de communication : le canal auditif, le canal visuel et le canal haptique.

Enfin deux autres niveaux sont distingués dans la taxonomie : un média et un style. Le média est un système représentationnel qui permet l'échange d'informations : la parole, le texte et le geste pour le mode langage et le son, le graphique et le déplacement pour le mode action. Le style d'interface est une classe de méthodes qui rend possible l'interaction. L'auteur distingue des styles différents pour chacun des modes. Le mode action possède trois styles : les environnements à fenêtres, les interfaces iconiques qui s'appuient sur une métaphore du monde réel et les formalismes graphiques (non-métaphoriques). Les styles du mode langage sont au nombre de cinq : les langages de commande, les langages de programmation (avec définition de procédure), les langues naturelles, les formulaires et les menus.

Cette taxonomie a l'avantage de traiter l'interface en sortie et en entrée. Elle identifie cinq niveaux d'abstraction à considérer dans la définition d'une modalité. De plus elle met explicitement en relation les canaux humains et ceux du système.

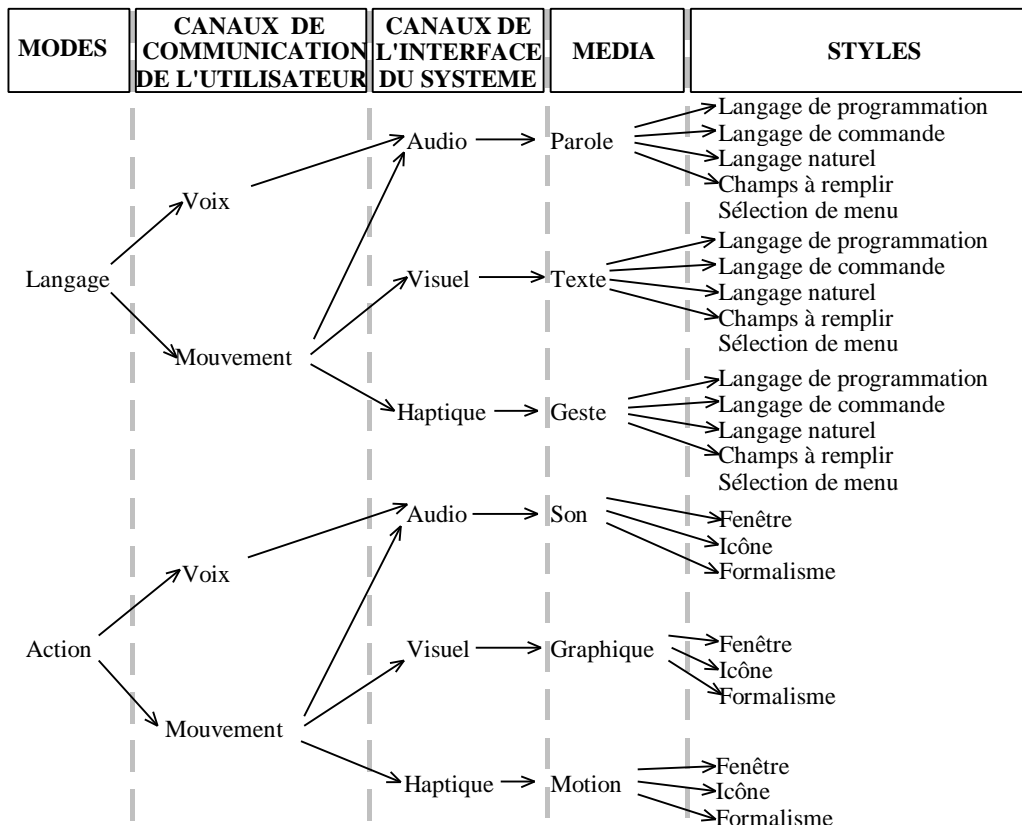


Figure 1 : Taxonomie des interfaces en entrée par Frohlich

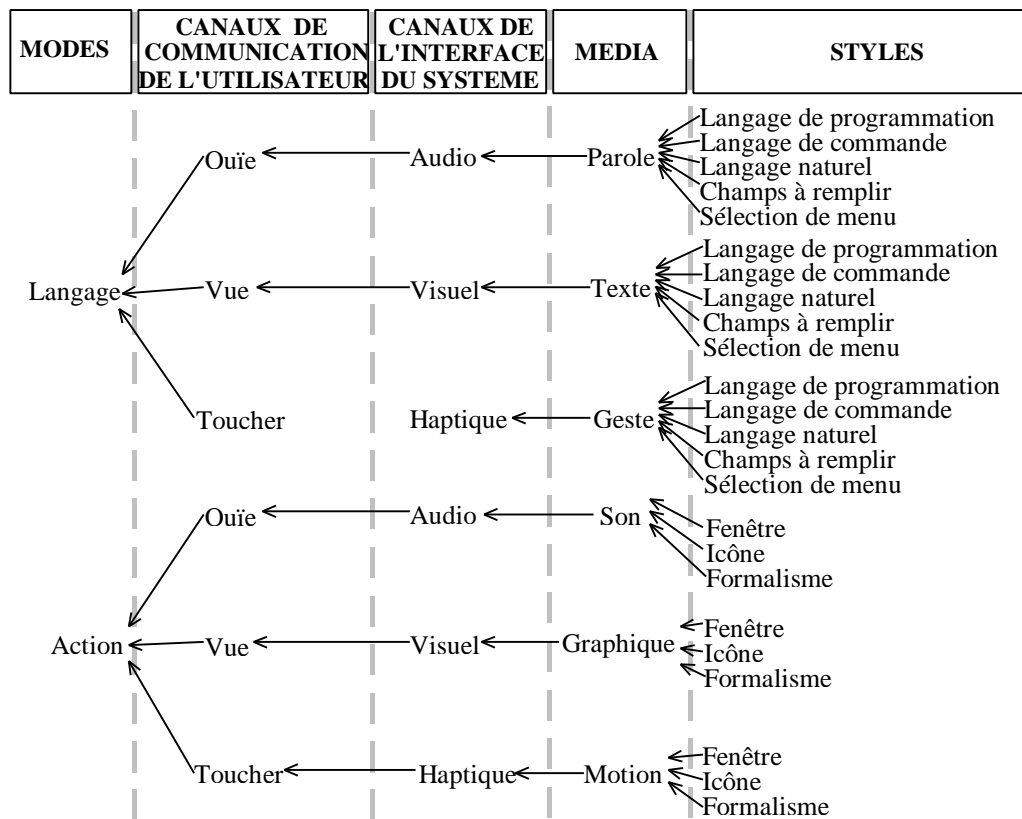


Figure 2 : Taxonomie des interfaces de sortie par Frohlich

### 4.3. Taxonomie de Bernsen

La taxonomie établie par Bernsen [Bernsen 93] est la seule qui est dédiée aux interfaces multimodales en sortie. Cette taxonomie vise à classer les modalités de sortie. Pour cela l'auteur identifie des caractéristiques des modalités. Ainsi dans son espace, une modalité est définie par deux traits : le média et le profil.

Le média est le sens humain intervenant dans la communication. Les valeurs possibles dans le cas d'un système informatique sont au nombre de trois :

- ? le son,
- ? le graphique,
- ? le toucher.

Le profil est un ensemble de quatre paramètres booléens :

- ? statique/dynamique,
- ? linguistique/non linguistique,
- ? analogique/non analogique,
- ? arbitraire/non arbitraire.

La caractéristique "statique/dynamique" repose sur la présence ou non d'une dimension temporelle dans la communication. Nous notons que la dimension spatiale souvent liée à la dimension temporelle est absente de la taxonomie de Bernsen. Pourtant les modalités liées à un écran peuvent être mono, bi ou tridimensionnelles. De plus des recherches sont menées sur la spatialisation du son avec des dispositifs sonores dédiés. Enfin les dispositifs à retour d'effort se caractérisent aussi par un certain nombre de degrés de libertés [Cadoz 94] que nous pouvons décrire comme autant de dimensions spatiales.

Une modalité "linguistique" est définie par un système structuré de symboles remplissant une fonction de communication. Ce système définit le vocabulaire du langage. Bernsen ne fait aucune distinction de complexité entre les différents langages possibles. Une telle distinction trouverait son utilité dans l'évaluation du temps d'apprentissage d'une interface.

La caractéristique "analogique" repose sur la notion de ressemblance avec la réalité. Cette caractéristique est à rapprocher du concept de métaphore du monde réel très utilisé en IHM. La métaphore du monde réel est cependant difficile à appliquer car l'objectif est de s'appuyer sur des concepts connus par l'utilisateur dans sa vie quotidienne. Cela dépend donc fortement des utilisateurs. Il nous semble donc difficile d'affirmer sans considération des utilisateurs qu'une modalité est analogique ou non.

Enfin Bernsen propose la caractéristique "arbitraire" pour différencier les modalités. L'arbitraire fait référence à l'absence de lien entre le contenu et le

---

contenant. Une modalité non-arbitraire repose donc sur une fonction de correspondance qui n'a pas été défini par le concepteur. Le risque d'utiliser une modalité arbitraire est que l'utilisateur ne soit pas capable de l'interpréter.

Grâce à ces cinq caractéristiques fondamentales, 48 classes de modalités sont définies. Vingt d'entre elles sont à exclure car elles correspondent à des incohérences dues au fait par exemple qu'une modalité linguistique ou analogique ne peut pas être arbitraire. Bernsen met ainsi en évidence 28 modalités pures de sortie.

Après avoir présenté la taxonomie de Bernsen qui identifie des caractéristiques pour décrire les modalités prises isolément, il convient maintenant d'étudier les propriétés CARE qui définissent des compositions entre modalités.

#### 4.4. Propriétés CARE

Les propriétés CARE sont au nombre de quatre et définissent les liens entre les modalités dans le contexte d'une interaction multimodale. C, A, R et E sont les initiales des quatre propriétés identifiées par [Martin 94] : la complémentarité, l'assignation, la redondance et l'équivalence. Ces propriétés sont appliquées à l'usage des modalités en entrée.

La définition de ces propriétés repose sur les notions d'état, de but, de modalité et de relation temporelle [Coutaz 95].

- ? Un état est un vecteur de variables observables, un ensemble de propriétés mesurables à un instant donné qui caractérise une situation.
- ? Un but est un état qu'un agent veut atteindre. Un agent peut être l'utilisateur, le système ou seulement un composant du système. Généralement, un agent est une entité capable d'initialiser le déroulement d'une action.
- ? Une modalité est une méthode d'interaction utilisée par un agent pour atteindre un but. Afin d'estimer le pouvoir d'expression d'une modalité, les auteurs dans [Coutaz 95] définissent une fonction d'accessibilité appelée "Reach" qui prend la valeur vraie pour deux états si une modalité permet de passer directement de l'un à l'autre. Une trajectoire d'interaction est une séquence d'étapes successives.
- ? Une relation temporelle caractérise l'utilisation dans le temps d'un ensemble de modalités. L'utilisation peut être simultanée ou en séquence.

A l'aide de ces définitions, les auteurs dans [Coutaz 95] définissent les quatre propriétés de façon formelle, développent leurs caractéristiques et les illustrent avec un exemple de système multimodale en entrée : MATIS [Nigay 94]. Dans ce rapport nous définissons informellement ces propriétés et les appliquons aux interfaces en sortie.

#### 4.4.1. *Assignment*

Cette propriété exprime l'obligation pour un état de départ et d'arrivée fixés d'utiliser une modalité. Cette propriété exprime l'absence de choix pour réaliser une tâche. Ceci n'est pas forcément un point négatif pour l'interaction. Elle dénote un manque évident de souplesse mais peut aussi être intéressante pour la robustesse en entrée comme en sortie : les fonctions critiques existent aussi pour les interfaces en sortie. L'exemple le plus courant est le message d'erreur. Ce dernier doit pouvoir être produit quel que soit l'état d'instabilité du système (déficiency du système de gestion du son ou du graphisme). La solution adoptée par Apple est l'utilisation du "debugger" en mode texte pour les messages d'erreur. De plus l'assignation d'une modalité en sortie à une tâche permet de garantir une stabilité de la forme des expressions de sortie.

#### 4.4.2. *Equivalence*

L'équivalence de modalités se définit pour un état de départ et d'arrivée fixés. L'équivalence d'un ensemble de modalités est vérifiée si chaque modalité permet d'atteindre l'état d'arrivée depuis l'état de départ en une seule étape. L'équivalence et l'assignation posent le problème du choix des modalités. Nous verrons au paragraphe 5 que dans le cas de l'interface en sortie, ce problème va au delà de l'existence ou non d'un choix. La relation d'équivalence permet de créer les classes de modalités qu'il est envisageable de mettre en œuvre pour atteindre un état souhaité.

#### 4.4.3. *Redondance*

La redondance dénote l'utilisation séquentielle ou parallèle de plusieurs modalités équivalentes. La redondance d'information en provenance de l'utilisateur implique la prise en compte d'une seule des modalités par le système. Par contre en sortie, la redondance s'avère une propriété beaucoup plus intéressante. Si le concepteur a choisi de montrer de différentes manières une même information, cela a pour conséquence de renforcer son observabilité et sans doute constitue un moyen pour attirer l'attention de l'utilisateur.

#### 4.4.4. *Complémentarité*

La complémentarité entre les modalités d'un ensemble  $M$  exprime le fait que pour passer d'un état à un autre il faille utiliser toutes les modalités de l'ensemble  $M$ . Cela signifie qu'aucune des modalités de l'ensemble  $M$  ne suffise à elle seule pour passer de l'état de départ à l'état d'arrivée. Toutefois, il n'est pas exclu qu'une modalité non contenue dans  $M$  puisse permettre cette transition. Les déictiques, caractérisés par des références entre modalités, sont un exemple d'utilisation complémentaire de modalités. L'utilisation complémentaire des modalités peut, comme pour la redondance, être parallèle ou séquentielle. En sortie, la complémentarité implique que l'utilisateur établisse les liens entre les informations présentées selon des

---

modalités différentes : il doit combiner les informations perçues pour interpréter l'expression de sortie.

#### 4.4.5. Conclusion sur les propriétés CARE

L'assignation et l'équivalence traduisent la présence ou non de choix des modalités en entrée comme en sortie. Au contraire la complémentarité et la redondance caractérisent l'usage de plusieurs modalités pour une tâche donnée. Cependant la redondance implique l'équivalence des modalités. Aussi nous considérons uniquement la complémentarité comme un constructeur de modalités complexes à partir de modalités simples ou "pures", si l'on reprend les termes utilisés par Bernsen (paragraphe 4.3). Ainsi l'assignation et l'équivalence peut se définir entre modalités "pures" et modalités complexes. Par exemple une modalité "pure" <m1> peut être équivalente à une modalité complexe <complémentaire (m2,m3)>. La souplesse qu'implique le choix entre modalités peut alors être accrue si les choix se font entre des modalités "pures" ou complexes. Les choix portent alors sur la modalité à utiliser mais aussi sur l'usage de plusieurs modalités. Dans le cas de la conception d'une interface en sortie, le concepteur peut alors se poser les questions suivantes :

? Faut-il utiliser des modalités "pures" ou complexes pour la tâche considérée ?

? Faut-il exploiter la redondance pour la tâche considérée ?

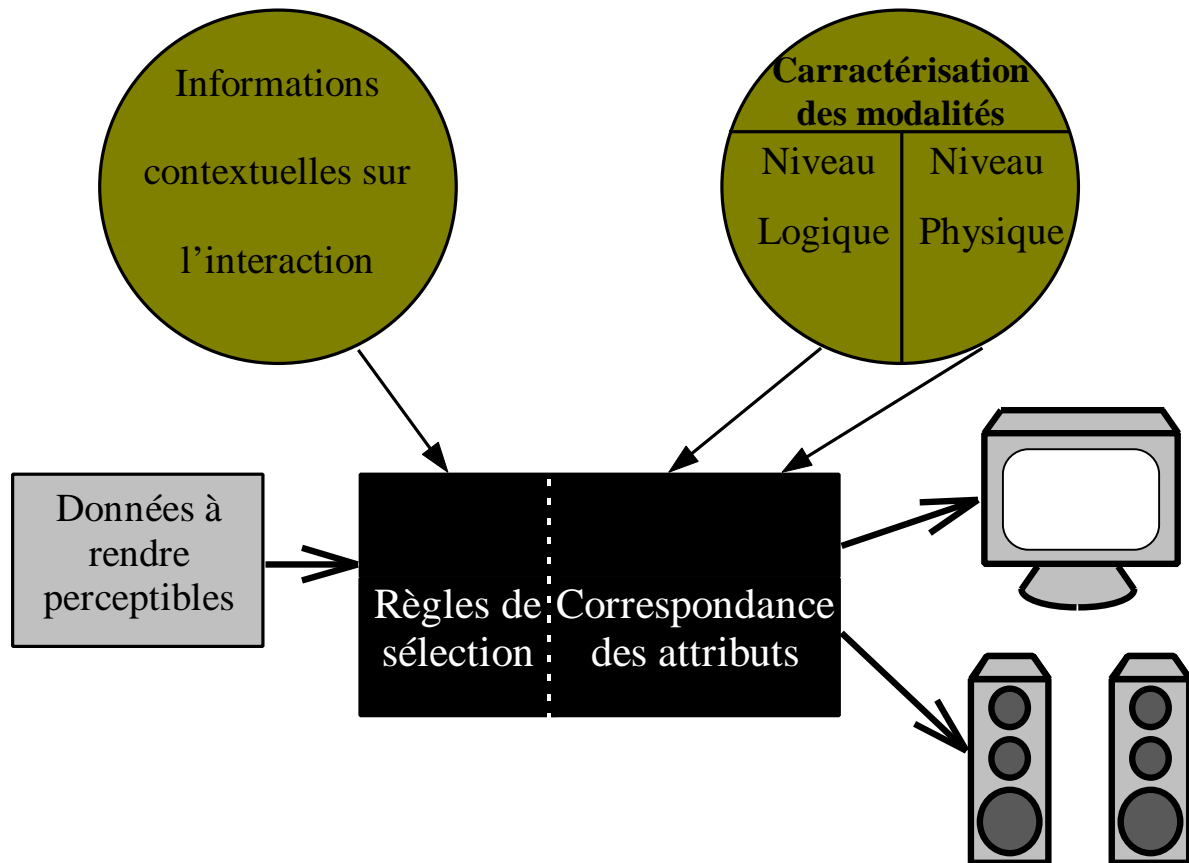
La réponse à ces questions ne peut que reposer sur un ensemble de critères qu'il convient d'identifier.

## 5. Critères relatifs à la multimodalité en sortie

La principale différence entre la conception d'un système multimodale en entrée et en sortie réside dans le choix des modalités. Dans le cas de la multimodalité en entrée, c'est l'utilisateur qui choisit la modalité, à condition que le système offre le choix. Au contraire, dans le cas de la multimodalité en sortie, le système doit choisir la ou les modalités les plus adéquates pour communiquer l'information à l'utilisateur.

Les taxonomies que nous avons présentées aux paragraphes précédents sont utiles pour caractériser les modalités et leurs usages. Ces caractéristiques permettent de guider le choix des modalités en sortie à utiliser. Cependant ces choix ne sont pas statiques et peuvent aussi dépendre du contexte d'interaction et en particulier de la tâche courante et des données à rendre perceptibles. Dans [Nigay 94] une définition du contexte d'interaction souligne la diversité des informations contenues. Celles-ci peuvent concerner la tâche courante, l'état courant de l'interface (comme le nombre de fenêtres à l'écran), l'environnement d'utilisation (tel que le niveau sonore de la pièce ou la position de l'utilisateur par rapport à l'écran) ou encore l'utilisateur et ses compétences.

A la Figure 3, nous mettons en relation le contexte de l'interaction, l'ensemble des modalités et de leurs caractéristiques, les données à rendre perceptibles et les règles de sélection et de composition pour définir la présentation multimodale.



**Figure 3 :** Conception d'une présentation multimodale.

Deux niveaux de caractéristiques des modalités, le niveau logique et le niveau physique sont distingués à la Figure 3. Ces deux niveaux correspondent respectivement aux deux composants d'une modalité  $\langle d,s \rangle$ , où  $d$  est le dispositif physique et  $s$  le système représentationnel (paragraphe 1). Par exemple les caractéristiques physiques sont la disponibilité, le débit ou la persistance. Au niveau logique nous utilisons les caractéristiques de Bernsen telles que linguistique, analogique ou arbitraire.

Les règles de sélection et de composition consistent à définir des valeurs pour les caractéristiques des modalités à partir des données à rendre perceptibles et le contexte d'interaction. Ces valeurs ainsi définies permettent alors de choisir le ou les modalités et donc de définir la présentation multimodale.

Grâce au schéma directeur de la Figure 3, nous pouvons identifier des règles de sélection et de composition qui sont encore très générales :



---

### **Règles de sélection**

- ? Les modalités à utiliser doivent utiliser au moins l'un des dispositifs disponibles (canal libre de l'interface du système au sens de Frohlich).
- ? Les modalités à utiliser doivent être adéquates à la tâche courante (identification d'un ensemble de tâches comme dans la taxonomie de Foley).
- ? Les modalités à utiliser doivent faire intervenir le canal le moins occupé et donc prendre en compte l'environnement d'utilisation propre à l'utilisateur (canal de communication de l'utilisateur au sens de Frohlich).
- ? Les modalités à utiliser doivent être en accord avec les préférences établies par l'utilisateur (modèle de l'utilisateur).

### **Règles de composition**

- ? Si la tâche courante est critique, comme un message d'alarme, plusieurs modalités doivent être utilisées de façon redondante.
- ? Si la présentation d'une donnée repose sur une modalité arbitraire, une modalité complémentaire et non arbitraire doit accompagner la première modalité. Un exemple simple de cette règle est la légende qui accompagne toujours une carte.

La précision des règles énoncées varient et il convient de les affiner. Certaines de ces règles sont trop générales pour être directement appliquées par le concepteur. De plus d'autres règles doivent être identifiées en vue de définir une méthode de conception complète.

## **6. Conclusion**

Dans ce chapitre, nous avons présenté les taxonomies existantes dans le but de mieux comprendre et cerner la multimodalité. Nous avons aussi identifiées les différences de conception entre une interface en entrée et en sortie. De plus les taxonomies dédiées aux interfaces en entrée ont été étendues aux interfaces en sortie.

L'étude de ces taxonomies a permis d'identifier des caractéristiques utiles pour distinguer les modalités. Ces caractéristiques sont cruciales pour pouvoir identifier des règles de choix et de composition des modalités, étape incontournable dans la conception de présentations multimodales. Aussi le dernier paragraphe présente notre cadre unificateur à la conception de présentation multimodale et un ensemble de règles de conception.

Dans le chapitre suivant, nous appliquons nos résultats sur la multimodalité en sortie au cas du rendu d'une grande quantité d'informations. Dans ce contexte, nous allons affiner nos règles de conception.



## **Chapitre III**

*Grande quantité d'informations  
et multimodalité*

Ce chapitre étudie les apports de la multimodalité aux systèmes de rendu de grandes quantités d'informations qui constituent notre domaine applicatif. Notre démarche de travail s'appuie sur notre étude de la multimodalité. Notre objectif consiste à sélectionner les modalités adéquates pour une interaction optimale. Pour ce faire, nous rappelons les problèmes liés aux grandes quantités d'informations au paragraphe 1. Nous proposons ensuite deux approches complémentaires pour produire des critères de sélection et de combinaison adaptés au rendu des grands espaces d'informations. Les deux approches sont explorées au paragraphe 3 et les critères proposés sont illustrés au paragraphe 4 avec des systèmes existants. Le paragraphe 5 reprend en détail la réalisation d'un système qui correspond à l'un des critères identifiés par notre étude : la représentation multiple.

## **1. Grandes quantités d'information**

De grandes quantités d'informations sont générées très fréquemment par les systèmes informatiques en raison de leur grande rapidité de traitement. Ces quantités d'informations sont de plus en plus fréquentes, dû aux progrès de la technologie des moyens de stockage (disques durs, bandes magnétiques, CD-ROM, etc.). Enfin la percée d'Internet à tous les niveaux de notre société (entreprises, particuliers, administrations, etc.) a rendu accessible à tout le monde de très grandes quantités d'informations.

La perception et la manipulation de ces grands ensembles soulèvent des problèmes d'utilisabilité. Le problème principal, identifié à la fin des années 80, est celui de la perte de la perception globale de l'espace d'informations lorsqu'on manipule une infime partie appelée centre d'intérêt. Nous nous trouvons donc face à des problèmes d'accessibilité et non de disponibilité de l'information. Pour illustrer nos propos, considérons la métaphore d'un labyrinthe. Il est très difficile de s'orienter dans un labyrinthe lorsque l'on se trouve à l'intérieure. Au contraire lorsque l'on a une vision globale du labyrinthe, vue de l'extérieur, la navigation est plus facile. Cette métaphore met en évidence la nécessité d'une vue globale de l'espace pour s'orienter. Bob Spence [Spence 97] dénote ce problème « focus + context » que nous traduirons par « centre d'intérêt + contexte d'interaction ». Le problème est de savoir comment représenter le centre d'intérêt, tout en montrant une vue globale, qui nous permette de situer le centre d'intérêt dans son contexte.

Pour apporter des éléments de réponse au problème du « centre d'intérêt + contexte d'interaction » notre approche consiste à exploiter la multimodalité. Au Chapitre II, nous avons étudié le mécanisme de sélection des modalités de sortie qui soient adéquates à la tâche à réaliser. Nous avons proposé un cadre général de conception pour la génération de présentations multimodales. La génération s'appuie sur des règles de sélection et de combinaison de modalités. Les taxonomies existantes nous ont permis d'identifier des caractéristiques des modalités nécessaires pour guider la sélection. Cette étude générale à la multimodalité doit maintenant être instanciée et

---

approfondie dans le cas du rendu de grandes quantités d'information. Pour cela nous adoptons deux approches complémentaires :

Une première approche consiste à étudier les propriétés d'ergonomie. Ces propriétés définissent l'utilisabilité générale de l'interface. Nous allons choisir les propriétés les plus pertinentes pour le cas du rendu de grandes quantités d'informations. L'étude de ces propriétés permet d'identifier des objectifs généraux en terme d'utilisabilité des modalités à utiliser.

Ainsi les propriétés d'ergonomie servent de premier moyen de sélection des modalités car certaines modalités sont clairement plus adaptées à vérifier certaines propriétés.

Une deuxième approche complémentaire à la première, consiste à approfondir les caractéristiques générales des modalités identifiées au chapitre précédent. Ceci nous conduit à définir des caractéristiques propres au rendu de grandes quantités d'informations.

Dans ce chapitre, nous commençons par énoncer la liste des propriétés ergonomiques, liste de référence pour notre première approche. Au paragraphe suivant, 3.1., nous détaillons les propriétés les plus pertinentes pour le rendu de grandes quantités d'informations. Nous montrons comment ces propriétés peuvent aider à choisir les modalités de sortie les plus adéquates. Au paragraphe 3.2., nous présentons de nouvelles caractéristiques des modalités qui sont propres au rendu de grandes quantités d'information : nous montrons comment ces caractéristiques sont utilisées pour la sélection de modalités en énonçant des règles de sélection et de composition. Nous illustrons nos résultats au paragraphe 4 en les appliquant à des systèmes existants. Nous constatons enfin que la représentation multiple de l'espace d'informations est un concept important qui est commun à nos approches (paragraphe 3) : la représentation multiple fait l'objet du paragraphe 5.

## **2. Propriétés d'ergonomie**

Les propriétés d'ergonomie constituent un raffinement du critère d'utilisabilité de génie logiciel. Ce sont les ergonomes qui, en essayant de qualifier l'utilisabilité d'un système, ont introduit ces propriétés. Elles ne sont pas encore définitivement établies et il existe de nombreuses listes. Nous adoptons l'approche décrite dans [Gram 96], où les propriétés sont classées en deux catégories : les propriétés de robustesse et de souplesse.

## 2.1. Propriétés de souplesse

<b>Propriété</b>	<b>Nom en anglais</b>	<b>Définition</b>
Atteignabilité	« Reachability »	Capacité du système à permettre à l'utilisateur d'atteindre le but désiré depuis l'état actuel.
Non-préemption	« Non-Preemptiveness »	Capacité du système à rendre le prochain but souhaité par l'utilisateur directement accessible.
Interaction à plusieurs fils	« Multi-Threading »	Capacité du système à permettre à l'utilisateur de réaliser plusieurs tâches de manière entrelacée.
Multiplicité de la représentation	« Representation Multiplicity »	Capacité du système à fournir plusieurs représentations d'un même concept.
Réutilisabilité des données d'entrée et de sortie	« I/O Re-use »	Capacité du système à réutiliser les données en provenance de systèmes extérieurs et à fournir des données qui seront réutilisables par d'autres systèmes.
Adaptabilité	« Customizability »	Personnalisation du système sur intervention explicite de l'utilisateur.
Migration de tâche	« Migratability »	Capacité de délégation dynamique de tâches entre le système et l'utilisateur (c.a.d. changement de l'acteur responsable de la tâche).
Multiplicité des rôles	« Human Rôle Multiplicity »	Capacité du système à connaître le rôle de l'utilisateur et d'utiliser cette donnée dans l'interaction.

## 2.2. Propriétés de robustesse

Propriété	Nom en anglais	Définition
Observabilité	« Observability »	Capacité pour l'utilisateur d'évaluer l'état interne du système.
Insistance	« Insistence »	Capacité du système à forcer la perception de l'état interne du système.
Honnêteté	« Honesty »	Capacité du système à rendre observable l'état interne du système sous une forme conforme à cet état et qui engendre une interprétation correcte de la part de l'utilisateur.
Curabilité	« Recoverability »	Capacité pour l'utilisateur de corriger une situation non désirée.
Prévisibilité	« Predictability »	Capacité pour l'utilisateur de prévoir, pour un état donné, l'effet d'une action.
Tolérance au rythme	« Pace Tolerance »	Capacité du système à laisser l'utilisateur décider quand il peut agir.

## 3. Propriétés de sélection relatifs aux grands espaces d'informations

En introduction à ce chapitre, nous avons annoncé nos deux approches complémentaires. La première approche nécessitait le rappel des propriétés d'ergonomie, qui a fait l'objet du paragraphe précédent. Ces propriétés seront utilisées au paragraphe 3.1 pour identifier des règles de sélections des modalités adaptées aux grandes quantités d'informations. Le paragraphe 3.2 est dédié à notre seconde approche qui consiste à approfondir l'espace de conception du chapitre II au cas du rendu de grandes quantités d'informations.

### 3.1. Propriétés d'ergonomie relatifs aux grands espaces d'informations

L'objectif global d'adéquation des modalités à la tâche reste trop vague. Par exemple dans le cas de barres de défilement, l'honnêteté est mieux vérifiée lorsque la taille de l'ascenseur à l'intérieure de la barre de défilement est proportionnelle à la taille du document. L'utilisateur interprète alors mieux la taille des déplacements dans l'espace d'informations.

Parmi les propriétés d'ergonomie, nous avons retenu cinq propriétés de robustesse pertinentes pour notre étude. Ces cinq propriétés sont issues de la liste du

paragraphe 2.2. Nous indiquons ci-dessous, pour chacune de ces propriétés, l'incidence sur le choix des modalités de sortie.

? L'observabilité est la capacité de l'utilisateur à évaluer l'état interne du système à partir de sa présentation. Toutefois l'observabilité totale n'est pas le remède au manque d'observabilité d'un système. Il faut bien discerner les variables internes qu'il est pertinent d'observer. Dans le cas de grandes quantités d'informations, nous avons retenu les variables suivantes :

- ? le centre d'intérêt ou « focus » en anglais,
- ? l'ensemble de l'espace d'informations,
- ? les liens entre les informations,
- ? la taille de l'espace d'informations.

L'observabilité d'une variable n'est pas une propriété booléenne. Certaines modalités peuvent être plus efficaces que d'autres pour rendre observables les variables citées ci-dessus. Le concepteur d'applications multimodales doit faire des choix en fonction des variables qu'il juge vraiment pertinentes à rendre observables et choisir les modalités les plus efficaces. Un exemple type d'application de cette propriété est le choix de la taille d'un objet à l'écran en fonction de l'importance qu'on lui accorde.

? L'honnêteté est la capacité de l'utilisateur à évaluer correctement l'état interne du système. Elle est garante de la bonne interprétation des données quand le système représente son état interne de façon inhabituelle. Il est crucial lorsque l'utilisateur navigue dans l'espace, qu'il puisse anticiper son déplacement. Si l'utilisateur n'est pas capable d'anticiper les résultats de son action de navigation, il va se perdre dans l'espace. Un autre exemple d'honnêteté est l'effet de perspective donné par les applications du type du mur en perspective (cf. paragraphe 4.3). Cet effet permet à l'utilisateur de comprendre que la taille des informations n'est pas proportionnelle à leur importance mais à leur emplacement par rapport au centre d'intérêt. Choisir des modalités honnêtes est donc une priorité pour le concepteur.

? La prévisibilité est la propriété qui indique que pour chaque action entreprise, le résultat peut être anticipé. Dans notre cas, elle est pertinente pour la navigation de l'espace des informations. Il est crucial lorsque l'utilisateur navigue dans l'espace, qu'il puisse anticiper son déplacement. Si l'utilisateur n'est pas capable d'anticiper les résultats de son action de navigation, il va se perdre dans l'espace.

? La curabilité exprime la possibilité de corriger un état non souhaité. Dans le cas des grandes quantités d'information, l'état indésirable le plus fréquent est d'être perdu. La curabilité de ce fait implique la possibilité offerte par le système de pouvoir revenir dans un endroit connu de l'espace d'informations. L'usage de modalités qui permettent de percevoir ces points de repères sont donc cruciaux. D'autre part il est envisageable d'utiliser une modalité de sortie dédiée uniquement à la présentation de ces points de référence, afin d'aider l'utilisateur à les percevoir. Par exemple



---

l'usage d'une modalité sonore serait intéressant : les points de référence fonctionneraient alors comme des balises sonores, les sons produits seraient dynamiques par rapport à l'emplacement courant donc au contexte d'interaction.

- ? L'insistance est la propriété qui indique que le système essaie de forcer la perception de l'utilisateur. Cette propriété est très utilisée pour les messages importants tels que les messages d'erreurs. Dans le cas de rendu de grandes quantités d'informations, certaines informations sont clairement plus importantes que d'autres (comme par exemple la pertinence calculée des réponses à un moteur de recherche). Il convient dans ces cas, d'utiliser des modalités qui traduisent cette insistance : cette dernière peut se décliner en caractéristiques des modalités à utiliser telles que « non éphémère », « inévitable » et « entretenu ».

Nous avons également retenu trois propriétés de :

- ? La multiplicité des représentations reflète la capacité du système à afficher les données de différentes manières. Dans le cas des techniques d'interaction, elle permet de faire face à la diversité des utilisateurs et des tâches. Une seule représentation ne pourra satisfaire tous les utilisateurs. La grande variété dans les utilisateurs incite à proposer des techniques différentes parmi lesquelles l'utilisateur pourra choisir. La multiplicité des représentations revient à proposer une interaction multimodale redondante et séquentielle (cf. paragraphe 4.4 du chapitre II).
- ? La réutilisabilité des données en entrée et en sortie est une propriété importante pour une bonne utilisation des systèmes en général. Dans notre domaine, elle permet d'insérer la technique d'interaction dans la démarche globale de l'utilisateur. en effet l'utilisateur peut alors directement utiliser les données affichées (données en sortie) pour spécifier des actions au système (données en entrée). Ce point est particulièrement important si l'on souhaite offrir la manipulation directe sur l'espace d'informations. L'utilisateur se déplace alors dans l'espace en manipulant l'espace lui-même.
- ? La réutilisabilité impose donc de choisir une modalité compatible avec les autres applications présentes dans l'environnement de l'utilisateur.
- ? L'adaptabilité est la propriété de souplesse qui indique la possibilité de paramétrer le système. Les techniques d'interaction nécessitent la vérification de cette propriété car elle permet à l'utilisateur d'adapter la technique d'interaction à ses préférences. Cette propriété impose entre autre au concepteur d'implémenter des modalités que l'utilisateur puisse paramétrer. Par exemple, un utilisateur daltonien peut ainsi choisir une modalité qui ne fasse intervenir que des couleurs qu'il perçoit.

Le respect de ces huit propriétés permet de prendre en compte un certain nombre d'aspects importants de la conception de systèmes de rendu d'informations. Toutefois, certaines questions soulevées restent assez vagues. Par exemple, nous prônons l'observabilité du centre d'intérêt et du contexte mais nous ne pouvons pas être plus

explicite sur la manière de choisir de telles modalités. Nous proposons donc d'explorer la seconde approche qui consiste à caractériser les modalités de sortie propres au rendu de grands espaces d'informations.

### 3.2. Caractérisation des modalités de sortie

Comme nous l'avons annoncé en introduction, notre seconde approche repose sur nos résultats du Chapitre II. Nous approfondissons et étendons les concepts identifiés pour le cas du rendu de grandes quantités d'informations.

Nous proposons tout d'abord de caractériser les dispositifs pour mettre en évidence les caractéristiques intéressantes des périphériques de sortie. Nous proposerons ensuite des critères de sélection du dispositif adaptés en utilisant ces caractéristiques. Ces critères s'appliqueront aussi aux modalités puisqu'une modalité est définie entre autre par le dispositif physique associé.

Après avoir caractérisé les dispositifs, il convient d'étudier les systèmes représentationnels, deuxième composant d'une modalité. De la même manière que pour les dispositifs, ces caractéristiques vont nous permettre d'énoncer des règles de sélection des modalités. Enfin nous étudions dans un troisième paragraphe des règles de composition qui reposent d'une part sur la complémentarité (paragraphe 4.4 du Chapitre II) et d'autre part sur nos caractéristiques des dispositifs et systèmes représentationnels.

#### 3.2.1. Caractérisation des dispositifs de sortie

Nous proposons cinq caractéristiques des dispositifs de sortie que nous illustrons du point de vue du système et du point de vue de l'utilisateur.

		Caractéristiques				
		Sens perceptif	Débit d'informations maximum	Capacité du dispositif	Persistance	Nombre de dimensions spatiales mises en jeu
Point de vue	Système	Dispositif de sortie : * haut-parleur * écran, imprimante * système à retour d'effort	Nombre d'informations par unité de temps	Quantité d'informations véhiculées par le dispositif.	Durée de vie des informations véhiculées	Degrés de liberté du dispositif de sortie $N_s$
	Utilisateur	Sens perceptif humain : * ouïe * vue * toucher	Capacité maximum de l'utilisateur (selon le sens perceptif)	Taille de l'extension de la mémoire à court terme	Durée de vie de l'extension de la mémoire à court terme	Nombre de dimensions perçues $N_u$ $N_u ? N_s$

Nous définissons les propriétés des cinq caractéristiques comme suit :

? Le sens perceptif définit trois classes de dispositifs. La première classe est celle des dispositifs sonores qui correspondent à l'ouïe humaine. La seconde classe est celle des dispositifs graphiques qui correspondent à la vue. La troisième classe regroupe

---

les dispositifs novateurs à retour d'effort. Comme Bernsen [Bernsen 93] et Frohlich [Frohlich 91] nous pensons que cette caractéristique est fondamentale pour un dispositif. Elle permet notamment d'exprimer des critères qui prennent en compte les utilisateurs handicapés (non-voyants ou sourds).

- ? Le débit maximum d'un dispositif dérive de ses limites physiques. Nous utilisons cette caractéristique pour définir des critères d'incompatibilité entre un dispositif et un flux d'information. Il s'agit donc d'un raffinement du critère de dynamique de la taxonomie de Bernsen. Par exemple, si un système physique envoie au système une donnée toutes les heures, le concepteur pourra choisir l'imprimante alors que si l'interface reçoit dix nouvelles données par seconde, l'écran devra être utilisé.
- ? La capacité d'un dispositif exprime la quantité d'informations qu'il rend disponible pour l'utilisateur à un instant donné. Cette caractéristique permet de différencier les dispositifs qui constituent un prolongement de la mémoire à court terme de l'utilisateur. Il est crucial de choisir une modalité de sortie adaptée qui ne dépassera pas la capacité du dispositif (comme une fenêtre qui dépasserait la taille de l'écran).
- ? La persistance est la durée pendant laquelle un dispositif peut maintenir une information. Cette durée, ajoutée à la durée de vie de la mémoire à court terme permet de choisir les modalités qui permettent à l'utilisateur de mémoriser les informations véhiculées par cette modalité. Cette caractéristique est à rapprocher du critère « non-éphémère » de Sellen [Sellen 92].
- ? Le nombre de degrés de liberté correspond au nombre de dimensions mises en œuvre par le dispositif. Cette caractéristique se raffine par le nombre de dimensions réellement perçues par l'utilisateur. Par exemple un écran qui a deux degrés de liberté permet des effets en trois dimensions. Nous pouvons exprimer comme règle que le dispositif employé dans une modalité doit toujours avoir au moins autant de degrés de liberté que l'information initiale.

Cette classification permet de caractériser les dispositifs de sortie de façon à différencier par exemple les écrans des imprimantes (débit meilleur pour les écrans mais persistance plus grande pour l'imprimante). Elle permet aussi de différencier les différents types de systèmes de rendu sonore par le nombre de dimensions. Les règles engendrées sont des règles restrictives. Toutefois, choisir le dispositif n'est pas la partie la plus compliquée dans le choix des modalités. Nous allons maintenant voir quelles sont les caractéristiques que nous pouvons dégager des systèmes représentationnels, deuxième composant d'une modalité.

### 3.2.2. *Caractérisation des systèmes représentationnels*

La caractérisation des langages d'interaction que nous proposons repose essentiellement sur le graphisme. Nous avons en effet constaté que la totalité des grands espaces d'informations reposaient sur des modalités graphiques. Ce reproche n'exclut toutefois pas la présence de plusieurs dispositifs au sens Multi-Disp de

l'espace M<sup>2</sup>LD puisqu'on peut imaginer deux dispositifs du même type intervenant dans l'interaction (par exemple deux écrans).

Les trois caractéristiques que nous identifions dans les systèmes représentationnels sont assez hétéroclites mais permettent de mettre en évidence la nature des techniques utilisées dans les systèmes.

- ? Nous différencions tout d'abord les rendus partiels et les rendus globaux de l'espace d'informations. Partiels et globaux indiquent si la modalité rend la totalité de l'espace d'informations ou seulement une partie.
- ? Une autre caractéristique importante est la précision du système représentationnel : précise ou vague. Cette différence ne peut se faire qu'en considérant la tâche de l'utilisateur manipulant l'espace d'informations. Nous définissons un système représentationnel comme précis si il permet de présenter assez précisément les informations sur un élément de l'espace pour réaliser la tâche. Cette caractéristique est à mettre en relation avec la propriété d'ergonomie d'observabilité.
- ? Nous présentons enfin une troisième caractéristique qui ne s'applique qu'aux modalités construites. Elle traduit l'utilisation ou l'absence d'une fonction de déformation lors de la construction de la modalité.

Les fonctions de déformation de l'espace ne sont pas des systèmes représentationnels mais des modificateurs de ces derniers. En cela nous pouvons les considérer comme des constructeurs à partir d'un système représentationnel donné. Dans le paragraphe 4, nous illustrons les fonctions de déformation en présentant des systèmes existants.

Nous proposons deux règles heuristiques de sélection qui reposent sur les trois caractéristiques ci-dessus.

- ? La modalité présentant le centre d'intérêt ne doit pas être déformée. En effet la déformation d'une partie de l'espace d'informations impose à l'utilisateur de comprendre la déformation et accroît donc la charge cognitive.
- ? La deuxième règle de sélection concerne les modalités globales qui servent à rendre perceptible la structure de l'espace d'informations. L'éventuelle déformation de cette modalité ne doit pas entraver cette perception. Cette règle dépend beaucoup des informations sur la structure dont l'utilisateur a besoin. Ainsi l'utilisateur peut se servir du contexte pour percevoir le contour de l'espace d'informations, la couleur, ou encore des tendances sur la taille des éléments.

### 3.2.3. *Combinaisons de modalités*

Aux deux paragraphes précédents, nous avons caractérisé les modalités prises isolément. En nous reposant sur les caractéristiques identifiées, nous étudions maintenant la combinaison de modalités. Pour cela, nous utilisons la complémentarité

---

identifiée comme opérateur de combinaison au Chapitre II. Nous proposons deux nouveaux types de complémentarité :

- ? La complémentarité temporelle définit si celle-ci est séquentielle ou parallèle (au même instant). Par exemple la représentation multiple de l'espace d'informations peut être au même instant dans des fenêtres distinctes ou encore séquentielle (passage d'une modalité à une autre).
- ? La complémentarité spatiale caractérise la frontière entre deux parties de l'espace présentées selon des modalités distinctes. La complémentarité spatiale peut alors être une simple juxtaposition ou une composition de fonction. Il est à noter que la juxtaposition de deux modalités risque d'entraîner une discontinuité visuelle.

Le problème de la représentation d'une grande quantité d'informations en rendant perceptible à la fois le centre d'intérêt et le contexte d'interaction s'exprime donc comme suit :

- ? Il faut toujours avoir en parallèle une modalité qui soit **précise** et une modalité (éventuellement construite) qui soit **globale**. Si cette règle ne peut pas être vérifiée, un relâchement de la condition de simultanéité peut être envisagé.

Nous proposons un autre type de règle qui reflète la continuité visuelle. Il s'agit en effet d'une motivation très importante dans notre domaine depuis la publication de [Mackinlay 95] qui a introduit l'expression « Smoothly Integrated » reprise dans de nombreux travaux. La continuité visuelle se vérifie en deux règles qui couvrent les deux types de complémentarité énoncés ci-dessus :

- ? Nous préconisons la continuité visuelle lorsque deux modalités sont combinées. Cela implique que deux modalités utilisées en parallèle pour représenter des espaces proches, doivent garantir que la présentation des deux espaces soient aussi proches.
- ? Nous proposons la règle de continuité temporelle comme une alternative au changement brusque de modalité à un instant donné. Lorsqu'une modalité laisse subitement place à une autre modalité très différente, l'utilisateur perd ses repères. Il faut donc garantir la continuité visuelle. Cette règle sera reprise plus en détail au paragraphe 5.

La distinction entre continuité visuelle temporelle et spatiale peut paraître subtile. Nous proposons une façon simple de les distinguer. La continuité spatiale est vérifiée si l'utilisateur peut balayer l'espace d'informations sans noter de rupture. Au contraire la continuité temporelle se vérifie si l'utilisateur fixant un point de l'espace ne distingue pas de rupture au court du temps.

#### **4. Exemples d'application de nos caractérisations**

Il existe des systèmes qui visent à résoudre le problème de la visualisation du centre d'intérêt dans son contexte en utilisant des objets d'interactions appropriés. Nous allons présenter ces systèmes selon nos caractéristiques. Pour chaque système,

nous proposons aussi de les confronter à nos règles de sélection et de composition des modalités.

#### 4.1. Lentilles magiques

Les lentilles magiques [Stone 94] (« magic lenses ») ne déforment pas le champ des informations mais filtrent son contenu. Sans lentille magique toutes les informations sont représentées à importance égale (le treillis est donc très confus). Les lentilles magiques permettent de filtrer certaines zones de l'espace par des fonction de zooms, de filtres sémantiques (Figure 4) ou encore de modification des couleurs. L'intérêt de ces filtres est de les combiner pour laisser apparaître un nombre minimal d'informations. Le déplacement des lentilles permet de balayer la totalité des informations et la superposition des lentilles permet de dégager une information particulière. Ce système est assez efficace sur les espaces denses d'informations comme les cartes géographiques ou météorologiques.

**Figure 4 :** Deux lentilles magiques superposées sur le plan d'une ville

La caractérisation des modalités nous donne le résultat suivant :

- ? une modalité par lentille magique qui est partielle, vague avec une fonction de déformation qui met en valeur un type particulier d'informations (filtre sémantique) ;
- ? une modalité pour l'intersection des lentilles qui est partielle, avec toute l'information nécessaire à la tâche de l'utilisateur donc précise ; la fonction de déformation est définie par la combinaison des filtres superposés ;
- ? Une modalité pour le « fond » qui est globale et vague.

Nous constatons qu'il existe bien une modalité précise (la seconde dans la liste) et une qui est globale (la troisième). De plus le passage d'une modalité à l'autre se fait presque sans discontinuité spatiale. La seule discontinuité se trouve dans la zone au

---

dessus de la lentille où se trouve la barre de titre. Notre dernière remarque sur ce système concerne la possible disparition éventuelle d'une modalité précise puisqu'il incombe à l'utilisateur de faire recouper les fenêtres.

En présentant ce système nous n'avons pas évoqué la fonction de zoom que peut exécuter une lentille magique car il s'agit d'un cas très différent. De plus les lentilles magiques de zoom sont très proches du système « DragMag ».

#### **4.2. Système « Drag Mag »**

Le système « Drag Mag » [Ware 95] est une amélioration des lentilles magiques agrandissantes qui permet de situer le centre d'intérêt dans le contexte. En effet les lentilles magiques agrandissantes cachent l'emplacement du centre d'intérêt à l'utilisateur (comme une loupe physique). Le « Drag Mag » quant à lui, déplace la zone « agrandie » et la situe dans le contexte en reliant la zone à agrandir et la zone agrandie par la vue en perspective (Figure 5). L'impression produite est celle d'un bâtiment vu du ciel dont la base (très lointaine) serait recopiée sur le toit (plus proche). Il est possible de déplacer la face arrière qui fait bouger le bâtiment dans son entier. Cette action ne modifie que l'emplacement de la zone agrandie et pas le coefficient d'agrandissement. Il est également possible de déplacer la face avant avec comme résultat la modification de la hauteur de la tour et donc du coefficient d'agrandissement.

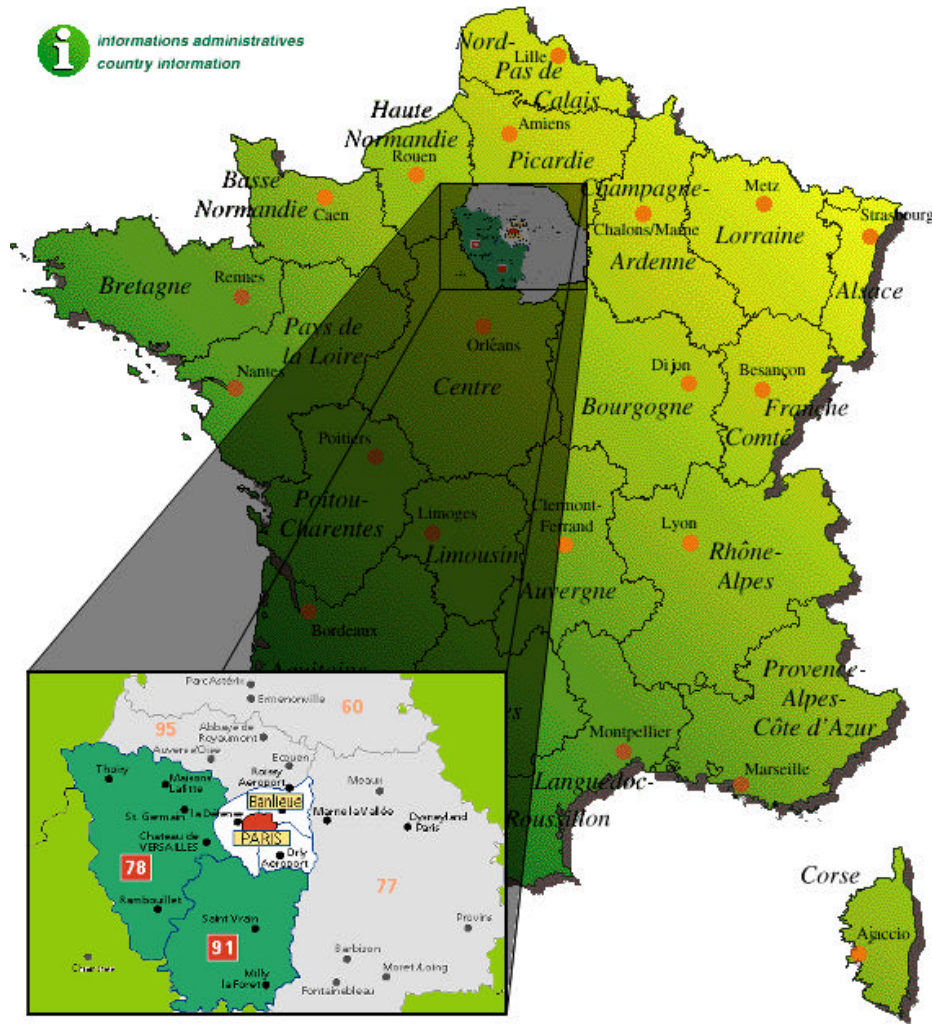


Figure 5 : Un « Drag Mag » sur la région Parisienne

Nous caractérisons les modalités de « Drag Mag » comme suit :

- ? une modalité de « fond » : quasi-globale, vague et sans déformation ;
- ? une modalité pour la zone agrandie, partielle, précise, et sans déformation ;
- ? une modalité pour les faces assombries qui donnent l'effet de perspective, partielle, générale et avec une déformation de perspective. Nous considérons cette modalité comme générale. En effet elle ne véhicule, comme information, que le lien entre les emplacements des deux autres modalités.

Nous constatons qu'il y a une modalité globale et une modalité précise. Toutefois une partie de la modalité globale est cachée par la modalité précise. Les auteurs de [Ware 95] proposait à cet effet un autre prototype dont les fenêtres ne se superposaient jamais et restaient en dehors du cadre. Si nous constatons ainsi une meilleure vérification du critère de rendu du centre d'intérêt et de son contexte, c'était au détriment de la navigation. En effet si on impose de sortir la fenêtre de zoom de la zone de travail, on ne peut plus se servir de la distance entre la zone agrandie et son



---

emplacement sur la carte comme facteur de grossissement. Le zoom se faisait donc au moyen d'une barre de défilement à l'utilisation moins naturelle.

Le critère de continuité visuelle est respecté entre la modalité globale et la modalité de perspective d'une part et entre la modalité de perspective et la modalité précise d'autre part. Toutefois, on peut reprocher que la continuité soit réduite à la seule structure et pas aux éléments de l'espace d'informations. Par exemple, on ne voit pas la continuité d'un cours d'eau ou d'une route sur une carte comme celle de la Figure 5.

### 4.3. Mur en perspective

Le mur en perspective (« Perspective Wall ») est une technique d'interaction qui permet de visualiser une grande quantité d'informations [Mackinlay 95]. Ces informations sont représentées sur un mur infini. Ce mur contient en général trois faces comme le montre la figure 6 :

- ? la face avant qui n'est pas déformée mais qui permet de visualiser une petite quantité d'informations sans distorsion,
- ? deux faces périphériques, vues en perspective (déformation des informations) mais qui offrent une grande liberté pour représenter le contexte.

Cette représentation est bien adaptée aux listes et aux diagrammes trop longs pour être représentés d'un seul tenant. Elle s'adapte aussi très bien aux représentations du temps puisqu'elle permet de rendre compte séparément du passé et du futur.

**Figure 6 :** Mur en perspective

Nous caractérisons les modalités du mur en perspective de la façon suivante :

- ? une modalité pour la face avant qui est précise, partielle et sans déformation ;
- ? une modalité pour chaque face qui est générale et vague avec une déformation de perspective.

Nous dégageons la modalité globale en additionnant les trois modalités pures proposées. Le critère de présentation du centre d'intérêt et du contexte est ainsi vérifié ainsi que celui de la continuité visuelle grâce à un astucieux montage géométrique des facettes. Nous remarquons enfin que le critère de non déformation du centre d'intérêt est vérifié mais pas celui concernant la structure de l'espace d'information.

#### **4.4. Vue en œil de poisson**

Une vue en œil de poisson (« Fisheye View ») est une vue 2D déformée d'un ensemble d'informations consistant à faire apparaître les informations avec une taille variable. Au premier plan (centre d'intérêt choisi par l'utilisateur), les informations sont de taille satisfaisante. Les autres informations de l'espace sont de taille inversement proportionnelle à leur distance au centre d'intérêt. On obtient ainsi une vue comparable à celle d'un objectif « grand angle » d'un appareil photographique. L'utilisateur peut, sans perdre de vue son centre d'intérêt, utiliser sa vision périphérique comme celle d'un poisson ou d'une mouche pour surveiller tout l'espace. Cette technique peut par exemple rendre méconnaissable une carte des États-Unis (non respect du critère de non-déformation de la structure) comme le montre la Figure 7. La vue en œil de poisson ne permet pas une comparaison aisée entre deux éléments distants. Pour palier à ce problème, le système devrait proposer une autre déformation centrée sur un autre centre d'intérêt. Cette technique (polyfocal display) pose toutefois de sérieux problèmes de navigation.

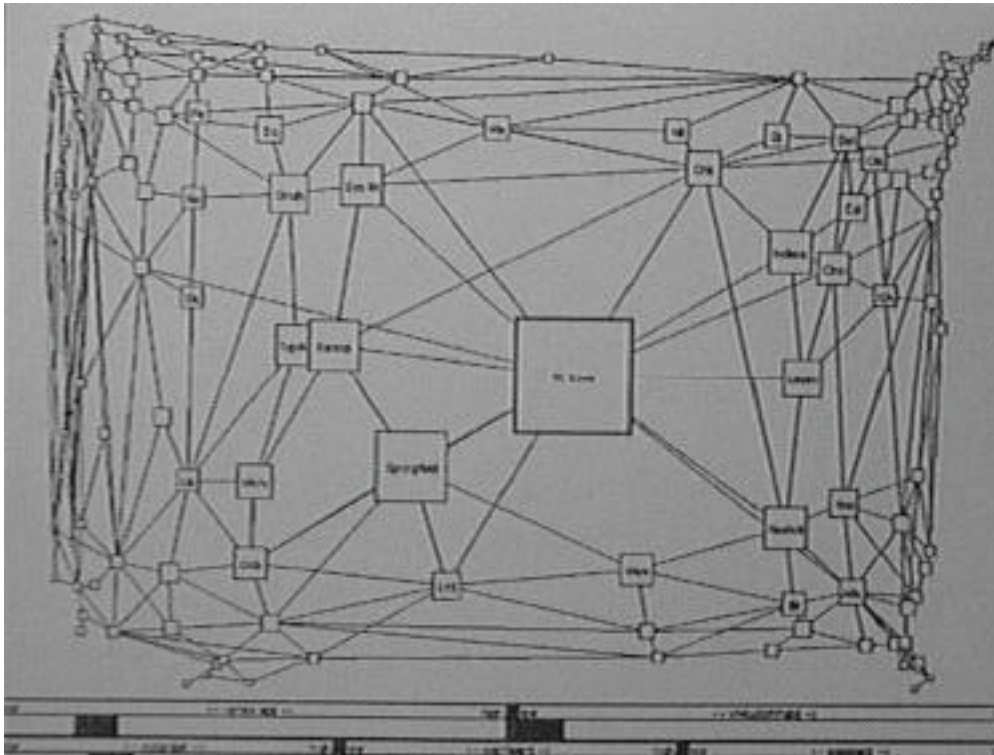


Figure 7 : Vue en œil de poisson sur une carte des Etats-Unis

La vue en œil de poisson peut être considérée comme une suite de modalités concentriques. A la figure 7, nous dégagons deux types de modalités :

- ? une modalité en forme de disque qui couvre la région de St Louis qui est partielle, précise et déformée ;
- ? un ensemble de modalités en anneaux concentriques (autour du disque) qui sont partielles, vagues et avec une déformation de la taille des éléments.

Comme pour le mur en perspective, nous caractérisons la somme des modalités envisagées pour construire une modalité globale. La continuité visuelle est respectée car le système n'offre aucune discontinuité spatiale et le changement de centre d'intérêt se fait de façon progressive. Toutefois, le centre d'intérêt et la structure de l'espace sont déformés.

#### 4.5. Technique de « zoom »

La technique du zoom permet de s'éloigner de l'ensemble des informations pour changer de centre d'intérêt puis de se rapprocher à nouveau comme dans le système ProjectX [PROJECT X] de la figure 8. Il permet de se situer mais ne permet pas la visualisation simultanée du détail et du contexte. A l'inverse de la vue en œil de poisson dans laquelle c'est l'information qui se déforme pour mettre en évidence le centre d'intérêt, ici c'est l'utilisateur qui se déplace. L'utilisateur se déplace dans l'espace d'informations pour se rapprocher de son centre d'intérêt ou pour s'en éloigner et obtenir une vue globale.

**Figure 8** : Technique de zoom, ProjectX

La technique de zoom peut être vue comme une superposition de modalités globales, partielles et sans déformation avec une modalité de premier plan qui est partielle, précise, et sans déformation. Nous pouvons constater que :

- ? Il existe une modalité globale, vague et sans filtre mais cette modalité n'est pas présente à tout moment.
- ? Il existe des modalités partielles, précises et sans filtres mais qui ne sont pas toujours présente.

La combinaison de modalités engendrées par la technique de zoom fait qu'il n'est pas possible d'avoir au même instant une vue globale et une vue précise. La modalité précise et la modalité globale ne sont pas présentes simultanément. La continuité temporelle est vérifiée puisque le passage de la modalité précise à la modalité globale (et vice-versa) passe par un grand nombre d'étapes intermédiaires.

#### **4.6. Arbre conique**

La technique du sapin ou de l'arbre conique, présentée à la figure 9, vise à ranger les éléments non pas linéairement mais sur un carrousel composé de cônes [Robertson 91]. Vus de profil, les cônes peuvent être tournés pour amener l'information pertinente au premier plan. L'information ainsi ramenée peut être un autre cône sur lequel l'utilisateur recommence l'opération. Il est possible de demander une information qui ne soit pas directement sur le cône au premier plan. Le carrousel tourne alors tous les cônes nécessaires pour amener l'information désirée au premier plan.

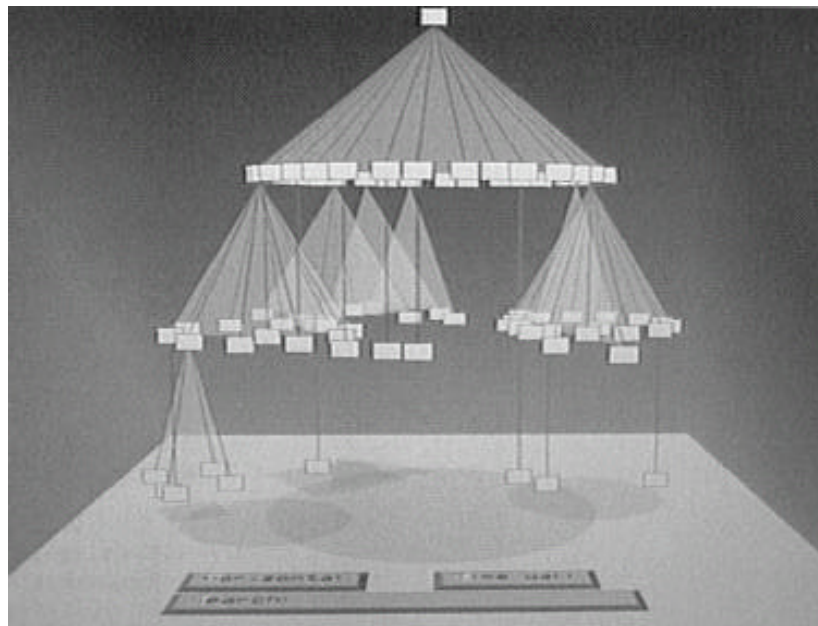


Figure 9 : L'arbre conique sur une arborescence de fichiers

La décomposition en modalités est assez compliquée pour cette technique d'interaction. La règle intéressante à vérifier pour ce système est la continuité visuelle temporelle qui intervient lorsque l'on change de centre d'intérêt. Tous les cônes du carrousel tournent sur leurs axes assez lentement pour que l'on puisse observer les changements dans l'espace d'informations. La continuité visuelle est donc bien respectée pour cette technique.

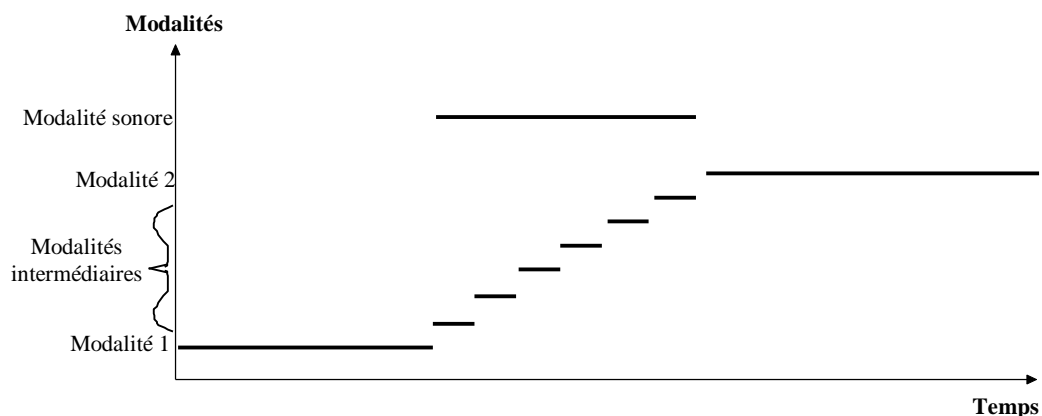
## 5. Représentation multiple

Afin de concrétiser notre approche de travail, nous proposons une solution qui réponde aux critères de sélection et de composition proposées. Nos règles de conception ne constituent pas un moteur de génération de présentations multimodales qui répondent à tous les problèmes de choix des modalités de sortie. La meilleure solution est alors de laisser le choix à l'utilisateur des modalités de sortie. L'utilisation des modalités se fera donc séquentiellement et laissera l'utilisateur choisir à quel moment il désire changer de modalité. Le critère qu'il nous semble très important de vérifier dans ce cas là, est la continuité visuelle temporelle lors du changement. Nous allons maintenant décrire en détail la conception de cette technique puis nous énoncerons les règles de conception que nous avons dégagées.

### 5.1. Principes de la représentation multiple

Tout d'abord nous considérons les modalités qui interviennent dans la représentation multiple. Chaque modalité, que nous concevons, est construite à partir d'une modalité globale et d'une modalité précise. D'autre part pour garantir la continuité visuelle, nous construisons des modalités intermédiaires lors du passage d'une modalité à une autre.

Le constructeur de modalités que nous proposons s'applique aux modalités graphiques. L'approche consiste à utiliser une succession de modalités graphiquement proches. Cette technique de fondu enchaîné est appelée « morphing » en anglais lorsqu'elle s'applique à une animation d'images. Elle assure que la continuité visuelle est respectée dans le temps. Le passage entre deux modalités utilisées séquentiellement est ainsi plus souple qu'un passage brusque (ou instantané). Nous préconisons toutefois l'utilisation d'une modalité sonore additionnelle pour délimiter le laps de temps du fondu afin de mettre en évidence l'instant où la représentation n'est pas stable.



**Figure 10 :** Utilisation des modalités lors du passage continue

Cette approche soulève un certain nombre de problèmes de mise au point : choisir la modalité de départ, configurer les modalités, etc. Afin de pouvoir répondre aux interrogations soulevées par ces problèmes, nous avons réalisé un système qui implémente la représentation multiple. La description de ce logiciel fera l'objet du chapitre suivant.

## 5.2. Règles de conception

Nous étudions ici la représentation multiple d'une grande quantité d'informations afin d'améliorer l'utilisabilité de l'interface. Aussi la représentation multiple ne doit pas augmenter la complexité d'utilisation. Pour cela nous définissons cinq règles de conception :

- ? l'existence d'une représentation de référence (déjà connue des utilisateurs et sans déformation) dans le choix des modalités,
- ? l'existence d'au moins une modalité construite qui allie précision et présentation globale de l'espace d'information,
- ? la continuité visuelle lors du passage d'une modalité à une autre,
- ? la facilité de changer de modalité,
- ? la réutilisabilité des outils de navigation.

---

Les deux premières règles dérivent de la description du système faite au paragraphe précédent. Nous avons affiné la règle « centre d'intérêt + contexte » en tenant compte du fait qu'une modalité ne doit pas forcément servir à toutes les tâches. Cette règle n'impose donc pas que toutes les modalités construites soient précises et globales.

La continuité, qu'elle soit visuelle, auditive ou encore tactile, est liée à la navigation. Il est crucial que l'utilisateur ne soit pas perdu dans l'espace suite à un changement de visualisation ("Où j'étais, où je suis, où je vais"). Il est à noter que la continuité est aussi utilisée lors de la navigation pour aider l'utilisateur à percevoir le chemin parcouru. Par exemple dans [Robertson 91] lorsque l'utilisateur sélectionne un noeud, le carrousel et son ombre s'animent pour amener le noeud sélectionné au premier plan. Nous appliquons ici ce principe de continuité lors du passage d'une visualisation à une autre.

Le changement de modalité définit une tâche articulatoire qui ne modifie pas l'état interne du système. Il convient donc que cette tâche soit facile à accomplir. De plus l'interface doit inciter (principe d'« affordance ») l'utilisateur à changer de modalité selon sa tâche afin d'exploiter au mieux la représentation multiple.

Notre dernière règle définit des contraintes de cohérence au niveau de l'interaction. Afin d'une part de favoriser l'apprentissage et d'autre part d'augmenter la prévisibilité, les outils de navigation (moyens de locomotion) doivent être cohérents d'une modalité à une autre. Cette règle incite en fait à factoriser les concepts de navigation identiques aux différentes représentations afin de les représenter par des objets d'interaction communs.

## **6. Conclusion**

Dans ce chapitre nous avons étudié les apports possibles de la multimodalité au rendu d'informations. Le principal problème est, nous l'avons vu, le choix des modalités. Nous avons pour cela été amené à proposer une caractérisation des modalités spécifiques au rendu de grands espaces d'informations. Les caractéristiques dégagées nous ont permis de proposer plusieurs règles de conception que nous avons illustrées. Enfin nous avons approfondie un concept de la multimodalité : la représentation multiple. Cette étude de cas amené à affiner nos règles de conception. Le chapitre suivant décrit le prototype VITESSE qui implémente la représentation multiple.





## **Chapitre IV**

*Systeme VITESSE*

Le système VITESSE est un système de visualisation d'informations. Nous avons donc dissocié le système de visualisation et le noyau fonctionnel. Nous avons implémenté, à titre d'exemple, un noyau fonctionnel qui nous permette de connecter VITESSE aux moteurs de recherche du World Wide Web. VITESSE est donc relié par le réseau à ces systèmes de recherche d'informations mais pourrait servir à visualiser d'autres types d'informations. Pour chaque réponse retournée par le moteur de recherche, VITESSE charge la page Web correspondante. VITESSE récolte ainsi d'autres informations que celles initialement proposées par le moteur de recherche. Nous récoltons donc un grand espace d'informations en provenance d'Internet.

Chaque page Web fait l'objet d'un rectangle dans la page de présentation. Les rectangles sont disposés en « escargot » dont le centre est occupé par les réponses les plus pertinentes et la périphérie par les moins pertinentes. Cette représentation est arbitraire au sens de Bernsen. Le paragraphe 1 présente la spécification de l'interface de VITESSE. Nous définissons l'interface de notre système de représentation multiple au paragraphe 1.1 et les outils de navigation au paragraphe 1.2. Nous terminons notre présentation de l'interface par le paragraphe 1.3 où nous expliquons les règles de conception énoncées au chapitre précédent que nous avons appliquées.

Le deuxième paragraphe est consacré à la réalisation logicielle de VITESSE. Nous avons utilisé le modèle d'architecture PAC-Amodeus pour organiser les composants de notre logiciel. Nous présentons donc d'abord l'architecture globale au paragraphe 2.1 et nous détaillerons ensuite le composant principal (appelé contrôleur de dialogue) au paragraphe 2.2. Nous concluons ce chapitre par un compte rendu de notre expérience avec le langage JAVA pour implémenter une architecture selon le modèle PAC-Amodeus. Comme au paragraphe précédent, nous séparerons l'architecture globale et le contrôleur de dialogue.

## **1. Conception de l'interface**

Afin d'analyser l'interface du système VITESSE [VITESSE], nous avons découpé la conception de l'interface en deux parties : le système de représentation de la grande quantité d'informations et les outils de navigation qui permettent de manipuler l'espace d'informations.

### **1.1. Systèmes de représentation**

Nous avons choisi de représenter l'espace d'informations selon différentes modalités. Ce choix était tout d'abord motivé par la comparaison que nous voulions effectuer des différentes possibilités de vue en œil de poisson. Nous avons choisi de représenter l'information dans un espace restreint. Cette contrainte se justifie par la grande diversité des écrans mais aussi par la difficulté des systèmes à gérer des mémoires vidéo de grande taille. La taille retenue est un carré de 600 pixels de côté.

---

Le système VITESSE implique donc l'utilisation d'un écran dont la résolution est supérieure à 640\*480 pixels. Le contenu de cette zone d'écran est donc variable mais pas le contenant qui doit rester de taille fixe pour mieux pouvoir comparer l'utilisabilité des différentes techniques d'interaction. Parmi les techniques d'interaction ou modalités utilisées, nous en distinguons une que nous considérons comme étant de référence.

### *1.1.1. Technique de référence*

Afin de tester l'utilisabilité de VITESSE, nous avons choisi une technique de référence. Nous avons choisi une technique qui est sans doute beaucoup plus répandue que les vues en œil de poisson : une technique utilisant le zoom et les barres de défilement. Cette technique est la seule que nous ayons implémentée et qui soit non-déformante. Nous prenons comme hypothèse que l'espace d'informations est constitué d'éléments organisés dans une structure de graphe [Furnas 97]. L'hypothèse d'existence d'éléments implique parfois un découpage de l'espace d'informations. Par exemple, on peut considérer une carte géographique comme un ensemble de pixels reliés avec les pixels voisins. Dans VITESSE, afin de rester général, nous avons décidé d'encapsuler les éléments d'information dans des rectangles à l'écran. La technique de Zoom (Figures 11a et 11b) permet de zoomer sur une dizaine de rectangles en ayant une vue précise du contenu (Figure 11a) ou bien d'obtenir une vue globale (Figure 11b) des 50 ou 100 rectangles presque sans aucune information sur le contenu. Un grand nombre de positions intermédiaires permettent à l'utilisateur d'utiliser le niveau de détail requis pour la tâche qu'il entreprend.

**Figure 11a :** La technique de référence, vue précise

**Figure 11b :** La technique de référence, vue globale

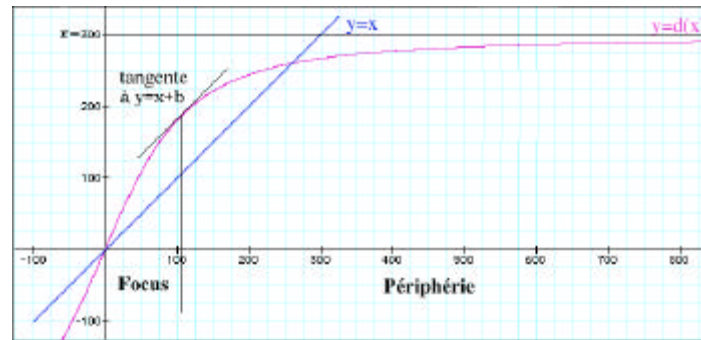
### *1.1.2. Techniques déformantes*

Les trois premières techniques déformantes qui ont été implémentées correspondent à celles proposées par [Sarkar 92] (Figure 13a et 15b). Nous avons ensuite développé d'autres techniques telles que la représentation en cylindre (Figure 14a et 14b) ou la représentation sous forme de bosse (Figure 13b) qui sont dérivées des techniques de [Sarkar 92] et de [Leung 94]. Pour réaliser ces techniques d'interaction, nous utilisons des fonctions de déformation que l'on applique à l'espace initial présenté à la Figure 11a et 11b. Cette approche ne va pas à l'encontre du critère d'honnêteté, à condition que l'utilisateur puisse effectuer les corrections nécessaires afin d'obtenir une carte mentale conforme à la structure des données. La vérification

du critère d'honnêteté fait appel à des théories de psychologie de la perception et de la cognition.

### 1.1.2.1. Fonction de déformation

La fonction de déformation,  $d$ , présentée à la figure 12 constitue le point de départ de notre conception.



$$d(x) = x \cdot \frac{\sqrt{(x^2 + z^2)(r^2 + (z - o)^2) + z^2(z - o)^2} + z(z - o)}{x^2 + z^2}$$

$$d^{-1}(x) = \frac{z \cdot x}{\sqrt{r^2 + x^2 + (z - o)^2}}$$

**Figure 12 :** Fonctions de déformation.

Dans les équations de la Figure 12,  $x$  représente la valeur de la distance en pixels que l'on cherche à déformer, tandis que  $r$ ,  $o$  et  $z$  sont les paramètres de la transformation.

$r$  désigne la valeur limite de la fonction. Par exemple, à la Figure 12, la fonction de déformation  $d(x)$  est comparée à l'identité avec  $r$  fixé à 300 pixels.  $r$  fixe donc la taille de la visualisation à l'écran : ainsi dans la visualisation sphérique de la Figure 13-b,  $r$  désigne le rayon de la sphère.  $o$  et  $z$  sont les paramètres qui déterminent la puissance de l'effet de "loupe" [Sarkar 92]. A la Figure 12,  $o$  et  $z$  ont respectivement la valeur 150 et 112 pixels.

A partir de la fonction de déformation,  $d(x)$ , nous calculons la fonction inverse  $d^{-1}(x)$  (Figure 12) nécessaire à l'interaction : en effet  $d^{-1}$  permet de calculer la position des clics-souris dans l'espace non déformé.

Nous expliquons d'abord comment appliquer la fonction  $d$  pour définir plusieurs représentations de l'espace puis introduisons de nouvelles fonctions de déformation obtenues par combinaison de fonctions.

### 1.1.2.2. Calcul de la représentation à partir de la fonction de déformation

Comme le montre la Figure 13, la fonction de déformation peut être appliquée aux coordonnées cartésiennes  $(x,y)$  (Figure 13a) des emplacements de l'espace initial

ou encore aux coordonnées polaires  $(r, \theta)$  (Figure 13b). Dans ce dernier cas, l'angle est conservé et seule la distance  $r$  est transformée. Intuitivement cette visualisation (Figure 13-b, coordonnées polaires) nous semble mieux satisfaire le critère d'honnêteté en vue de l'analogie avec une demi-sphère. Elle s'apparente à une métaphore du monde réel, comme lorsque l'on fait tourner un globe. A l'opposé la transformation cartésienne conserve l'alignement en  $x$  et en  $y$  de l'espace initial. Ainsi si l'espace de départ contient des colonnes, celles-ci seront perçues par l'utilisateur.

**Figure 13a** : Représentation en  
œil de poisson cartésienne

**Figure 13b** : Représentation en  
œil de poisson polaire

De nouvelles représentations sont obtenues en combinant les deux approches précédentes. Par exemple la Figure 14a représente l'espace déformé obtenu en appliquant une transformation cartésienne sur l'axe des abscisses et une transformation polaire sur l'axe des ordonnées. On obtient alors une représentation qui offre les avantages des deux approches : métaphore du monde réel tout en respectant l'alignement en  $x$  et  $y$  de l'espace de départ.

**Figure 14a** : Représentation en cylindre vu de  
l'extérieur

**Figure 14b** : Représentation en cylindre vu de  
l'intérieure

Une autre approche consiste à appliquer la fonction de transformation et sa réciproque. Bien que la fonction réciproque aille à l'encontre de l'observabilité du centre d'intérêt (effet de loupe inverse), sa combinaison avec la fonction de transformation offre l'avantage de rendre plus observable le voisinage du centre d'intérêt. L'exemple de la Figure 14b, est obtenu en combinant  $d$  (cartésien) sur l'axe des abscisses et une fonction dérivée de  $d^{-1}$  (polaire) sur l'axe des ordonnées. Le centre d'intérêt est présenté de façon moins précise qu'à la Figure 14a, mais son voisinage est plus observable.

### 1.1.2.3. Combinaison de fonctions de déformation

A partir de deux fonctions de transformation  $d$  et  $d'$ , nous définissons une nouvelle fonction,  $d_{\text{nouv}}$ , en les combinant selon une fonction caractéristique  $c$  dont le domaine de valeurs est  $[0,1]$  :

$$d_{\text{nouv}} = c(x)d(x) + (1-c(x))d'(x)$$

Pour éviter une rupture visuelle (fonction de déformation discontinue), deux solutions sont possibles :

?  $c$  est continue.

?  $c$  est discontinue (valeurs 0 ou 1) et passe de 0 à 1 au niveau de l'abscisse,  $V$ , du point d'intersection des deux fonctions  $d$  et  $d'$ . La nouvelle fonction se définit alors comme suit :

$$d_{\text{nouv}} = \text{si } (x < V) \text{ alors } d(x) \text{ sinon } d'(x)$$

Nous combinons deux fonctions de transformation pour obtenir les avantages des deux visualisations. Par exemple à la Figure 15a, la sphère tronquée est obtenue en combinant la fonction identité ( $\text{Id}$ ) et la fonction  $d$  de la Figure 12 :

$$d_{5a} = \text{si } (x < V) \text{ alors } \text{Id}(x) \text{ sinon } d(x)$$

La visualisation en bosse de la Figure 15b est la fonction duale de la fonction précédente :

$$d_{5b} = \text{si } (x < V) \text{ alors } d(x) \text{ sinon } \text{Id}(x)$$

**Figure 15a** : Représentation en sphère tronquée

**Figure 15b** : Représentation sous forme de bosse

Nous avons montré comment définir plusieurs représentations d'un même espace à partir d'une fonction de déformation ou de la combinaison de deux fonctions. La validation de ces visualisations ne pourra se faire que par expérimentation avec des utilisateurs.

---

Généralement, les techniques déformantes transforment les rectangles initiaux en polygones quelconques. Il devient alors difficile de les remplir avec de l'information (textes ou images) car les techniques d'affichage sont alors très coûteuses en temps de calcul. Nous avons donc choisi de ne pas déformer le texte. De plus les techniques de déformation risquent de rendre le texte illisible (« aliasing ») et les techniques pour empêcher ce problème (« anti-aliasing ») sont aussi coûteuses en temps de calcul. Nous proposons d'inscrire les informations les plus importantes en priorité à l'intérieure de l'espace délimité par les polygones. Cette politique va à l'encontre de l'effet de perspective produit par certaines techniques (par exemple la transformation sphérique). La propriété du texte de toujours rester « droit » peut être considéré comme une propriété « magique » [Dieberger 94] de la métaphore de la perspective.

### *1.1.3. Passage d'une représentation à une autre*

L'une des propriétés intéressante de VITESSE est la possibilité de changer de représentation en cours d'utilisation. Le passage s'opère sans discontinuité visuelle et se combine avec une modalité sonore afin de rendre compte de la criticité du passage d'une représentation à une autre. Pour sélectionner une représentation, l'utilisateur doit pouvoir :

- ? visionner très rapidement les visualisations qui lui sont offertes,
- ? les reconnaître facilement,
- ? savoir quelle représentation est actuellement utilisée.

Pour répondre à ces critères, nous avons choisi d'implémenter des onglets contenant une icône de la technique de représentation. L'onglet représentant la technique en cours est plus sombre et plus enfoncé que les autres qui apparaissent comme des boutons. Nous pensons que les boutons incitent plus à être enfoncés que des onglets en arrière plan. Cette partie de l'interface peut être considéré comme faisant partie des outils de navigation. Toutefois le couplage avec le système de représentation multiple est si étroit qu'il n'est pas possible de l'en dissocier.

## **1.2. Outils de navigation**

Afin de naviguer dans un des espaces d'information, l'utilisateur doit pouvoir le manipuler. VITESSE offre les mêmes outils de navigation quelle que soit la technique de représentation courante. Nous espérons que cette mise en commun permet un meilleur apprentissage du système pour les utilisateurs qui n'ont pas l'habitude des systèmes de navigation. Avant de présenter chaque objet d'interaction qui intervient dans l'interface (paragraphe 1.2.2. et 1.2.3.), nous introduisons les deux styles possibles de manipulation : la manipulation directe et la manipulation indirecte.

### *1.2.1. Manipulation directe et indirecte de l'espace d'information*

La manipulation indirecte s'oppose à la manipulation directe par l'apparition d'un intermédiaire. Dans le cas des systèmes informatiques, l'intermédiaire est le système qui reçoit des ordres de l'utilisateur selon un langage et manipule l'espace d'informations en conséquence. Ce paradigme d'interaction implique un apprentissage plus difficile dû à la barrière linguistique mais souvent augmente l'efficacité car le plan d'actions revient au système contrairement à la manipulation directe. Nous avons relevé trois types de manipulation indirecte pour les grandes quantités d'informations :

- ? L'utilisateur ne manipule pas l'information mais sa position dans l'espace d'informations (un simulateur d'avion par exemple). Cette technique est surtout utilisée par le biais des touches du clavier mais peut aussi s'appliquer à la souris.
- ? Le pointeur de la souris indique quelle est la zone à agrandir et le système agrandit cette zone de l'espace d'informations. Il y a donc une association entre pointeur de la souris et centre d'intérêt de l'utilisateur.
- ? L'utilisateur sélectionne sa zone d'intérêt et le système déplace l'espace d'informations de façon à faire apparaître cette zone dans le centre d'intérêt. Cette technique nécessite parfois plusieurs actions pour obtenir l'effet souhaité.

La manipulation directe propose à l'utilisateur de manipuler directement les objets de l'espace d'informations comme il le ferait avec des objets du monde réel. L'utilisateur peut ainsi déplacer les objets avec la souris. L'avantage de ce paradigme d'interaction est la facilité d'apprentissage due aux ressemblances avec les concepts du monde réel (par exemple prendre un objet et le glisser dans la corbeille).

Nous avons choisi d'utiliser la manipulation directe pour la facilité d'apprentissage. Dans VITESSE, l'utilisateur peut ainsi déplacer l'information en sélectionnant un point d'accroche puis en déplaçant l'espace d'informations. Cette technique d'interaction implique de nombreux déplacements avec la souris mais offre une grande liberté d'action à l'utilisateur. Nous notons néanmoins trois défauts à cette technique :

- ? Il est précis au « pixel près » alors que l'intérêt de l'utilisateur se porte sur les informations.
- ? Il n'est pas toujours possible d'effectuer le déplacement voulue en une seule fois en particulier lorsque l'espace n'est pas entièrement visible à l'écran.
- ? Il est seulement possible de prendre en compte deux dimensions. Cette limite est due aux limitations des dispositifs de pointage utilisés.



---

### 1.2.2. Objets d'interaction supplémentaires

Compte tenu des problèmes soulevés au paragraphe précédent, nous avons été amené à proposer d'autres outils de navigations (Figure 16) pour palier aux déficiences de la manipulation directe :

- ? quatre boutons (droite, gauche, haut et bas) dédiés au déplacement de l'espace d'informations : le déplacement est exactement égal à la taille d'un élément.
- ? deux ascenseurs permettant de se déplacer directement d'un bord à l'autre de l'espace d'informations.
- ? des boutons de zoom avant et zoom arrière (lorsque c'est nécessaire pour la technique d'interaction).

**Figure 16 :** Outils de navigation dans VITESSE

Enfin nous proposons deux autres types d'outils de navigation. Le premier consiste en des raccourcis claviers des fonctions correspondant aux boutons. Le dernier outil de navigation consiste en un bouton de recentrage qui permet de revenir à l'état de départ, c'est à dire au centre de l'espace d'informations.

### 1.2.3. Objets pour paramétrer la technique de visualisation

Nous avons été amené à introduire deux autres outils pour le confort d'utilisation du système.

Le premier d'entre eux est une barre de défilement « à tout faire » qui sert dans quatre des sept représentations. Cet objet d'interaction sert à modifier la frontière entre deux modalités comprises dans la même représentation. Dans certains cas elle peut avoir un comportement ressemblant à celui du zoom (métaphore de la sphère tronquée).

Le second outil est réservé aux utilisateurs avertis. Nous avons donc choisi sciemment d'en rendre l'accès difficile. Il s'agit d'une fenêtre contenant des barres de défilement qui modifient les paramètres de la fonction mathématique utilisée pour les déformations. L'activation de cette fenêtre se fait en double-cliquant sur le fond de l'espace d'informations (l'espace gris entre les rectangles contenant les informations).

### **1.3. Propriétés vérifiées**

Ce paragraphe montre dans quelle mesure les principaux critères proposés au chapitre III sont vérifiées dans VITESSE.

Le système VITESSE a été créé pour vérifier les critères de continuité spatiale et temporelle. En utilisant la technique de fondu enchaîné, le critère de continuité temporelle se trouve vérifié. Le critère de continuité spatiale est également vérifié puisque nous nous avons implémenté et combiné des techniques de vues en œil poisson, reconnues pour leur continuité visuelle.

VITESSE vérifie le critère de coexistence de plusieurs modalités globales et d'au moins une modalité précise puisque il propose sept modalités dont six intègrent un modalité globale et une modalité précise simultanément.

Le critère de facilité de changement de la représentation est vérifié dans VITESSE depuis que nous avons remplacé les cases à choix multiples par des boutons contenant une icône de la technique de présentation. En effet la correspondance entre la technique employée et le bouton est beaucoup plus facile à établir qu'avec une case à cocher étiquetée d'un nom scientifique.

La réutilisabilité des outils de navigation est très grande dans VITESSE puisque à part la technique de référence (paragraphe 1.1.1), toutes les autres techniques utilisent les mêmes outils de navigation décrits au paragraphe 1.2.2.

Nous pouvons également remarquer que VITESSE vérifie le critère d'adaptabilité puisque les utilisateurs experts peuvent accéder aux paramètres des techniques de déformation en double cliquant sur le fond de l'espace d'informations. Par ailleurs, tous les utilisateurs peuvent faire varier le paramètre qui définit la limite entre les modalités lorsque c'est pertinent (barre de défilement étiquetée « truncating »).

## **2. Réalisation logicielle**

Même si nous n'avons pas les contraintes de qualité des produits commerciaux, nous devons produire un code modifiable et portable. Ces exigences nous ont incités à mener une étude sur l'architecture globale de VITESSE.

### **2.1. Architecture logicielle globale**

Le premier critère de qualité à vérifier est la portabilité vis à vis du moteur de recherche. VITESSE doit être indépendant du moteur. Ce critère de qualité nous a amené à considérer le modèle d'architecture ARCH [ARCH].

Afin d'atteindre une grande modularité du code qui participe à la modifiabilité du code, nous avons décidé d'appliquer le modèle PAC-Amodeus [Nigay 94] qui allie les avantages du modèle ARCH à celui d'une approche multi-agent, hautement modulaire.

### 2.1.1. Architecture globale

Comme le montre la Figure 17, PAC-Amodeus préconise cinq composants logiciels : le Contrôleur de Dialogue (CD) composant principal est organisé en une hiérarchie d'agents PAC.

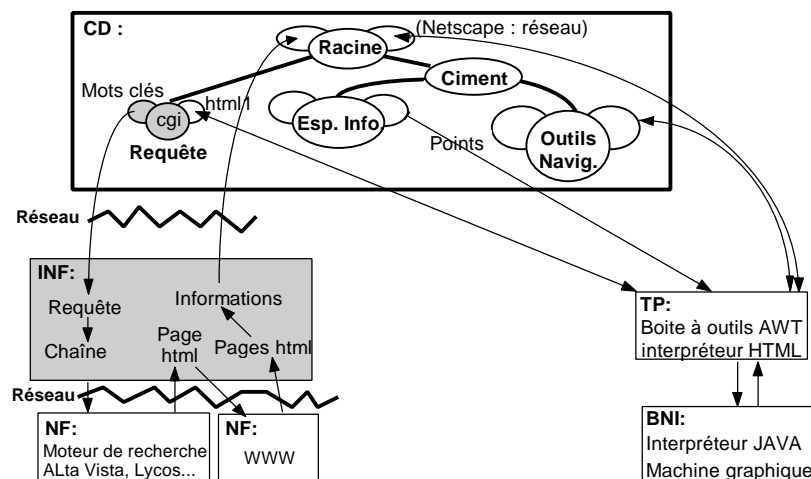


Figure 17 : Architecture globale du système VITESSE.

Le Noyau fonctionnel (NF) implante les concepts du domaine applicatif : le moteur de recherche. C'est le rôle de l'Interface avec le Noyau Fonctionnel (INF) de communiquer avec le NF par des envois de requête. L'INF sert d'adaptateur et permet de rendre le Contrôleur de Dialogue indépendant du moteur de recherche utilisé. De plus l'INF peut transformer les informations reçues du NF : par exemple l'INF vérifie la validité des réponses (liens inexistant) et calcule le nombre d'occurrences des mots de la requête pour chaque réponse.

A l'opposé de l'arche, le composant bas niveau d'interaction (BNI) désigne la plate-forme d'accueil logicielle et matérielle. Il contient en particulier l'interpréteur JAVA. Aucun code supplémentaire n'a été nécessaire pour ce composant. Le composant technique de présentation (TP) définit les règles de correspondance entre objets de présentation et objets d'interaction. Il comprend en particulier la boîte à outils AWT. Ce composant vise à rendre le contrôleur de dialogue indépendant des outils de développement.

Le Contrôleur de Dialogue, clé de voûte du système, prend en charge l'enchaînement des tâches et gère chaque fil de dialogue grâce à une hiérarchie d'agents PAC. Notre technique d'interaction à plusieurs facettes est donc réalisée dans le CD sous la forme d'une hiérarchie d'agents. Avant de détailler cette hiérarchie, il est important de noter que chaque agent est indépendant du Noyau Fonctionnel et des outils de développement. Ainsi le code de nos techniques d'interaction est réutilisable et ne dépend pas ni d'un moteur de recherche, ni d'une boîte à outils.

### *2.1.2. Réalisation des Techniques d'Interaction : le contrôleur de dialogue*

Pour définir la hiérarchie d'agents PAC du contrôleur de dialogue, nous appliquons les règles heuristiques de mise en oeuvre présentées dans [Nigay 94]. Nous appliquons la règle « vue multiple » qui permet de réaliser le critère de représentation multiple. Cette règle s'énonce comme suit :

? chaque technique de présentation est gérée par un agent PAC, « Visualisation »,

? un agent père « Vue Multiple » assure la cohérence entre les vues.

Nous obtenons la hiérarchie de la Figure 18. L'abstraction de l'agent « Vue Multiple » maintient l'espace des informations non déformé. Sa présentation définit la palette qui permet de changer de visualisation. Son contrôle maintient le type de visualisation courante sélectionnée.

Chaque agent « Visualisation » contient dans son contrôle la fonction de déformation. L'abstraction maintient l'espace déformé selon la fonction. Les données ne sont pas dupliquées, seules les positions des emplacements sont maintenues. La présentation gère la visualisation. A un instant donné, un seul agent « Visualisation » est créé. Au changement de visualisation, un agent « Animation » qui est éphémère est créé dynamiquement pour gérer le passage d'une visualisation à l'autre. Cet agent « Animation » reçoit de son agent père « Vue Multiple » l'espace déformé courant et la fonction de déformation transitoire à appliquer. Après  $n$  affichages ( $n=10$  dans VITESSE), cet agent renvoie l'espace déformé final à l'agent « Vue Multiple ». Ce dernier détruit l'agent « animation » et crée alors dynamiquement un nouvel agent « Visualisation2 » en lui fournissant sa fonction de déformation et l'espace courant déformé.



JAVA a été le premier langage intégré dans les butineurs populaires tels que Netscape ou Internet Explorer. Grâce aux principes de modularité des architectures, ces logiciels se sont par la suite développés vers des systèmes de « plugs-in » qui permettent aujourd'hui d'intégrer d'autres types de langages dans un butineur. Le mixage entre World Wide Web et code exécutable est plus ancien que l'apparition des langages comme JAVA puisque dès les premières versions, les butineurs pouvaient faire appel à l'exécution de programmes sur la machine serveur. L'innovation apportée par JAVA est l'exécution locale (sur la machine de l'utilisateur) de l'application qui permet une véritable interactivité. JAVA et ses concurrents tels que Tcl-Tk sont des langages qui sont interprétés par une Machine Virtuelle (VM).

Toutefois, les programmes exécutés sur le serveur, appelés CGI-BIN, continuent d'exister en parallèle des applications JAVA aussi dénotées applets. Dans le paragraphe suivant, nous montrons comment les composants d'une application en JAVA peuvent s'organiser selon le modèle d'architecture ARCH puis nous étudions ensuite en détail le composant contrôleur de dialogue.

### 2.2.1. Architecture générale et JAVA

Nous avons vu l'architecture possible d'une application Web avec le logiciel VITESSE. La figure 20 propose une généralisation de l'architecture de VITESSE s'appuyant sur les composants de ARCH et qui puisse s'appliquer à tout type d'applet. Cette architecture permet d'identifier les composants existants (VM JAVA) et les supports à utiliser pour créer les composants manquants (un CGI-BIN pour créer un accès au réseau). Enfin cette architecture permet de critiquer le mécanisme puisqu'elle met en évidence certains points litigieux.

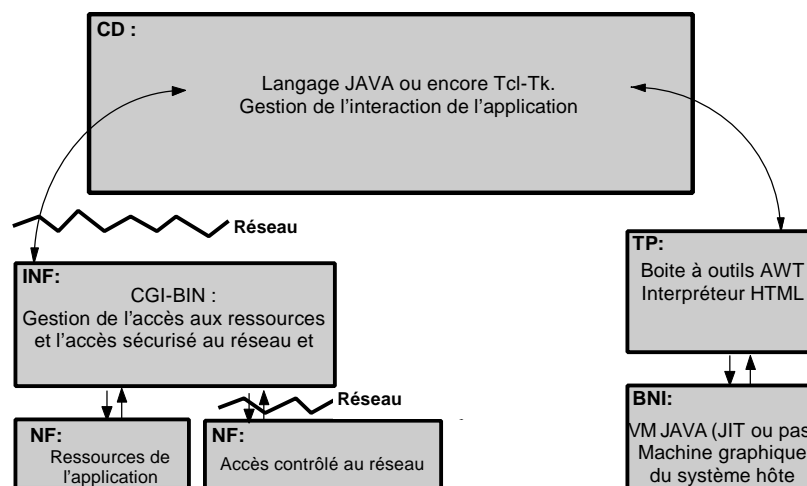


Figure 20 : Contexte d'un applet JAVA.

Les critiques qui peuvent être faites à ce genre de systèmes sont les suivantes :

- ? La boîte à outil AWT se trouve dans le composant Techniques de présentation (TP). Or les objets d'interactions existants sur les différentes machines n'ont pas le

---

même comportement. D'une part, pour l'utilisateur, il est intéressant de garder les objets d'interaction habituels. D'autres part le concepteur doit parfois traiter des comportements différents selon le système hôte (BNI).

- ? La nécessité de passer par la machine serveur pour se connecter au réseau Internet ralentit les connexions.
- ? L'impossibilité de maintenir des ressources sur le disque dur de la machine locale oblige les systèmes qui suivent l'architecture de la figure 20 à mettre ses ressources dans le noyau fonctionnel alors que certaines ressources ne dépendent parfois que de l'interface.

Ce type d'application, bien que très prometteur par certains aspects, est encore jeune et nécessite d'atteindre une certaine maturité. En terme d'interface, le mécanisme d'application intégrée dans une page Web nous libère de nombreuses contraintes de compatibilité.

### 2.2.2. Architecture du contrôleur de dialogue et AWT

Le contrôleur de dialogue de la figure 20 prône l'utilisation de JAVA ou de tout autre langage de script pour son implémentation. Notre expérience du langage JAVA nous amène maintenant à voir comment concilier la boîte à outils AWT et un modèle à agent comme PAC.

Nous avons vu comment les agents PAC se prêtent à une structuration arborescente. JAVA intègre aussi une arborescence pour les objets d'interaction. Ces composants appelés « layout » reposent sur la structure spatiale des objets d'interaction. Chaque layout est délimité par une surface rectangulaire qui contient d'autres « layouts », noeuds de l'arbre, ou bien des objets d'interaction (boutons, barres de défilement, etc.) qui correspondent aux feuilles de l'arbre. JAVA propose plusieurs types de « layout » ainsi que la possibilité de créer soi-même de nouveaux layout. Afin d'illustrer l'utilisation des layout, la figure 21 montre quelques exemples appliqués à un ensemble de boutons. Nous constatons que l'arborescence d'objets d'interaction correspond partiellement à celle des agents PAC du contrôleur de dialogue.

**Figure 21** : « Layouts » JAVA

Un aspect important des agents PAC est la décomposition en composant de Présentation, d'Abstraction et de Contrôle. La réalisation logicielle d'un agent PAC en terme d'objets peut se faire de plusieurs manières :

- ? Nous pouvons créer trois objets par agent PAC, un par composant.
- ? nous pouvons implémenter un agent PAC par un seul objet.

Nous avons choisis cette dernière solution qui est motivée, entre autre, par un inconvénient du système des applets incorporées dans une page Web : les connexions. En effet le butineur qui intègre l'application ouvre une connexion pour chaque objet et chaque ressource de l'application. Multiplier les objets revient donc à multiplier les connexions or pour chacune d'elles il faut payer le prix d'un temps dit de « latence » qui pénalise l'interactivité finale du système. Il convient donc de garder un nombre d'objets inférieur au seuil critique d'une centaine d'objets.

### **3. Evaluation et améliorations futures**

VITESSE nous a permis d'identifier les problèmes de réalisation logicielle inhérents à la représentation multiple d'un grand espace d'informations. Nous avons également pu mettre en évidence des résultats sur la navigation comme le choix d'une manipulation directe ou indirecte. Nous avons pu faire cohabiter des techniques de visualisation existantes. VITESSE peut ainsi servir comme outil de comparaison entre technique de visualisation. VITESSE a également servi à mettre au point de nouvelles représentations (métaphores du cylindre ou de la bosse).

Tous ces résultats nécessitent une validation par l'expérimentation. Il est en effet difficile d'affirmer sans évaluation que ces techniques d'interaction sont bien adaptées à l'utilisateur car elles font intervenir la perception et la cognition. L'évaluation



permettra donc de valider nos propos mais également de quantifier l'amélioration par rapport aux systèmes existants.

Un autre travail futur sur VITESSE concerne plus spécifiquement la programmation qui s'est faite avec la première version du langage JAVA. L'apparition de nouvelles versions de ce langage remet en question un certain nombre de reproches que nous avons formulés au paragraphe 2.2. La traduction de VITESSE vers ces nouvelles versions devrait ainsi nous permettre d'améliorer l'architecture logicielle mais aussi des aspects de l'interface.



## **Chapitre V**

*Conclusion : bilan et perspectives*

Nous concluons le rapport de D.E.A. par un bilan de notre travail puis nous exposons les perspectives que nous envisageons tant sur le plan théorique que sur le plan de la réalisation logicielle.

## **1. Bilan et contribution**

Le stage de D.E.A. nous a permis d'explorer les possibilités engendrées par l'utilisation de techniques d'interaction multimodale pour mieux rendre perceptible un grand espace d'informations.

Notre étude s'est appuyée sur un état de l'art de la multimodalité. Les taxonomies des systèmes multimodaux présentées au Chapitre II mettent en évidence le peu d'études sur la multimodalité en sortie, du système vers l'utilisateur. La communauté a clairement beaucoup plus travaillé sur la multimodalité en entrée de l'utilisateur vers le système. Nous avons montré comment certaines taxonomies des modalités en entrée peuvent être étendues pour caractériser les modalités de sortie et nous avons exploité la seule taxonomie existante des modalités en sortie. L'objectif de cet état des lieux est d'identifier des caractéristiques des modalités de sortie. Ces derniers sont nécessaires afin de pouvoir sélectionner les modalités de sortie qui sont pertinentes pour la tâche courante et adaptées à l'utilisateur et au contexte d'interaction. Comme nous ne pouvons embrasser tout le domaine de la multimodalité de sortie dans une étude de D.E.A., nous nous sommes ensuite intéressés à exploiter ces caractéristiques dans le cas particulier d'une grande quantité d'informations.

Au Chapitre III, nous exploitons donc les caractéristiques des modalités de sortie du Chapitre II pour les appliquer au cas du rendu d'une grande quantité d'informations. Pour cela, nous présentons d'abord les problèmes liés au cas d'une grande quantité d'informations afin de fixer les objectifs : le problème du rendu simultané du « centre d'intérêt + contexte d'interaction ». Nous présentons ensuite des règles de sélection et de combinaison des modalités de sortie selon le contexte d'interaction et l'espace d'informations à rendre perceptible. Ces règles découlent de critères d'ergonomie que nous avons identifiés comme cruciaux pour le cas d'une grande quantité d'informations. Elles constituent un premier pas vers un espace de conception d'interfaces multimodales en sortie. Cet espace de conception nous semble faire cruellement défaut et être nécessaire pour progresser dans ce domaine d'étude.

Le chapitre IV présente notre système VITESSE, comme un application pratique des résultats présentés au Chapitre III. VITESSE est un système sur le WWW qui offre une interface multimodale en sortie pour rendre perceptible l'ensemble des réponses fournies par un moteur de recherche, à la suite d'une requête soumise. Le grand espace d'informations que nous avons choisi est donc constitué des réponses des moteurs de recherche. Nous avons appliqué nos règles pour la conception de l'interface. VITESSE a aussi été l'objet d'une étude sur l'architecture logicielle : nous avons appliqué le modèle d'architecture PAC-Amodeus de l'équipe pour le confronter aux problèmes de la multimodalité en sortie et de la portabilité d'un moteur de

---

recherche à l'autre. De plus le langage de programmation utilisé JAVA a soulevé des problèmes intéressants en terme d'architecture.

## **2. Perspectives**

Faisant suite à cette étude, nos propositions s'organisent en deux parties : les extensions et les prolongements à plus long terme.

### **2.1. Perspectives à court terme**

Afin de valider nos résultats, il est crucial de maintenant mener une évaluation ergonomique du système VITESSE. La mise en place de l'évaluation nécessite l'aide d'un spécialiste en ergonomie. En effet il convient de mettre en place le protocole de l'expérimentation et les scénarios ainsi que de choisir un ensemble de sujets représentatifs. Cette évaluation est planifiée pendant l'été 1997.

Cette évaluation implique des modifications logicielles de VITESSE. Le code de VITESSE doit être instrumenté afin de capter et d'enregistrer les actions de l'utilisateur pour une analyse ergonomique ultérieure. Un format de capture a été déjà élaboré dans l'équipe par un autre étudiant de D.E.A. Outre l'instrumentation, d'autres présentations multimodales peuvent être définies, que nous envisageons de rajouter avant l'évaluation. Enfin dans le cadre du contrat avec France Télécom il est prévu de connecter VITESSE à un ensemble d'informations plus limité que celui du WWW, et qui concerne uniquement des présentations d'hôtels. Ceci permettrait de mettre en place des scénarios d'utilisation beaucoup plus précis.

### **2.2. Perspectives à long terme**

Nous envisageons à plus long terme de nombreux travaux. Certains poursuivent mon étude de D.E.A., tandis que d'autres correspondent à des approches différentes mais complémentaires du domaine

#### **Méthode de conception**

Lors de notre étude de D.E.A. nous avons identifié des caractéristiques des modalités en sortie et des règles de sélection en vue de la conception de présentations multimodales. Ceci constitue un premier pas qu'il convient de prolonger. L'objectif serait alors d'identifier tous les paramètres nécessaires au choix des modalités de sortie qui soient en accord avec la tâche à réaliser et adaptées à l'utilisateur. L'approche serait de mettre en place une méthode de conception qui dans un premier temps serait limitée aux problèmes liés au rendu d'une grande quantité d'informations avant d'essayer de généraliser à tous types de tâche et de contexte.

#### **Moteur général de génération de présentations multimodales**

La mise en place d'une méthode de conception pourrait alors déboucher sur la définition d'un moteur de génération de présentations multimodales. En effet il est envisageable d'identifier des règles de sélection et de combinaison des modalités qui soient générales et programmables. Ce moteur pourrait alors constituer une partie d'un squelette d'application multimodale ou être intégré à un générateur d'interfaces multimodales. Ce travail devra s'accompagner d'une étude sur les modèles d'architecture logicielle pour les interfaces multimodales afin de définir les interfaces logicielles de notre moteur.

### **Collecticiel : navigation à plusieurs**

Une tâche-utilisateur liée au rendu de grandes quantités d'informations est la navigation dans l'espace. Nous avons lors de notre étude de D.E.A. décrit les facteurs intervenant dans la navigation. Une piste de recherche qui nous semble nouvelle est d'exploiter le fait que la navigation et le filtrage des informations peut se faire à plusieurs dans le cas d'un collecticiel. La navigation reposerait alors sur les compétences de plusieurs utilisateurs qui pourraient partager leurs expériences et leurs connaissances de l'espace. Outre des problèmes d'architecture logicielle, cet étude pourrait aboutir à de nouvelles techniques d'interaction adaptées à un groupe d'utilisateurs.

# Références bibliographiques

[ARCH]

The UIMS Workshop Tool Developers, A Metamodel for the Runtime Architecture of an Interactive System. In SIGCHI Bulletin, 24, 1 (Jan 92), pp. 32-37

[Bernsen 93]

Bernsen N., Taxonomy of HCI Systems: State of the Art, ESPRIT BR GRACE, deliverable 2.1, (1993)

[Buxton 83]

Buxton W., Lexical and pragmatic considerations of input structures, Computer Graphics 17 (1), (1983), pp. 31-37.

[Cadoz 94]

Cadoz C., Le geste canal de communication homme/machine. La communication instrumentale. Technique et science informatiques. Volume13-n°1/1994, pp. 31-61.

[Card 90]

Card S., Mackinlay J. and Robertson G., The Design Space of Input Devices, I-n proceedings of CHI'90 (April 1-5 1990, Seattle, Washington), pp. 117-124.

[Coutaz 95]

Coutaz J., Nigay L., Salber D., Blandford A., May J., Young R., Four easy pieces for assessing the usability of multimodal interaction: The CARE properties. In proceedings of Interact'95 (1995, Lillehammer, Norway), pp. 115-120.

[Dieberger 94]

Dieberger A. Navigation in Textual Virtual Environments using a City Metaphor. PhD thesis. <http://www.lcc.gatech.edu/faculty/dieberger/> (1994).

[Foley 84]

Foley J., Wallace V., Chan P., The human Factors of computer Graphics interaction technics, IEEE computer Graphics and Applications, 4(11), (1984), pp. 13-48.

[Frohlich 91]

Frohlich D., The Design Space of Interfaces, Multimedia Systems, Interaction and Applications, In Proceedings of 1<sup>st</sup> Eurographics Workshop (April 18-19 1991, Stockolm, Suède), pp. 53-69.

[Furnas 97]

Furnas G., Effective View Navigation. In proceedings of CHI'97 (March 22-27 1997, Atlanta, Georgia ), pp. 367-374.

[Gargan 88]

Gargan R., Sullivan J., and Tyler S., Multimodal Response Planning : an Adaptative Rule Based Approach., CHI'88.

[Gram 96]

Gram C. et Cockton G. Design Principles for Interactive Software. Chapman & Hall 1984, pp. 133-187.

[Leung 94]

Leung Y. and Aperley M., A Review and Taxonomiy of Distortion-Oriented Presentation Techniques. In ACM ToCHI (Vol 1. No2, June 1994), pp. 126-160

[Mackinlay 95]

Mackinlay J., Robertson G. et Card S. The Perspective Wall : Detail and Context Smoothly Integrated. In Proceedings of CHI'91 (April 27-May 2, 1991, New Orleans, Louisiana), pp. 173-179.

[Martin 94]

Martin J., Cadre d'étude de la multimodalité sur les types et buts de coopération entre modalités. In Proceedings Interface to Real&Virtual Worlds, EC2 Publ. pp. 97-106.

[Stone 94]

Stone M., Fishin K. and Bier E., The Movable Filter as a User Interface Tool. In CHI'94 Proceedings (April 24-28, 1994 Soston, Massachusetts). pp. 306-312.

[Nigay 93]

Nigay L. et Coutaz J., A Design Space For Multimodal Systems : Concurrent Processing and Data Fusion. In INTERCHI'93 Proceedings (April 24-29 1993, Amsterdam, Netherlands).

[Nigay 94]

Nigay L., Conception et modélisation logicielles des sytèmes interactifs. Thèse de l'université J. Fourier. (soutenu le 28 janvier 1994 à Grenoble).

[Nigay 97a]

Nigay L. et Vernier F., Navigable Worlds. CHI'97 Workshop on Navigation in Electronic Worlds (March 23-24 1997, Atlanta, Georgia),  
<http://iihm.imag.fr/nigay/WK-NAVIGATION/>

[Nigay 97b]

Nigay L. et Vernier F., Multiple representations of navigable worlds. Soumis et accepté à : Information Visualisation Workshop in Paris (24-25 June 1997, Paris, France).

[PROJECT X]

Apple Research Labs Technology., Project X,  
<http://www.atg.apple.com/research/tech/projectx/>



[Robertson 91]

Robertson G., Mackinlay J. et Card S. Cone Trees : Animated 3D Visualisations of Hierarchical Information. In Proceedings of CHI'91 (April 27-May 2, 1991, New Orleans, Louisiana), pp. 189-194.

[Sarkar 92]

Sarkar M. and Brown M., Graphical fisheye views of graphs. In CHI'92 Conference proceedings (May 3-7 1992, Monterey, California), pp. 83-91.

[Sellen 92]

Sellen A., Kurtenbach G., Buxton W., The Prevention of Mode errors Through Sensory Feedback, Human Computer Interaction, Lawrence Erlbaum, Vol. 7, No2, (1992), pp. 141-164.

[Spence 97]

Spence B., The acquisition of Insight. <http://www.ee.ic.ac.uk/hp/staff/spence.html>

[Stone 94]

Stone M., Fishkin K. and Bier E. The Movable filter as a user interface tool. In Proceedings of CHI'94 (April 24-28, 1994, Boston, Massachusetts), pp. 306-312.

[Vernier 96]

Vernier F., Etude et mise en œuvre de techniques d'interaction pour les moteurs de recherche du World Wide Web. Rapport de magistère 2<sup>ème</sup> année, 43 pages.

[Vernier 97]

Vernier F. et Nigay L., Représentation Multiple d'une Grande Quantité d'Information. A paraître dans IHM'97 (10-12 Septembre 1997, Futuroscope de Poitiers, France).

[VITESSE]

VITESSE, Visualisation & Interaction Techniques to Enhance Superscalar Search Engine. <http://iihm.imag.fr/vernier/Vitesse.html>.

[Ware 95]

Ware C. et Lewis M. The DragMad Image Magnifier. In Proceedings of CHI'95 Conference Companion (May 7-11 1995, Denver, Colorado), pp. 407-408.