

# Interfaces multimodales : composition et caractérisation des modalités de sortie

Frédéric Vernier et Laurence Nigay

Laboratoire CLIPS-IMAG, Equipe IHM  
BP 53 38041 Grenoble cedex 9, France

{[frederic.vernier](mailto:frederic.vernier), [laurence.nigay](mailto:laurence.nigay)}@imag.fr

## RESUME

Cet article a trait à la conception d'interfaces multimodales en sortie. Les deux points étudiés sont la composition de modalités de sortie et la caractérisation d'une modalité de sortie. En particulier, nous proposons un espace de caractérisation des compositions de modalités, identifiant vingt cinq classes de composition. La seconde partie de cet article présente MulTab, un système multimodal en sortie. Ce système permet d'illustrer les compositions de modalités au sein d'un système de visualisation d'un grand tableau de données.

**MOTS CLES :** Multimodalité en sortie, Espace de conception, Composition, Caractérisation.

## INTRODUCTION

Les interfaces multimodales font l'objet d'études depuis que les concepteurs ont conçu des systèmes mettant en œuvre plusieurs dispositifs en entrée (permettant à l'utilisateur d'interagir avec le système). Les études en ingénierie de l'interaction Homme-Machine s'intéressent également à l'interaction en sortie (du système vers l'utilisateur) mais le terme multimodalité y est rarement appliqué. Le terme multimédia est plus souvent retenu, et désigne davantage les supports de l'information échangée (médiats) plutôt que son contenu. Notre étude se consacre à l'interaction multimodale en sortie qui peut être analysée selon le dispositif mis en œuvre et le langage utilisé [12]. Si le concept de dispositif est très courant, celui de langage d'interaction nécessite d'être illustré. Le système VITESSE [14] illustre deux langages en entrée reposant sur le même dispositif : le clavier. Dans VITESSE, ce dispositif peut être utilisé pour saisir du texte (une requête) et pour naviguer dans un espace d'information (le résultat d'une requête). Le langage d'interaction de la saisie d'une requête est la langue pseudo-naturelle. Le langage d'interaction de l'outil de navigation est analogique et repose sur l'emplacement des touches sur le clavier. Ainsi, une pression de la touche "Esc" transporte l'utilisateur vers le coin gauche supérieur de l'espace car cette touche se situe en haut à gauche du clavier. De même, la touche "H" le transporte au centre de l'espace. Pour le concepteur du système interactif, cette modalité "navigation" ne peut pas être considérée comme une autre forme de la modalité "saisie

de texte". Elle met en œuvre une programmation complètement différente où seul le pilote du périphérique est commun. La programmation d'un système interactif nécessite de considérer, à un plus haut niveau d'abstraction, les informations échangées entre le système et l'utilisateur. Aussi, dans la suite de cet article, nous considérons qu'une modalité, en entrée comme en sortie, désigne un moyen de communication qui met en œuvre un dispositif physique et un langage d'interaction.

## CONCEPTION D'UNE INTERFACE MULTIMODALE EN SORTIE

En sortie, l'interaction multi-dispositifs est un domaine assez pauvre à cause du manque de variété dans les dispositifs de sortie. Les manettes à retour d'effort sont principalement utilisées pour les simulateurs et les jeux. Les haut-parleurs constituent la seule alternative classique aux écrans. Cependant les modalités sonores (exploitant des haut-parleurs) soulèvent des problèmes tels que :

- Le volume des haut-parleurs peut être réglé (éventuellement à 0) par l'utilisateur, sans que le système en soit informé.
- Le volume de l'environnement sonore de l'utilisateur peut être plus élevé que celui des haut-parleurs.
- Le contexte d'interaction de l'utilisateur (lieu de travail par exemple) peut lui imposer un volume très bas, voire nul.

Pour ces raisons, les modalités sonores sont rarement exploitées dans l'interaction, excepté dans certains types de logiciels comme la création musicale où elles sont inévitables. Dans les interfaces traditionnelles, les modalités sonores sont généralement utilisées pour répéter l'information graphique (redundance et insistance). Dans les interfaces pour personnes handicapées, les modalités sonores constituent une solution lorsque les modalités visuelles ne sont pas utilisables.

Afin de cerner l'espace des possibilités du concepteur d'interfaces multimodales en sortie, nous considérons de considérer l'interaction multi-langages comme l'interaction multi-dispositifs. Ceci permet de traiter les compositions de modalités graphiques (par exemple un

tableau de chiffres et un histogramme) qui peuvent offrir les mêmes apports de complémentarité et de redondance que l'interaction multi-dispositifs.

Adoptant ce point de vue, deux éléments de conception doivent être pris en compte :

- la composition de modalités,
- la caractérisation d'une modalité simple ou composée en vue de sa sélection.

Bien que ces deux éléments de conception soient fortement liés[10], nous les considérerons néanmoins séparément à des fins analytiques. Ils sont essentiels à la conception d'une interface multimodale. Pour chacun de ces deux éléments, nous proposons des cadres de conception.

### COMPOSITION DE MODALITES

Notre travail sur la composition est à mettre en relation avec les espaces existants sur la composition de modalités comme CARE [4], TYCOON [9] et MSM [12]. Dans les espaces TYCOON et CARE, les concepts de complémentarité et de redondance sont centraux et relèvent d'une composition sémantique. Dans l'espace MSM [12] les aspects articulatoires et sémantiques sont identifiés par un axe "niveau d'abstraction" pour différencier Multimédia et Multimodal. Enfin la distinction entre modalités séquentielles et parallèles est présente dans ces trois espaces.

Notre approche consiste tout d'abord à distinguer les cinq aspects suivants de la composition :

- la composition temporelle de modalités,
- la composition spatiale de modalités,
- la composition articulatoire des modalités (niveau dispositif),
- la composition syntaxique des modalités (niveau langage d'interaction),
- la composition sémantique des modalités.

Les compositions spatiales reflètent la volonté du concepteur de placer les modalités à des endroits spécifiques de l'environnement de l'utilisateur. Nous considérons également la composition temporelle qui caractérise les enchaînements possibles de modalités. La composition articulatoire des modalités dénote une composition à un niveau d'abstraction plus bas, entre les flux d'information transitant par les dispositifs associés aux modalités. La composition syntaxique considère les relations entre les langages d'interaction utilisés. Enfin, la composition sémantique concerne la composition des informations véhiculées par les modalités.

La seconde étape de notre approche consiste à considérer, les cinq aspects de la composition (temporel,

spatial, articulatoire, syntaxique et sémantique) selon les cinq schémas issus des relations de Allen [1] et illustrés par les Figures 1 à 5.

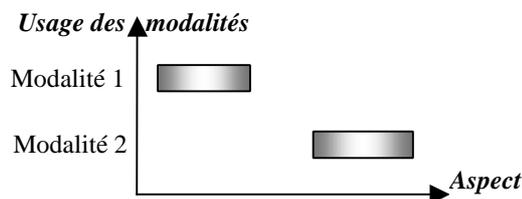


Figure 1 : Composition de modalités éloignées.

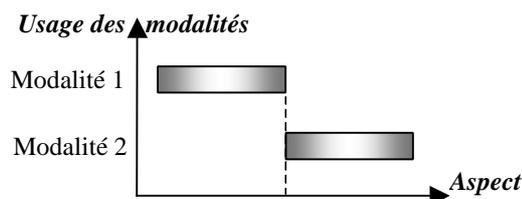


Figure 2 : Composition de modalités avec un point de contact.

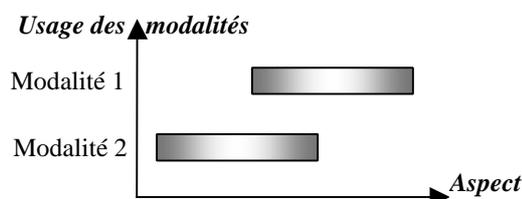


Figure 3 : Composition de modalités avec une intersection non vide.

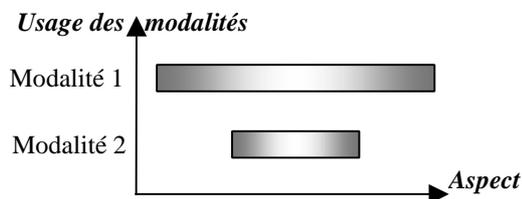


Figure 4 : Composition de modalités dont l'une est plus étendue et englobe l'autre.

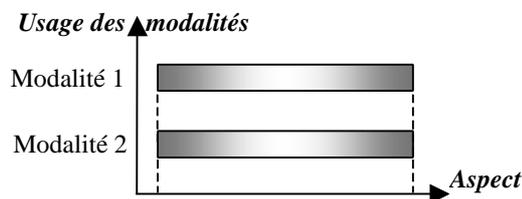


Figure 5 : Composition de modalités de même étendue.

L'application des cinq schémas de composition aux cinq aspects de composition définit 25 types de compositions, présentés dans le Tableau 1.

## Schéma de composition

Aspect de composition	Composition					
Temporelle	Anachronique	Séquentielle	Concomitante	Coïncidente	Parallèle / Simultanée	
Spatiale	Disjointe	Adjacente	Intersectée	Imbriquée	Recouvrante	
Articulatoire	Indépendante	Fissionnée	Fissionnée et dupliquée	Partiellement Dupliquée	Dupliquée	
Syntaxique	Différente	Complétion	Divergente	Extension	Jumelle	
Sémantique	Concurrente	Complémentaire	Complémentaire et redondante	Partiellement redondante	Totalement redondante	

**Tableau 1:** Application des schémas de composition  
aux cinq aspects de composition : temporel, spatial, articulatoire, syntaxique et sémantique.

### Composition temporelle de modalité

La séquentialité et le parallélisme dans l'usage des modalités ont été largement étudiés [4, 11]. Au contraire, les trois autres schémas de composition temporelle du Tableau 1 sont nouveaux pour la composition de modalité et constituent donc un affinement significatif.

La composition anachronique de modalités se distingue de la séquentialité par l'intervalle de temps entre les modalités. La durée minimale de l'intervalle pour passer d'une composition anachronique à une composition séquentielle est fixée par le concepteur en fonction de ses besoins. L'intervalle nécessaire pour séparer deux messages sonores sans rapport entre eux, et que le concepteur veut donc anachroniques, doit par exemple être de plusieurs secondes.

La composition par concomitance intervient quand une modalité en remplace une autre, avec un certain intervalle de temps pendant lequel les deux modalités sont présentes. Comme pour la composition anachronique, la durée de cet intervalle de transition dépend des besoins du concepteur. Une telle composition est importante pour aider et accompagner l'utilisateur dans la transition d'une modalité à une autre. Elle nécessite cependant que le ou les dispositifs puissent rendre en parallèle ces deux modalités pendant quelques instants.

Enfin, la composition coïncidente implique que l'une des deux modalités n'apparaisse que lorsque la seconde est déjà présente. Cette même modalité se termine avant que la seconde soit achevée. Une telle relation est nécessaire lorsque le concepteur met en œuvre une modalité qui fait sens seule et une seconde modalité qui ne fait sens qu'au regard de la première. Le système PPP [2] utilise cette composition : le message sonore n'est joué que lorsque

l'animation pointe sur un objet de l'interface comme une image.

### Composition spatiale de modalités

L'aspect spatial de la composition est un trait plus particulier aux modalités de sortie. En effet si les aspects temporel et sémantique de la composition des modalités d'entrée sont importants [4, 9, 12], l'aspect spatial est souvent négligé car les dispositifs d'entrée sont souvent différents pour chaque modalité. En sortie, le problème est différent car les modalités ne partagent essentiellement que deux dispositifs : l'écran et les haut-parleurs. Si nous considérons que deux modalités peuvent recouvrir le même lieu (comme la surface d'un écran ou l'espace de propagation d'un son), c'est que les sens perceptifs humains (vision et ouïe) nous permettent de percevoir plusieurs éléments au même endroit. La transparence par exemple nous permet de percevoir deux informations affichées au même endroit d'un écran. De même, lorsque le son n'est pas spatialisé, tous les sons perçus au travers des haut-parleurs nous semblent provenir du même endroit.

Une composition disjointe signifie que l'utilisateur peut percevoir une distance entre les modalités. A contrario, une composition adjacente signifie que l'utilisateur perçoit les modalités en des lieux différents mais comme côte à côte. Les compositions intersectée, imbriquée et recouvrante dénotent trois types de transparence. La composition intersectée définit par exemple le fondement des lentilles magiques ("Magic lenses") [13]. Une lentille magique est translucide et recouvre une partie de l'espace d'information sur laquelle s'applique l'opération de la lentille. Lorsque deux lentilles s'intersectent, leurs effets s'accumulent.

La composition imbriquée est illustrée par des exemples générés automatiquement par le processus décrit dans [10] : un rectangle (modalité imbriquée) dénote à tout instant la zone du schéma (modalité englobante) qui est en cours d'explication orale.

Enfin le système des "Pixels Miroir" [15] est un exemple de composition recouvrante où la totalité de l'écran est utilisée par deux modalités différentes : le document et le reflet de l'utilisateur.

Ces cinq types de composition sont importants pour le concepteur d'une interface graphique, généralement limitée par l'espace disponible à l'écran. Les compositions recouvrante et imbriquée utilisent moins d'espace, mais induisent généralement un effort de perception et d'interprétation de la part de l'utilisateur (problème de continuité visuelle).

### **Composition articulatoire de modalités**

La composition articulatoire concerne la composition des flux d'information qui transitent par les dispositifs associés aux modalités. Directement liés à la composition articulatoire, les moteurs de fusion pour les interfaces multimodales en entrée soulèvent des problèmes d'ingénierie. Les auteurs de [11] proposent le terme de fission qui paraît plus adapté à l'interaction en sortie.

Les cinq cas de composition articulatoire affinent ce terme de fission.

Les compositions, nommées indépendante et dupliquée, aux deux extrêmes du Tableau 1 traduisent respectivement l'absence de lien entre les couches bas niveau des modalités ou au contraire une duplication de toute l'information entre les deux modalités. Par exemple la composition des deux modalités exploitant deux haut-parleurs d'un système monophonique est dupliquée.

La composition fissionnée repose sur une fission d'un flux d'information. Les systèmes tels que les murs d'images (un même film découpé sur plusieurs écrans) mettent en œuvre une fission à bas niveau d'abstraction d'un flux vidéo découpé en sous-parties contiguës. Chaque partie est jouée sur un dispositif différent donc une modalité distincte.

Les compositions dupliqués et fissionnées sont rarement mises en œuvre et nécessitent un contexte spécial. Nous considérons par exemple comme dupliquées et fissionnées deux vidéo issues d'une même source mais dont chacune est sous-titrée dans une langue différente.

La composition partiellement dupliquée est par contre plus fréquente. On constate en effet que certains dispositifs ont la capacité de présenter des informations identiques mais en quantité ou qualité différentes.

L'écran d'une station de travail et d'un ordinateur de poche peuvent par exemple présenter la même image mais celle de l'ordinateur de poche sera tronquée par rapport à celle de la station de travail. Un document Postscript sorti sur imprimante peut également avoir une meilleure résolution que son double affiché à l'écran.

### **Composition syntaxique de modalités**

Tandis que la composition articulatoire se réfère au dispositif de la modalité, la composition syntaxique a trait à son langage d'interaction. Nous distinguons cinq schémas de composition syntaxique qui vont de la composition de modalités dont le langage d'interaction est le même (composition jumelle) jusqu'à la composition modalités aux langages d'interaction différents, comme l'exemple de VITESSE décrit en introduction. La composition de complétion dénote l'utilisation de modalités dont les langages ont des pouvoirs d'expression complémentaires. Le langage pseudo naturel complète par exemple celui des dates afin de produire des énoncés tel que "*la réunion aura lieu le 10/03/2000*". La composition divergente indique l'utilisation de langages d'interaction ayant une partie du vocabulaire en commun. C'est le cas par exemple du langage écrit et du langage parlé. Enfin nous proposons la composition d'extension qui indique des langages aux pouvoirs d'expression imbriqués. Par exemple, le texte mis en forme avec des attributs tels que l'italique ou le soulignement est un langage plus étendu que le texte brut, sans formatage.

### **Composition sémantique de modalités**

Un aspect central à considérer dans une composition de modalités est la relation entre la sémantique des informations véhiculées par les deux modalités. Les formes de multimodalité qui ont été le plus étudiées sont sans nul doute la complémentarité et la redondance [4, 10]. La concurrence entre modalités, identifiée dans l'espace de conception de [11], n'est généralement pas considérée comme un cas de multimodalité car une analyse séparée des deux modalités est préférable. La composition que nous appelons partiellement redondante et celle que nous appelons complémentaire et redondante nécessitent un complément d'explications.

La composition complémentaire et redondante véhicule des informations dont une partie seulement se répète. Une telle composition de modalités graphiques est souvent nécessaire dans les vues multiples. Chaque vue présente des informations différentes. Une partie des informations est cependant répétée dans chacune des vues afin de permettre à l'utilisateur d'établir un lien entre les vues. Par exemple le système "MagniFind"<sup>1</sup> [8] permet à l'utilisateur d'interagir sur une vue listée et détaillée de ses fichiers (taille, date, etc.) et en même

---

<sup>1</sup> Disponible sur Internet : <http://www.inxight.com/>.

temps sur une vue en arbre hyperbolique qui ne permet pas de visualiser les fichiers ni les détails sur les répertoires, mais qui permet de voir de manière synthétique la liste des sous-répertoires. Ainsi, ces deux modalités graphiques, à première vue complémentaires, présentent pourtant une partie redondante puisque la liste des répertoires de plus haut niveau apparaît dans les deux vues.

Nous appelons une composition des modalités M1 et M2 partiellement redondante lorsque M1 véhicule une partie du sens véhiculé par M2 sans en reprendre la totalité. C'est un aspect très courant dans le domaine de la visualisation d'une grande quantité d'information car les vues synthétiques de l'espace, proposées dans de nombreux systèmes, n'offrent en général qu'une redondance partielle avec la vue principale de l'espace.

### Aspects et schémas : un espace unificateur

En conclusion une composition de modalités de sortie fait intervenir l'une des cinq compositions pour les aspects temporels, spatiaux, syntaxiques et sémantiques. Ces quatre axes indépendants impliquent donc 5<sup>4</sup> relations différentes entre modalités. Nous notons la particularité de la troisième ligne du Tableau 1, composition articulatoire, qui n'est pas indépendante des deux dernières lignes (compositions syntaxique et sémantique). Certaines combinaisons ne sont pas possibles. Ainsi, si une composition repose sur un flux d'information dupliqué vers deux dispositifs de sortie, les deux modalités ne sont pas sémantiquement complémentaires. La composition fissionnée et dupliquée et la composition partiellement dupliquée sont également restrictives. A l'opposé, les deux compositions indépendantes et fissionnées autorisent toutes les compositions aux niveaux syntaxique et sémantique.

L'espace de conception que nous proposons complète et unifie les espaces TYCOON, CARE et MSM. Notre espace est plus complet puisqu'il propose 25 types de compositions organisées selon cinq aspects fondamentaux que nous avons conjugués avec cinq schémas de composition. Ces compositions de modalités permettent de créer de nouvelles modalités dites composées. Outre la composition de modalités qui caractérise l'usage de plusieurs modalités, il convient de caractériser les modalités qu'elles soient simples ou composées.

### CARACTERISATION D'UNE MODALITE

Caractériser les modalités est nécessaire pour leurs sélections ou choix. La sélection d'une modalité est un problème principalement lié à l'interaction en sortie, car en entrée le concepteur peut faire migrer ce choix vers l'utilisateur qui est sans doute l'acteur le plus apte à choisir la modalité adaptée à sa tâche et au contexte. En sortie, l'utilisateur peut être soit concerné par le choix de

modalité soit peut désirer ne pas être sollicité. L'espace CARE propose deux concepts qui dénotent le choix ou l'absence de choix des modalités en entrée : l'équivalence et l'assignation. De plus les propriétés CARE peuvent être étudiées selon le point de vue de l'utilisateur (UCARE [4]). En sortie, nous distinguons trois acteurs : le concepteur, l'utilisateur et le système. Chaque acteur peut considérer que deux modalités sont équivalentes ou qu'une modalité est assignée à une tâche. Les choix de l'utilisateur sont limités par ceux offerts par le système et les choix du système par ceux effectués par le concepteur. Cette distinction implique trois types d'interfaces multimodales : aucune adaptation (choix des modalités effectué par le concepteur), adaptativité (choix calculé par le système) et adaptabilité (choix effectué par l'utilisateur).

Pour effectuer un choix, les modalités doivent être caractérisées et donc classées. O. Bernsen [3] proposent quatre propriétés booléennes appelées "profil d'une modalité" :

- statique ou dynamique,
- linguistique ou non-linguistique,
- analogique ou non-analogique,
- arbitraire ou non-arbitraire.

Ces propriétés constituent une aide pour le choix d'une modalité. Cependant elles n'impliquent pas un choix évident d'une modalité particulière puisqu'elles ne véhiculent pas une échelle de valeur (une modalité analogique n'est pas meilleure qu'une modalité non-analogique ou langagière).

En complément de la taxonomie de Bernsen, nous considérons les critères d'ergonomie identifiés<sup>2</sup> dans [7]. Ces critères traduisent l'utilisabilité des systèmes. Les choix effectués par le concepteur ou le système doivent pouvoir être reliés à un ou plusieurs de ces critères. Par exemple deux modalités vérifient le critère d'observabilité (capacité du système à rendre observable des variables internes) mais l'une d'elles est préférable pour le critère d'honnêteté (capacité du système à inciter la bonne interprétation du rendu). De plus une modalité arbitraire ne favorise pas le critère d'honnêteté. C'est par exemple le cas du son perçu à la réception d'un message électronique qui n'est pas "honnête" pour un utilisateur novice, alors que le message textuel "vous avez reçu du courrier" (modalité langagière) répond mieux au critère d'honnêteté. Une modalité composée de deux modalités redondantes favorise le critère d'insistance (capacité du système à mettre en valeur les informations les plus pertinentes pour forcer leur perception).

Nous avons proposé des traits de caractérisation d'une modalité, qu'elle soit simple ou composée. Ces traits sont cruciaux pour permettre un choix, qu'il soit effectué par

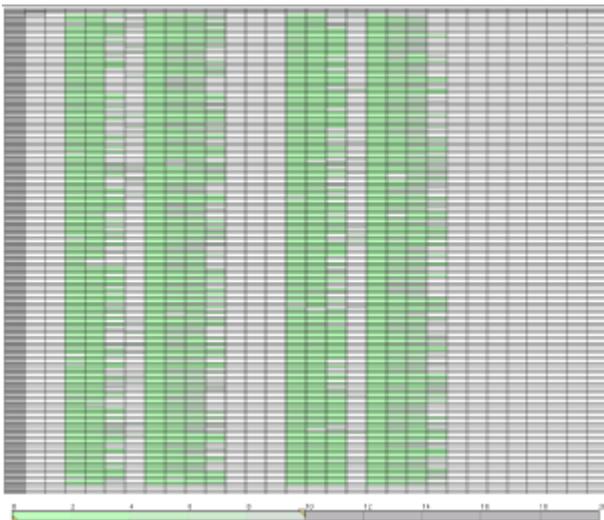
<sup>2</sup> <http://iihm.imag.fr/docs/coutaz.book/coutaz.properties.html>

le concepteur, par le système dynamiquement, ou par l'utilisateur. Les compositions et traits de modalités ont été illustrés séparément par des systèmes existants. Nous illustrons maintenant plusieurs compositions et traits par un seul système illustratif : MulTab. MulTab permet la manipulation de tableaux de données. Par le biais de ce système, nous illustrons les compositions spatiales, temporelles, syntaxiques et sémantiques. Le bas niveau d'abstraction (composition articulatoire) n'est pas traité par MulTab car un seul dispositif, l'écran, est utilisé.

## MULTAB

MulTab (MULTimodal TABLE) permet la manipulation de larges tables de données numériques ou alphanumériques, avec un maximum de 20000 cellules. Le système a été conçu en considérant les critères d'ergonomie d'observabilité, d'honnêteté, d'atteignabilité, d'insistance et de vue multiple.

MulTab offre une modalité de visualisation globale de la table, comme le montre la Figure 6 (observabilité). Comme les cellules apparaissent trop petites pour y inscrire une valeur numérique, nous avons utilisé un trait visuel, la couleur, pour traduire la valeur de la cellule.



**Figure 6 :** Colorisation des cellules dont la valeur est comprise entre 0 et 10.

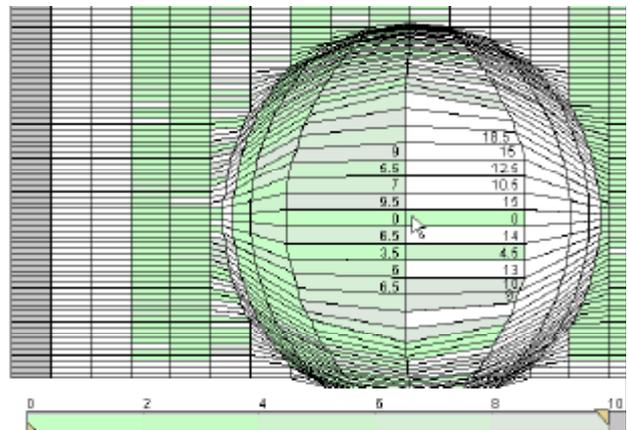
Pour favoriser le critère d'insistance, la teinte de la couleur est proportionnelle à la valeur. C'est une modalité non-linguistique au sens de Bernsen. Le critère d'observabilité a guidé notre conception de l'interacteur (visible au bas de la Figure 7) qui permet de sélectionner l'intervalle de valeurs à colorier selon un dégradé. Deux barres de défilement auraient été équivalentes pour le critère d'atteignabilité mais pas pour le critère d'observabilité. La modalité que nous avons retenue permet en effet de savoir visuellement, pour une couleur du dégradé, quelles sont les valeurs associées. Le critère d'honnêteté est aussi vérifié, car les curseurs, en forme de triangles inversés, incitent l'utilisateur à comprendre que

la borne inférieure ne peut pas dépasser la borne supérieure.

Si nous considérons maintenant la composition des deux modalités, celle liée au tableau colorié et celle de l'interacteur lié à la colorisation, nous la décrivons comme :

- parallèle dans le temps,
- adjacente dans l'espace,
- d'extension de syntaxe (le tableau étant plus étendu),
- complémentaire pour la sémantique.

Si l'utilisateur ne disposait que de ces deux modalités pour visualiser et éditer son tableau de données, il commettrait un grand nombre d'erreurs de navigation dues aux ressemblances des couleurs des cellules contenant des valeurs proches. Les couleurs ne sont présentées que pour donner des points de repère indispensables à la navigation. MulTab propose donc une troisième modalité illustrée, dans sa forme de demi sphère à la Figure 7. Cette Modalité, que nous appelons "focus", est dédiée à l'affichage en chiffre de la valeur contenue dans quelques cellules. Elle est linguistique au sens de Bernsen et sa forme est analogique. Le critère de multiplicité du rendu et le critère d'honnêteté nous ont conduit à produire la solution de la Figure 7, où la présentation de toutes les cellules est continue : en effet il n'y a pas de discontinuité visuelle entre la présentation des cellules selon la modalité "focus" et celle selon la modalité globale du tableau.



**Figure 7 :** Affichage d'une modalité détaillée (focus) en forme de demi sphère.

Ainsi, la composition de la modalité globale dédiée au tableau et de la modalité "focus" est :

- parallèle dans le temps,
- adjacente dans l'espace,
- d'extension pour la syntaxe,
- partiellement redondante pour la sémantique,

Comme le montre la Figure 8, l'utilisateur peut également faire apparaître deux petites demi sphères lorsqu'il a besoin de consulter le nom de la colonne ou de la ligne de la cellule au centre du focus. La Figure 8 illustre

également la possibilité de créer plusieurs centres d'intérêt (foci) afin de comparer des cellules.

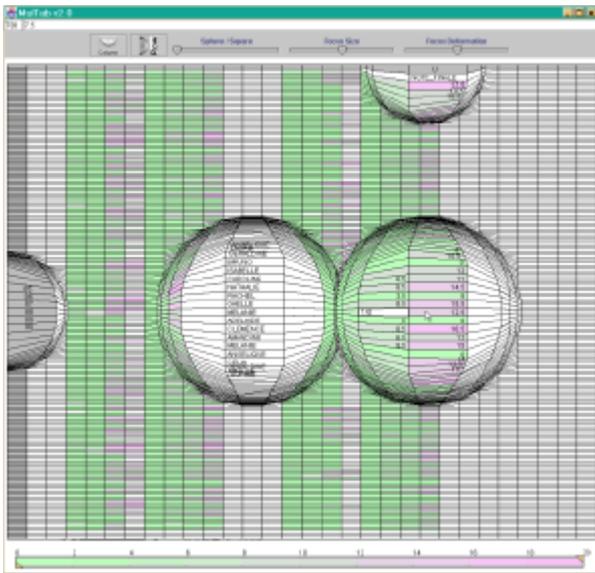


Figure 8 : Vue générale de l'application MulTab au moment où 4 foci sont utilisés par l'utilisateur.

Tous les foci secondaires ainsi créés sont liés au premier par une contrainte de type "même colonne centrale ou même ligne centrale". Cette technique basée sur les travaux des vues en œil de poisson [6] met donc en œuvre une composition de deux modalités, de même profil et même syntaxe, qui est :

- coïncidente dans le temps,
- complémentaire pour la sémantique.

L'aspect spatial dépend de l'utilisateur qui peut :

- garder les demi sphères rapprochées (composition intersectée),
- éloigner les deux modalités (composition disjointe).

La composition intersectée a impliqué d'implémenter un algorithme dédié dont le résultat est illustré à la Figure 9.

De plus un point important de conception de la modalité "focus" est sa forme graphique. L'aspect sphérique est un héritage des travaux de [6] sur les vues en œil de poisson.

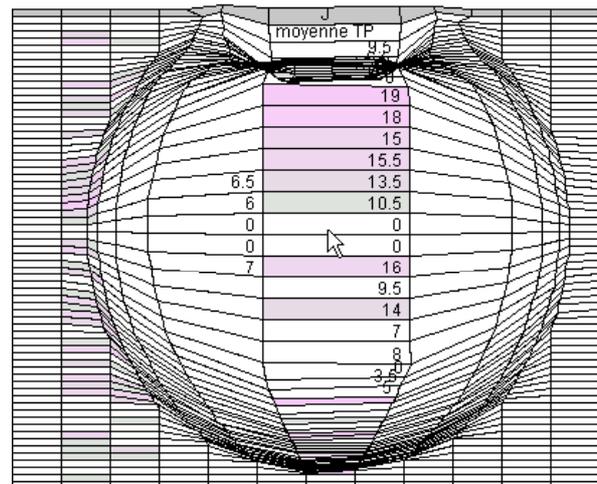


Figure 9 : Composition spatiale de deux foci rapprochés : le focus de l'espace et le focus sur le haut de la colonne.

Dans [5], une évaluation expérimentale a été menée pour comparer une forme circulaire et une forme rectangulaire pour une modalité détaillée incluse dans une modalité globale. Leur prototype Fovea, est reproduit avec permission à la Figure 10. Cette expérimentation a fait ressortir de meilleures performances de navigation avec la forme rectangulaire mais une préférence des utilisateurs pour la forme circulaire.



Figure 10 : Système Fovea avec un focus circulaire à gauche et un focus rectangulaire à droite.

De tels résultats expérimentaux ne permettent ni au concepteur ni au système d'effectuer la sélection de la forme de la modalité "focus". Dans MulTab, la sélection de la forme est donc laissée à l'utilisateur qui, à l'aide d'une glissière, change la forme sphérique en forme pyramidale en passant par des formes intermédiaires, comme l'illustre la Figure 11.

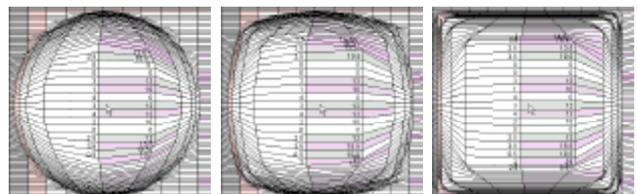


Figure 11 : Différentes formes de modalité "focus" : sphérique à gauche, pyramidale à droite, hybride entre les deux.

## CONCLUSION

Dans cet article, nous avons considéré la conception d'interfaces multimodales en sortie selon deux angles d'étude : la composition de plusieurs modalités et la caractérisation d'une modalité. La composition de modalités permet de construire de nouvelles modalités et donc d'étendre les possibilités de conception. La caractérisation d'une modalité simple ou composée est elle importante pour permettre le choix d'une modalité, que celui-ci soit effectué par le concepteur, le système ou l'utilisateur.

Nous avons d'abord défini un espace de conception formé de deux axes (aspect et schéma) pour identifier vingt cinq compositions différentes. Nous avons ensuite présenté notre espace de caractérisation qui s'appuie sur les profils d'une modalité [3] et les critères d'ergonomie [7]. L'une des contributions importantes de nos deux espaces est d'unifier et de compléter les approches existantes, que ce soit pour les interfaces en entrée ou en sortie. Nous avons enfin illustré ces éléments de conception par un exemple : MulTab.

Une première évaluation expérimentale des modalités "focus" a été conduite avec le système VITESSE [14] et une évaluation expérimentale de MulTab est en cours. De plus nous entrevoyons pour étendre les capacités interactionnelles de MulTab de multiples perspectives basées sur l'interaction multi-dispositifs. En particulier, nous envisageons l'utilisation de plusieurs écrans pour étendre la quantité d'information observable et l'utilisation du son pour renforcer la perception de certaines informations telles que le passage du centre d'intérêt d'une ligne à une autre. La mise en œuvre de ces modalités nous permettra de confronter les éléments de conception identifiés au niveau de la composition articulaire.

Pour la complétude de l'étude, il convient aussi d'étudier les relations entre une modalité d'entrée et de sortie, puisque c'est de leur couplage qu'émerge l'interaction. Peut-on appliquer notre espace de conception pour la composition d'une modalité d'entrée avec une modalité de sortie ?

## BIBLIOGRAPHIE

1. Allen J. *Maintaining Knowledge about Temporal Intervals*. Communication of the ACM. Volume 26. N°11. Novembre 1983. pp. 832-843.
2. André E., Rist T., et Müller J. *Guiding the User through Dynamically Generated Hypermedia Presentations with a Life-Like Presentation Agent*. Actes de la conférence IUI'98 (San Francisco, CA), pp. 21-28.
3. Bernsen N. *A revised generation of the taxonomy of output modalities*. Projet Esprit AMODEUS Working Paper RP5-TM-WP11, 1994.
4. Coutaz J., Nigay L., Salber D., Blandford A., May J. et Young R. *Four easy pieces for assessing the usability of multimodal interaction: The CARE properties*. Actes de la conférence Interact'95 (Lillehammer, Norway), 1995, pp. 115-120.
5. Coutaz et al. *CoMedi: Using Computer Vision to Support Awareness and Privacy in Mediaspaces*. Dans le volume "extended Abstract" des actes de la conférence CHI'99 (Pittsburgh, USA), 1999, pp. 13-16.
6. Furnas, G. *Generalized fisheye views*, Actes de la conférence CHI'86 (Boston, USA), 1986, pp. 16-23.
7. Gram C. et Cockton G. *Design Principles for Interactive Software*. Chapitre 2 du livre édité par Chapman & Hall, 1984, pp. 25-51.
8. Lamping, J., Rao R., et Pirolli P. *A focus+context technique based on hyperbolic geometry for visualizing large hierarchies*. Actes de la conférence CHI '95 (Denver, USA), 1995, pp. 401-408.
9. Martin J.C., *Six primitive types of cooperation for observing, evaluating and specifying cooperations*. Actes de la conférence AAAI Fall (Cape Cod, USA), 1999.
10. Maybury M. *Intelligent Multimedia Interfaces*. Menlo Park, CA: AAAI Press, 1993.
11. Nigay L. et Coutaz J. *A Design Space For Multimodal Systems: Concurrent Processing And Data Fusion*. Actes de la conférence INTERCHI'93 (Amsterdam, Hollande), 1993, pp. 172-178.
12. Nigay L. et Coutaz J. *Espaces conceptuels pour l'interaction multimédia et multimodale*. TSI, spécial Multimédia et Collecticiel, AFCET & Hermes Publ., Volume 15 N°9, 1996, pp. 1195-1225.
13. Stone M., Fishkin K. et Bier E. *The Movable filter as a user interface tool*. Actes de la conférence CHI'94 (Boston, USA), 1994, pp. 306-312.
14. Vernier F. et Nigay L. *Représentations Multiples d'une Grande Quantité d'Information*, Actes de la conférence IHM'97 (Poitiers, France), 1997, pp. 183-190.
15. Vernier F., Lachenal C., Nigay L. et Coutaz J. *Interface Augmentée par effet Miroir*. Actes de la conférence IHM'99 (Montpellier, France), 1999, pp. 158-165.