

# Interaction multimodale

## Usage, utilisabilité

**Jean Caelen**



*CLIPS*

**Communication Langagière et  
Interaction Personne-Système**

Fédération IMAG

BP 53 - 38041 Grenoble Cedex 9 - France

# Les problèmes fondamentaux

- Adéquation des modalités aux tâches
- Adéquation aux usages et aux profils  
d'utilisateurs
- Fusion des entrées
- Fission des sorties

# Adéquations...

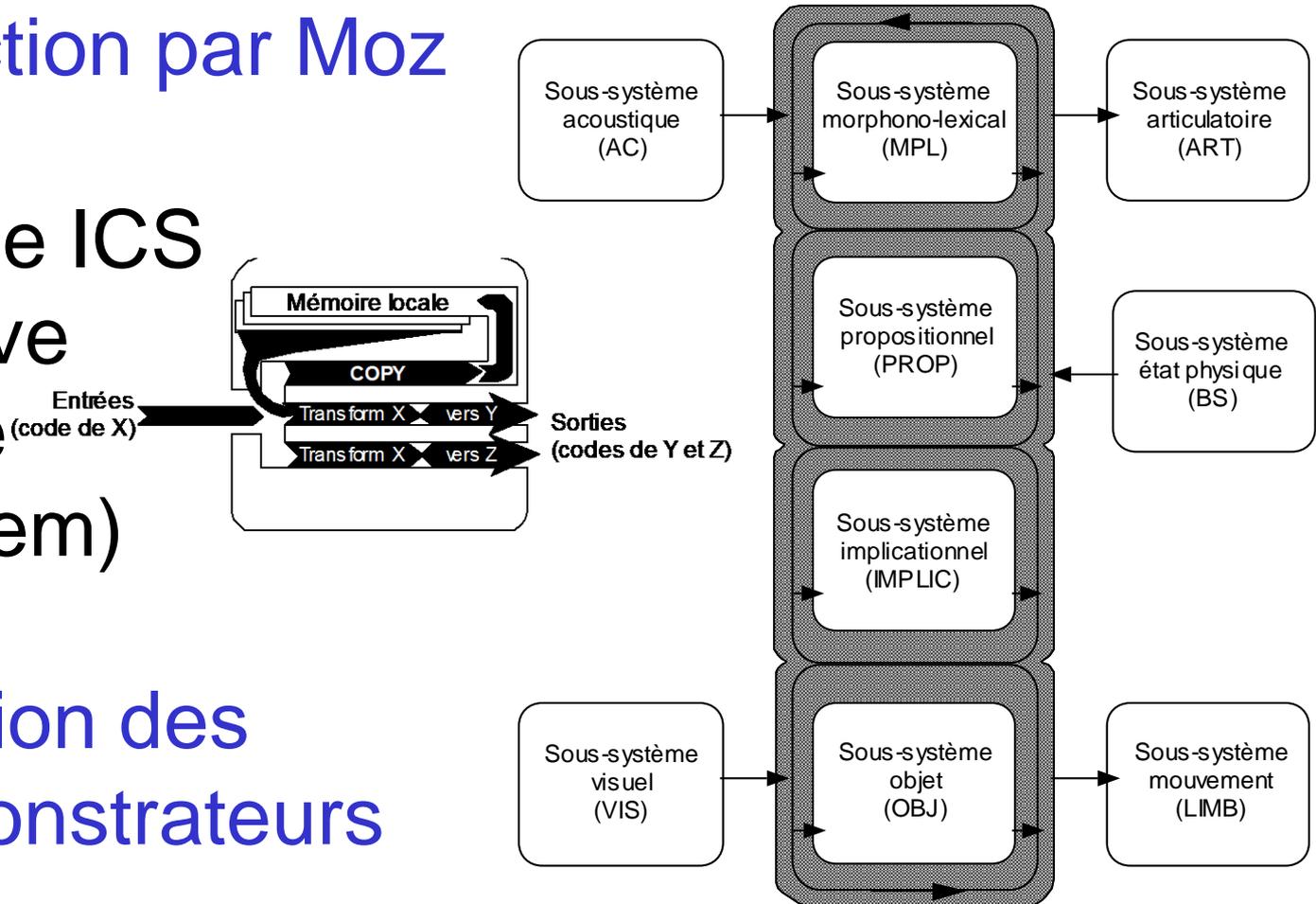
- Quelle est l'adéquation d'une forme d'interaction à une tâche donnée ? à un utilisateur donné ?
- Comment tenir compte des contraintes environnementales pour optimiser l'efficacité et la fiabilité des commandes de l'utilisateur mais aussi celles des messages restitués par la machine ?
- Quels liens faut-il établir entre les modalités de sortie et les modalités d'entrée ?
- Comment circonscrire un espace d'interaction pertinent au regard de la tâche ?

# Les méthodes

La prédiction par Moz

Le modèle ICS  
(Interactive  
Cognitive  
Sub-system)

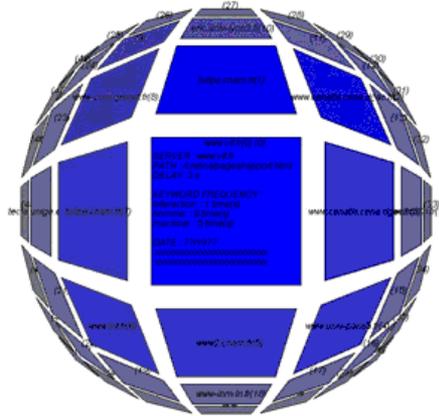
L'évaluation des  
démonstrateurs



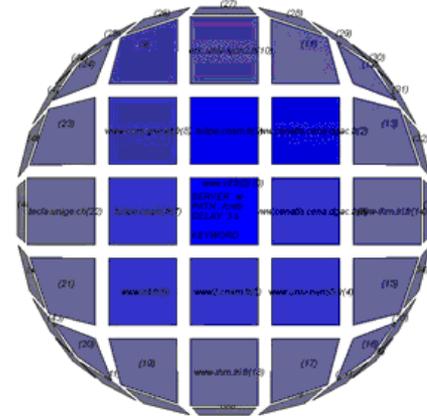
# Évolution de la notion de modalité

- **Modalités sensorielles**
- **Modalités de présentation**
- **Modalités d'action**

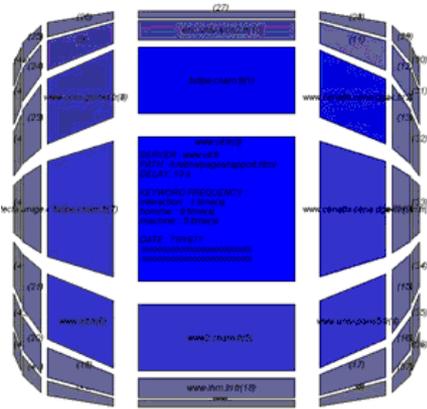
# Modalités de présentation



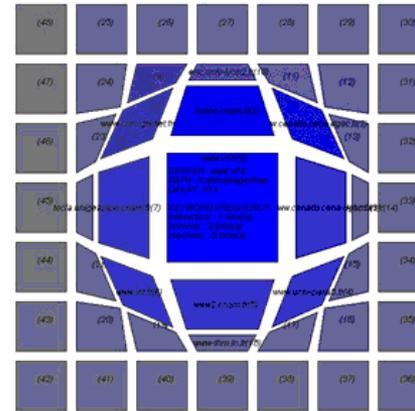
Sphère



Sphère tronquée



Cylindre



Colline

# Évolution de la notion de modalité

- **Modalités sensorielles**
- **Modalités de présentation**
- **Modalités d'action**

# 10 mythes (S. Oviatt)



***Ce n'est pas parce qu'une interface est multimodale que les utilisateurs vont utiliser la multimodalité.***

Dans QuickSet la multimodalité est utilisée dans 20% du temps d'une session de travail. Les utilisateurs passent d'un mode à l'autre sans raison apparente et restent unimodaux certainement pour des raisons de confort personnel. Cependant les commandes spatiales sont plus fréquemment multimodales ainsi que les informations de taille, de forme des objets, de nombres, de lieux et d'orientations. La richesse sémantique de l'action favorise la multimodalité.

# 10 mythes (S. Oviatt)

2

## ***Le pattern parole-pointage n'est pas le plus intéressant.***

Depuis le fameux « mets ça là » de Bolt, la multimodalité a été centrée sur le paradigme de l'interaction synergique. Dans ce paradigme la parole est considérée comme mode sémantique dominant et le geste de désignation comme subordonné. En fait cette conception est une survivance du concept clavier/souris (c'est-à-dire de sélection sur une icône ou un menu), bien plus pauvre qu'une interaction qui utiliserait les mouvements gestuels, les expressions faciales ou corporelles, etc. Par exemple des études avec un stylo/voix montrent que la multimodalité est de 14% plus utilisée qu'avec une entrée souris/voix. L'utilisation des déictiques est aussi plus fréquent de 20%.

# 10 mythes (S. Oviatt)



*La multimodalité ne signifie pas obligatoirement « parallélisme ».*

En effet on a constaté que bien souvent le geste précède la parole (99% des cas), même lorsque les deux modes dénotent des informations synchrones comme les déictiques. Le degré d'anticipation dépend de la langue. Il n'y a finalement que 25% des énoncés qui sont véritablement simultanés : synchronie ne signifie pas simultanété.

# 10 mythes (S. Oviatt)



*La parole n'est pas un mode « de base » dans un système multimodal.*

Cela n'est vrai que sur le plan historique. Depuis il y a bien des systèmes qui utilisent la main et le regard par exemple comme modes d'entrée, notamment dans les systèmes militaires. Le problème général de la multimodalité ne se pose donc pas en termes de *commande+sélection*, la commande étant linguistique et la sélection manuelle. Le problème ne se pose pas non plus en terme de *source principale/source secondaire* dans lequel on utiliserait la source secondaire dans le cas où la source principale serait dégradée.

# 10 mythes (S. Oviatt)

5

*Le langage multimodal ne diffère pas du langage unimodal.*

On peut dire seulement que le langage utilisé en contexte multimodal est syntaxiquement moins complexe, que les énoncés sont plus courts et que le débit est moins hésitant. Les ellipses sont plus fréquentes et les constructions linguistiques sont moins ambiguës, car les énoncés sont plus compacts. Il semble que ces propriétés rendent le langage multimodal plus apte à une intégration dans un système homme-machine.

# 10 mythes (S. Oviatt)

## 6 L'interaction multimodale ne favorise pas la redondance.

On pourrait croire le contraire, mais cela ne va pas dans un sens d'économie du point de vue de l'utilisateur. Celui-ci va donc privilégier la complémentarité. Même dans le cas d'échec puis d'essais de correction, l'usage de la redondance n'augmente pas de façon significative. La redondance n'est pratiquement utilisée que dans le sens d'une recherche de fiabilité

# 10 mythes (S. Oviatt)



**Les erreurs sur un mode ne sont pas compensées par un autre mode.**

Il est illusoire de penser que l'on va masquer les insuffisances d'un mode (par exemple les erreurs de reconnaissance de la parole) par un autre mode. En réalité les erreurs se cumulent d'un mode à l'autre. Mais les utilisateurs optimisent l'usage d'un mode au profit de tel autre, après expérience faite de ses performances, ce qui rend somme toute, par effet indirect de l'usage, l'interaction plus robuste. Dans quelques cas cependant, lorsque une double incertitude se produit dans les deux modes d'entrée, il est parfois possible de recouper l'information sur un critère de cohérence sémantique.

# 10 mythes (S. Oviatt)



**8 Les utilisateurs n'organisent pas « leur » multimodalité de la même manière.**

Pour les uns, ce qui est séquentiel, est parallèle chez les autres. Tel mode est dominant chez les uns, et ne l'est pas chez les autres. Tel mode est persistant, etc.

# 10 mythes (S. Oviatt)

## 9 Les modes ne sont pas équivalents.

Leur pouvoir d'expression est différent sans parler de leur pouvoir perceptuel, qui paraît plus évident. Cela signifie que le geste (et inversement la parole) ne peut tout exprimer dans une interaction, il y a des limitations cognitives. Même si parfois on eut rapprocher deux modes, ils n'en diffèrent pas moins par leurs propriétés différentes : précision, latence, etc. Certains modes sont plus inconscients ou passifs que d'autres : la direction du regard par exemple.

# 10 mythes (S. Oviatt)

10

■ **Un système multimodal n'est pas plus efficace qu'un autre.**

On croit souvent qu'un système multimodal sera plus efficace qu'un système monomodal, car on pourra faire plusieurs choses en même temps, se reposer en passant d'un mode à l'autre, réduire la charge perceptive et cognitive, économiser le temps de planification, etc. Des expériences ont prouvé le contraire : une commande multimodale est souvent plus longue à exprimer qu'une commande monomodale, car il y a un coût dû à la multimodalité (par exemple la multimodalité produit un débit de parole plus saccadé et des hésitations plus fréquentes).

**Le projet SUPRATEL**  
**Un logiciel de téléphonie**  
**multiservice, multimodal**

Callee :

Surname  
Firstname  
Tel

**Carraux**

Eric

76 44 55 97

Secret



Mirror

Hang up

Call



Videophone

Add to directory

Add Tel

Direct

Domicil... ▼



Carraux Eric  
76 44 55 97



Bis

Carraux Eric  
76 44 55 97

Direct

Agence ,... ▼



Bis

Anton, St... ▼



A  
B  
C  
D  
E  
F  
G  
H  
I  
J  
K  
L  
M

**Carraux  
Eric**



76 44 55 97



25 66 33 23 ▼

**Coutaz  
Joelle**



76 25 65 32 ▼



12345679

W  
O  
P  
Q  
R  
S  
T  
U  
V  
W  
X  
Y  
Z

Edit

Add

Supress

Add Tel

Add Fax

# Exemples de commandes

**commande orale :**  "ouvrir" (si un document est déjà sélectionné)

 "ouvrir le document X"

**commande gestuelle :**  clic souris sur le bouton concerné

**commande mixte:** **8** clic sur un document +  "ouvrir"

 "ouvrir" + **8** clic sur un document +  "ce document"

 "ouvrir" + **8** clic sur un document

# Exemple de scénarios

- Simulation d'une interruption

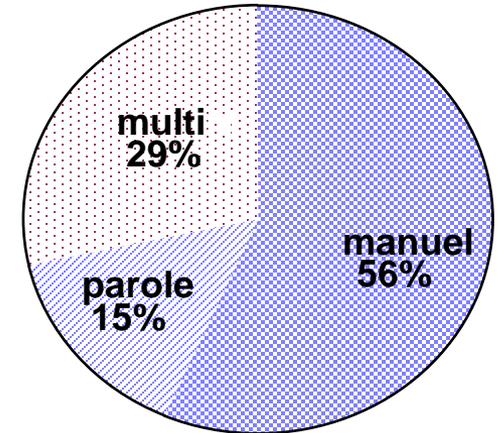
Tâche initialisée par le sujet	SCÉNARIO 4
Le numéro de Mireille PARRAN a changé, effectuer les modifications dans le répertoire : 74-23-45-28 Si un fax arrive, consultez le	Pendant la saisie, envoi d'un fax de Marie Lenoir 74-23-45-89

Tâche initialisée par le sujet	SCÉNARIO 3
téléphoner à Mme Camille MORIN pour demander l'heure et le jour de la prochaine réunion	inciter visiophonie (ne possédez vous pas la visiophonie inciter miroir (êtes vous sur que vous êtes bien cadré, je vous vois mal ) date par exemple : 20 décembre 10h30

# Résultats globaux

## Réalisation des tâches

- 166 tâches
- 93 manuelles
- 25 orales
- 48 multimodales



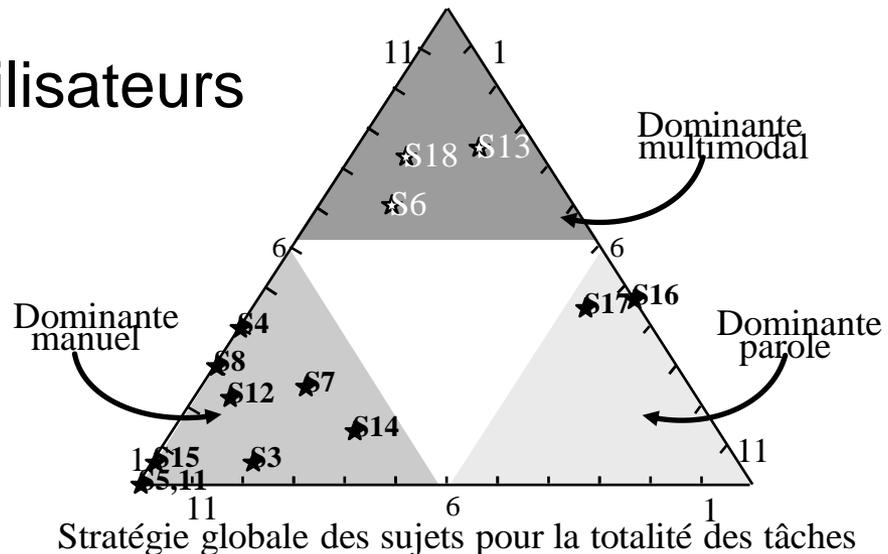
Pourcentage  
des différents types de tâches

- habitude des interfaces graphiques
- parole + multimodalité < manuel

# Stratégies des sujets

## Réalisation des tâches

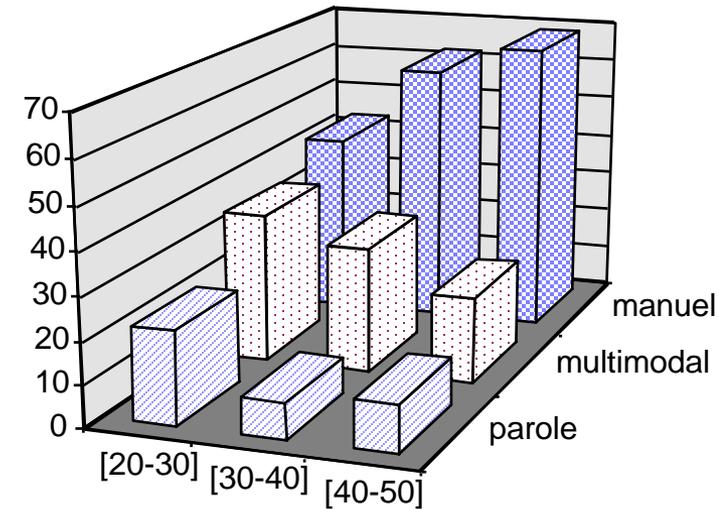
- Utilisation dominante d'une modalité
- Trois types d'utilisateurs



# Stratégies des sujets

## Réalisation des tâches

- pas de différence / sexe
- pas de différence / profession
- différence / âge



Utilisation des modalités  
en fonction de l'âge

**Influence du facteur "habitude"**

# Conclusion

## Réalisation des tâches

- Tâches multimodales : 29% ; tâches orales : 15%
- Stabilité entre les sessions
- Multimodalité et parole semblent convenir à la complexité sémantique, mais habitudes de communication par le graphique
- **Stratégie utilisateur**
  - Comportement inter-sujet variable ; influence de l'âge dans l'utilisation de la parole
- **Tâches multimodales**
  - Utilisation spontanée des deux modalités
  - Alternance de modalité comme mode de récupération des erreurs
  - Réalisations séquentielles ou alternées

**Abandon de la multimodalité dans ce projet**

**Le projet RICOM**  
**Recherche d'Information**  
**Collaborative et Multimodale**

# Gains attendus

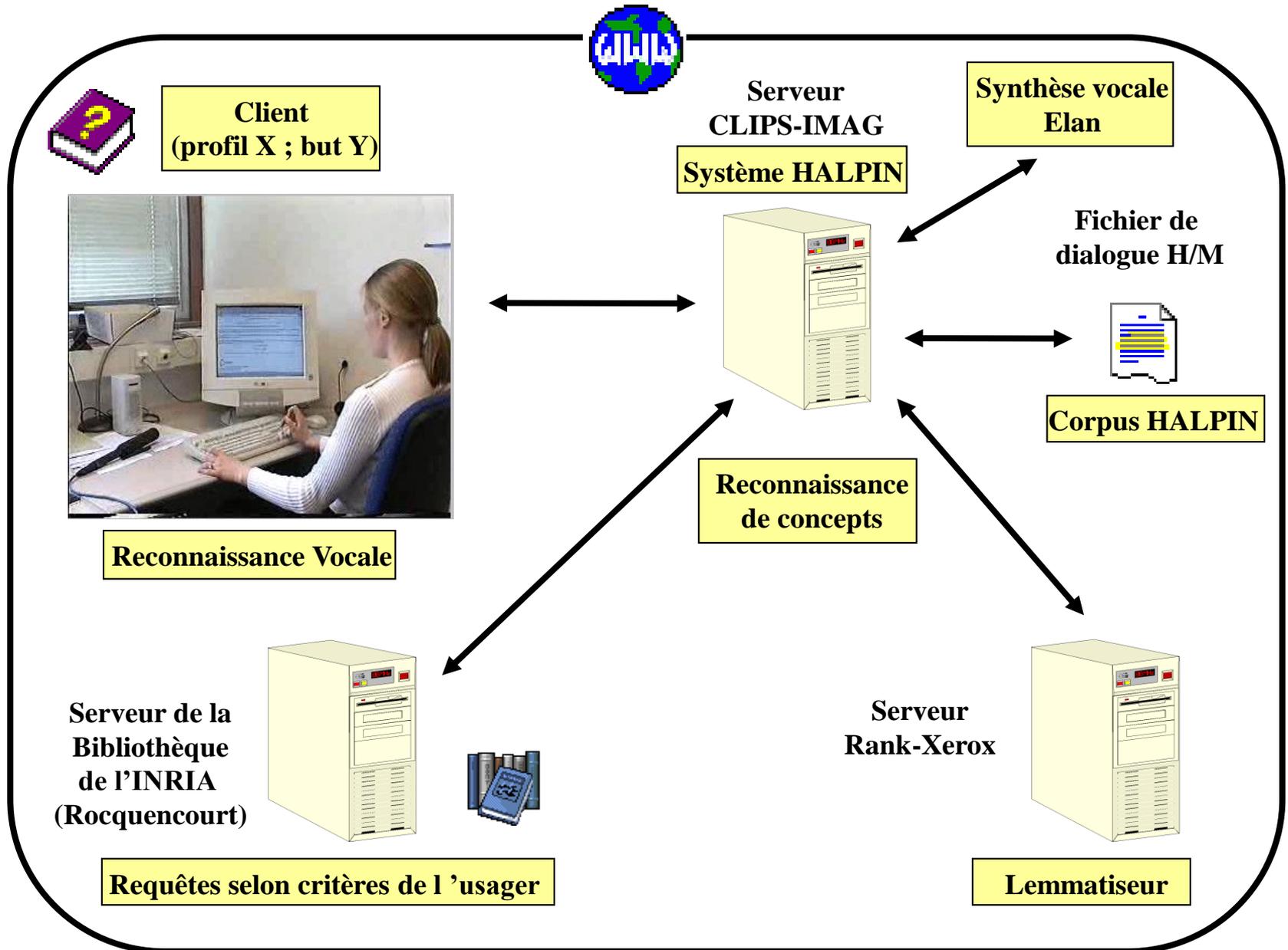
## Mode d'accès

- Navigation difficile => **multimodalité + dialogue**
- Expression limitée des requêtes => **souplesse du LN**
- Profils utilisateurs non pris en compte par les navigateurs actuels => **stratégies de dialogue**
- Univers référentiel et pragmatique limités et connus par la machine => **focalisation thématique**

## Buts des utilisateurs

- Butiner selon un fil thématique (motivation = curiosité)
- Accéder rapidement aux informations souhaitées (motivation = recherche d'un document donné ou recherche d'un ensemble de documents en relation thématique)

# Halpin : architecture du système



**Historique du dialogue**

que je vous aide à formuler votre requête ?  
[H:1] ben pourquoi pas  
[M:2] Pouvez-vous me donner le thème recherché, entre guillemets s'il vous plaît ?  
[H:2] je recherche de documents relatifs à " Internet " SVP  
[M:3] J'ai trouvé 145 documents ayant comme thème Internet .  
Voulez-vous affiner votre requête ?  
[H:3] j' aimerais bien  
[M:4] Afin d'affiner, vous devez choisir parmi les critères : Auteur, Titre, Thème, Année, ou Type de document.  
[H:4] je vous affiner avec le thème  
[M:5] Le thème actuel est Internet . On pourrait y ajouter par exemple communication par satellite ; transmission optique ; atm ; pcs ; technologiecommunication ; . Pour ajouter un thème, donnez le entre guillemets, sinon répondez Non pour passer à autre chose.  
[H:5] je souhaite ajouter " communications par satellite "  
[M:6] J'ai trouvé 2 documents ayant comme thème Internet communications par satellite . Consultez-les sur la page de droite. Est-ce que cela vous convient ?

**Réponse de la machine (hyperliens)**

1. [1997 IEEE International Conference on Communications towards the knowledge millenium : conference recor - IEEE - 1997](#)
2. [Computers and communications proceedings - IEEE computer society pre - 1997](#)

**Zone de détail d'un document**

Congrès : [IEEE international conference on communications ; 1997 ; Montréal, CA](#)  
Titre : 1997 IEEE International Conference on Communications  
Titre : towards the knowledge millenium : conference record  
Titre : vol. 1 of 3

parfaitement

Année en cours = [vide]  
Auteur en cours = [vide]  
Thème en cours = Internet communications par satellite  
Titre en cours = [vide]  
Type en cours = Tous les documents

Ré-écouter le son  
Stopper le son  
Profondeur de Recherche = 5

**Reconnaissance Vocale pour HALPIN**

Quitter

je recherche de documents relatifs à " Internet " SVP j' aimerais bien je vous affiner avec le thème je souhaite ajouter " communications par satellite " parfaitement

Connect Disconnect Begin Dict Stop Dict Direct Indirect

**Boutons pour les interactions vocales**

**Reconnaissance vocale IBM ViaVoice**

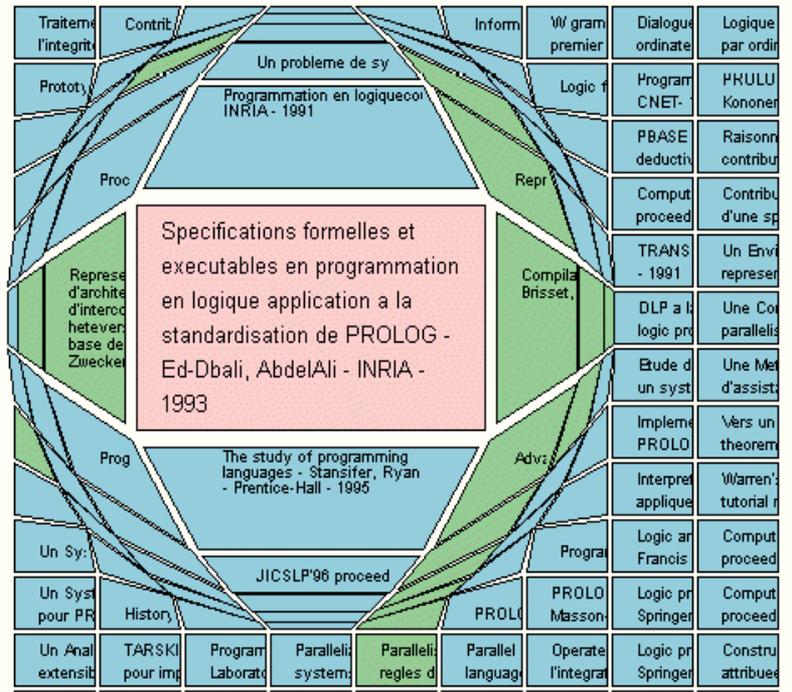
- Boîte d'historique du dialogue Halpin -  
 [M:\*] Bienvenue sur le système HALPIN Documentaire.  
 [M:0] Que souhaitez-vous faire ? 1 : Retrouver un document que vous connaissez, 2 : Rechercher un ensemble de documents, ou 3 : Découvrir ce site ?  
 [H:0] L'option numéro 2 je vous prie  
 [M:1] Rechercher un...  
 D'accord. Souhaitez-vous  
 formuler votre requête  
 [H:1] avec plaisir  
 [M:2] Pouvez-vous  
 entre guillemets s'il  
 [H:2] oui, c'est "Prolog"  
 [M:3] J'ai trouvé 271 documents ayant comme thème Prolog. Voulez-vous affiner votre requête ?

**Présentation des résultats en œil de poisson (avec hyperliens)**

**oui, avec la date**

Année en cours = (vide)  
 Auteur en cours = (vide)  
 Thème en cours = Prolog  
 Titre en cours = (vide)  
 Type en cours = Tous les documents

Ré-écouter le son  
 Stopper le son  
 Profondeur de Recherche = 5



**Zone de détails d'un document**

Auteur(s) : [Morel, Eric](#)  
 Titre : Environnement de programmation parallele  
 Titre : application au langage Prolog  
 Titre : Eric Morel  
 Éditeur : TMAG

# Halpin : Evaluation du système avec des sujets

## *Hypothèses a priori :*

- Choix **dialogue / navigation**, *nous pensons que le dialogue* sera préféré dans le cas d'une **recherche ciblée** de document(s) et la navigation dans tous les autres cas.
- Choix **affichage de liste / affichage en œil de poisson**, *nous pensons qu'à partir d'un certain nombre de documents (à déterminer), l'œil de poisson est plus pertinent* du fait que l'utilisateur a une vue plus globale sur l'ensemble des résultats et qu'il peut facilement **naviguer** sur les différents ouvrages.

# Halpin : l'expérience

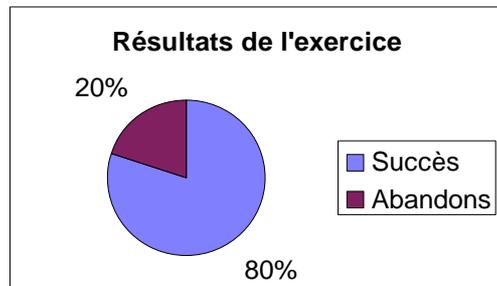
- **Test d'hypothèses**
  - dialogue vs navigation
  - présentation linéaire vs affichage en œil de poisson
- **10 sujets (+ 2 en pré-tests)**
- **5 tests (exercices) étaient proposés aux sujets**
- **Chaque exercice a un but précis :**
  - test 1** : informations précises
  - test 2** : informations imprécises
  - test 3** : informations dérivées
  - test 4** : informations enfouies (sous-thèmes)
  - test 5** : présentation linéaire classique vs. Œil de poisson
- **Questionnaire en fin de session**

# Exemple : Halpin, test 4

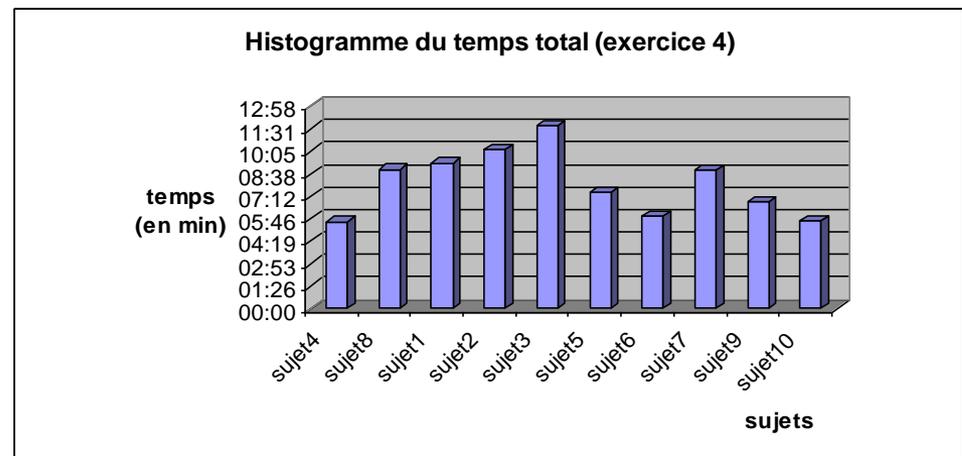
Consigne : Trouvez 5 sous-thèmes différents, dans les documents ayant pour thème la *programmation par objets*.

But : tester la stratégie d'affinement (dialogue vs. navigation)

Stratégie : tous les sujets dialoguent (recherche du thème) puis affinent sur l'œil de poisson.



L'affinement se fait par navigation. 2 abandons dénotent le conflit dialogue/navigation



# Halpin : résultats d'évaluation

- **Les sujets privilégient chaque fois que cela est possible l'utilisation du dialogue pour exécuter les différentes tâches** (ces résultats ont été confortés par le sondage). Cela ne dépend pas de la tâche mais plutôt de l'utilisateur.
- La **prédominance relative du dialogue** est peut-être due (??) :
  - ❖ à l'attrait de la synthèse vocale (qui favorise l'interaction verbale),
  - ❖ à la difficulté de changer de modalité une fois le dialogue engagé (l'interaction commence toujours par le dialogue),
  - ❖ à la facilité d'adaptation aux formes d'interaction dialogiques et à la maîtrise de la langue maternelle, qui ne nécessitent pas d'apprentissage particulier.
- L'**indifférence à la forme de présentation** des résultats (présentation linéaire vs. Œil de poisson)

# Perspectives de R&D : trois points clés

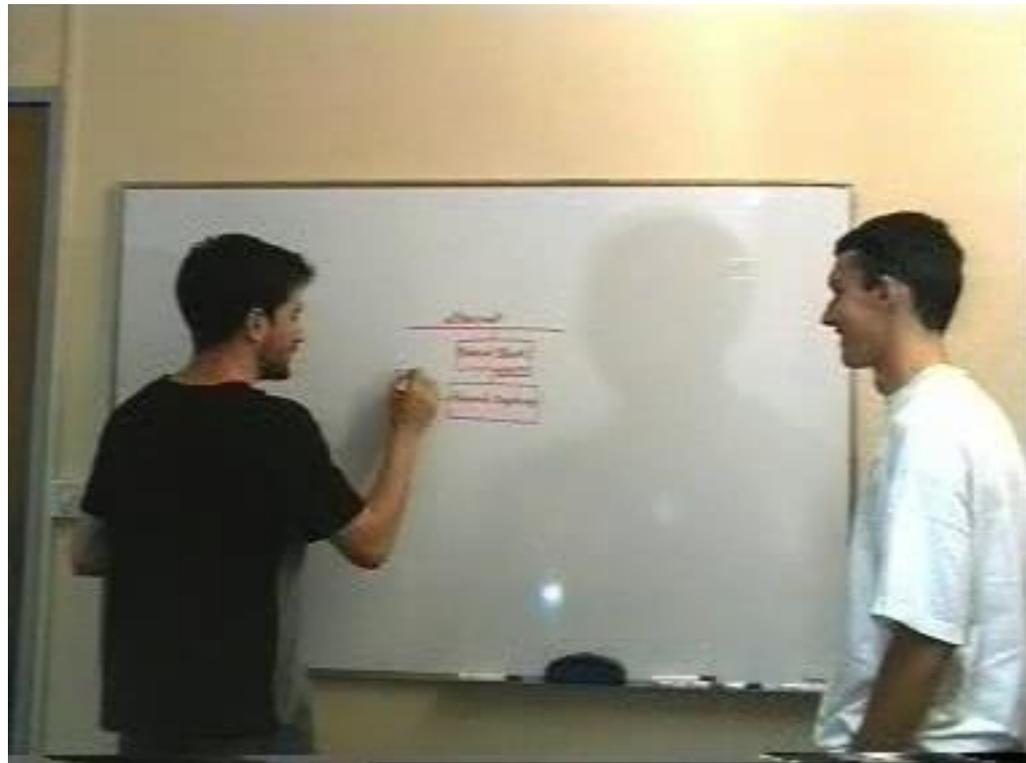
- **Physicalité**
- **Mobilité**
- **Interactivité plastique**

# Physicalité

- **Les principes de la Réalité Augmentée**
  - Conservation des objets qui nous sont familiers
  - Amplification fonctionnelle par le calcul électronique
- **L'ordinateur évanescent mais doué d'ubiquité**
  - Disparition de la boîte grise
  - Capacités de calcul réparties dans l'environnement

# Physicalité

- Exemple : Le Tableau Magique



# Mobilité

- **Mobilité de l'utilisateur mais le dispositif d'interaction est fixe (bornes interactives)**
  - **contexte d'interaction connu et système figé**
  - **migration des données personnelles de l'utilisateur**

# Mobilité

- **Mobilité de l'utilisateur et du dispositif d'interaction**

- **contexte d'interaction dynamique**
- **tâches dépendantes de la localisation.**
- **variabilité de la pertinence d'une tâche et des concepts observables**

- **Modélisation de modalités : équivalence fonctionnelle avec adaptation dynamique**

- **modalité des retours d'informations**
- **modalité selon la connectivité : dégradation courtoise**

# Interactivité « plastique »

- **Interactivité**
  - En grand : le mur augmenté (cave)
  - En petit : le téléphone portable
- **Tâche accomplie avec**
  - un seul système à la fois (en grand ou en petit)  
en équivalence fonctionnelle
  - plusieurs systèmes à la fois (en grand et en petit)  
en complémentarité ou en redondance

# Interactivité plastique

- **La plasticité des Interfaces : analogie avec certains matériaux**
- **Une forme d'adaptation**
  - **au contexte d'interaction**
  - **à la variété des dispositifs d'interaction**
- **La multimodalité est un élément clé dans la plasticité**