

# Modélisation de la communication multimodale

*Vers une formalisation de la pertinence*



Frédéric Landragin

2 avril 2003

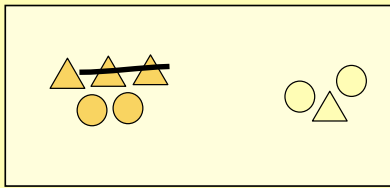


2

## Plan de l'exposé

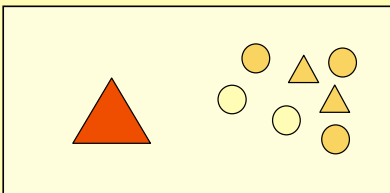
- Le problème : la résolution de la référence multimodale en modélisant les modalités et les facteurs cognitifs impliqués
- Les concepts :
  - les domaines de référence
  - la pertinence
  - la saillance
- Le modèle :
  - la résolution de la référence
  - l'architecture logicielle
  - l'instanciation du modèle
- Conclusion et perspectives

## Problème : plusieurs modalités...



« ces trois triangles » → {▲, ▲, ▲}

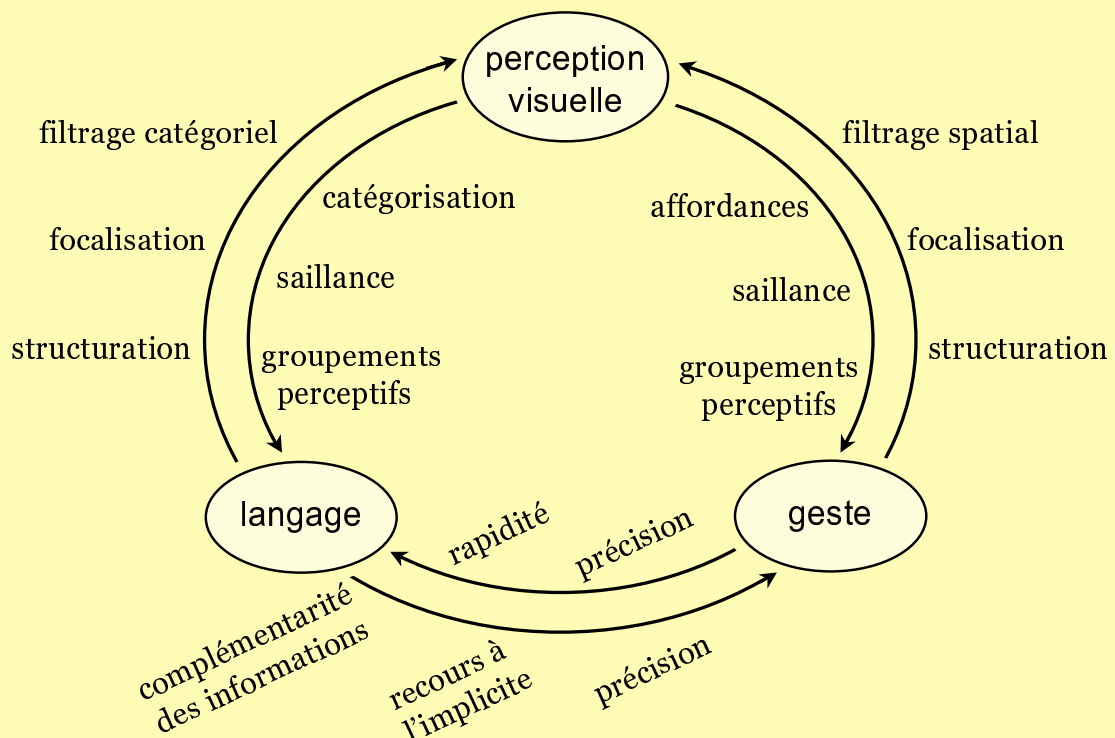
« les deux cercles » → {●, ●}



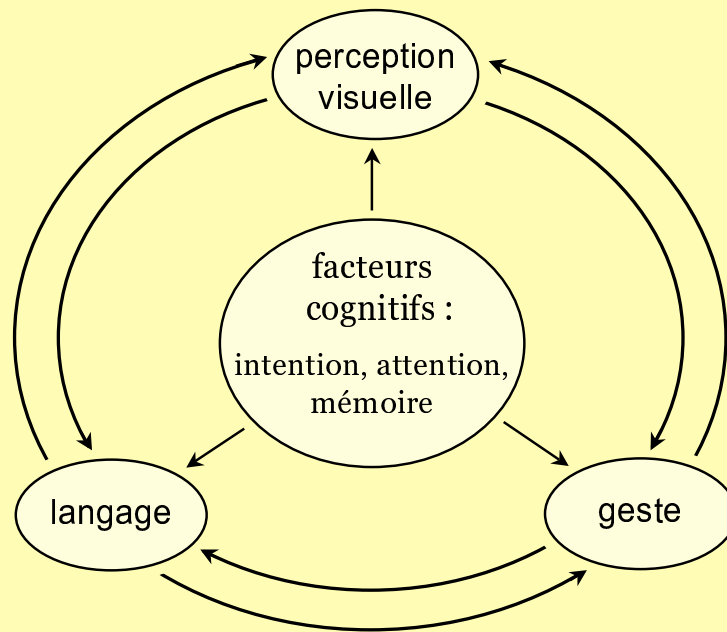
« le triangle » → {▲}

⇒ Analyse et modélisation des modalités

## ...qui interagissent



## ...qui interagissent



⇒ Analyse et modélisation des facteurs cognitifs

## Approches (1)

- Du côté de l'informatique : trop informatique...
  - peu d'exploitation des finesses du langage ou de l'implicite [Reithinger 87]
  - pas de fusion multimodale à un niveau sémantique [Oviatt 97], [Coutaz 91]
  - la résolution de la référence n'est pas le problème prioritaire
  - quelques modèles d'architecture [Sabah 96]
- Du côté de la linguistique : trop linguistique...
  - pas d'interaction fine entre geste et langage, contexte visuel négligé
  - geste souvent considéré comme déterminant directement les référents
  - quelques modèles pour la détermination [Corblin 87], [Kleiber 94]
- Du côté de la psychologie : trop flou...
  - beaucoup de théories, beaucoup d'incompatibilités, peu de formalisations
  - quelques principes de base importants [Wertheimer 23], [Miller 56]

## Approches (2)

- Du côté des approches pluridisciplinaires :  
modélisation de facteurs cognitifs dans un cadre rigoureux
  - [Clark 83, 86] = référence aux objets vue comme un processus coopératif soumis au principe de l'effort minimal ; influence de la saillance lors de la production d'expressions référentielles
  - [Grosz & Sidner 86] = modèle gérant une structure discursive, une structure intentionnelle et une structure attentionnelle
  - [Wright 90] ; [Haji\_ová 93] ; [Heeman & Hirst 95] = saillance
  - [Beun & Cremers 98] = modèle gérant saillance, attention spatiale et attention fonctionnelle

Mais pas de véritable intégration des 3 modalités...

## Objectifs

- Modélisation cognitive du processus d'interprétation des références multimodales :
  - tenant compte de l'interaction entre modalités
  - tenant compte des facteurs cognitifs
  - fondée sur le modèle des « domaines de référence »
  - avec une approche pluridisciplinaire
- Elaboration de systèmes de dialogue homme-machine :
  - formalisation la plus fidèle possible de théories cognitives
  - intérêt particulier pour l'intégration

## Les domaines de référence (1)

- L'identification des référents passe par l'identification du sous-ensemble contextuel auquel ils appartiennent
- Différentes sources de contexte se confrontent, chacune pouvant donner lieu à une structuration en sous-contextes
  - le contexte langagier (l'historique du dialogue)
  - le contexte visuel partagé (le support de la communication)
  - le contexte de tâche (en dialogue homme-machine finalisé)
- Mêmes caractéristiques  $\Rightarrow$  confrontation, intégration  
[Landragin, Salmon-Alt & Romary 02]

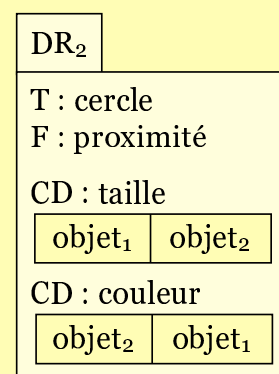
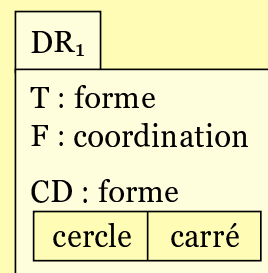
## Les domaines de référence (2)

T = type

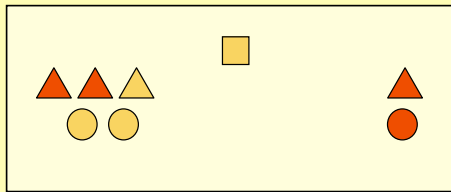
F = facteur de groupement

CD = critère de différenciation

partitions



## Construction de domaines visuels



- Groupement par proximité

niveau 1 : 5 + 1 + 2

niveau 2 : 6 + 2

- Groupement par similarité

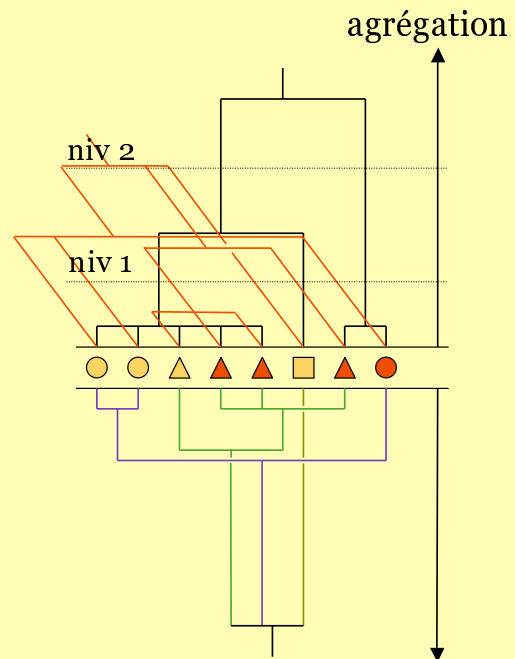
niveau 1 : 3  $\blacktriangle$  + 2  $\circ$  + 1 + 1 + 1

niveau 2 : 4  $\triangle$  + 3  $\circ$  + 1

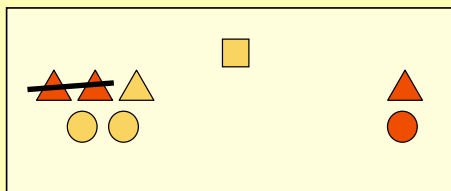
- Groupement par continuité

niveau 1 : 3  $\triangle$  + 1 + 1 + 1 + 1 + 1

niveau 2 : 4  $\triangle$  + 3  $\circ$  + 1



## Construction de domaines visuels



- Groupement par proximité

niveau 1 : 5 + 1 + 2

niveau 2 : 6 + 2

- Groupement par similarité

niveau 1 : 3  $\blacktriangle$  + 2  $\circ$  + 1 + 1 + 1

niveau 2 : 4  $\triangle$  + 3  $\circ$  + 1

- Groupement par continuité

niveau 1 : 3  $\triangle$  + 1 + 1 + 1 + 1 + 1

niveau 2 : 4  $\triangle$  + 3  $\circ$  + 1

sim niv 1 + prox niv 1



sim niv 2 + prox niv 1



sim niv 2 + prox niv 1 + cont niv 1



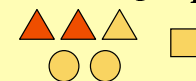
sim niv 2 + prox niv 2 + cont niv 1



sim niv 2 + prox niv 2 + cont niv 2



sim niv 3 + prox niv 2 + cont niv 3



## La pertinence

- La théorie de la Pertinence [Sperber & Wilson 95]
  - la présomption de pertinence : le locuteur n'a pas seulement l'intention de transmettre un message, il a aussi l'intention de le transmettre de telle façon que son intention soit reconnue
    - ⇒ l'interlocuteur s'appuie sur l'hypothèse que le message est pertinent.
  - la nature ostensive et inférentielle de la communication : l'interprétation fait intervenir des mécanismes de **décodage** du message et d'**inférence**.
- Les notions d'effets contextuels et d'effort de traitement
  - une hypothèse est d'autant plus pertinente dans un contexte donné que ses effets contextuels y sont plus importants ;
  - une hypothèse est d'autant plus pertinente dans un contexte donné que l'effort nécessaire pour l'y traiter est moindre.

## La pertinence pour classer des domaines

- **Problème** : ordonnancement des domaines de référence construits
- **Approche** :  
exploiter un critère de pertinence
- **Méthodologie** :
  - faire jouer à la fois un critère de simplicité cognitive et d'informativité
  - choix du critère de la Pertinence avec les notions d'effets et d'effort
  - calcul du ratio : effets/effort
 [Landragin, De Angeli, Lopez, Wolff & Romary 02]

$$\frac{\text{effets contextuels}}{\text{effort de traitement}} \quad \begin{array}{l} \text{(fonction des possibilités offertes)} \\ \text{(fonction de la complexité du contexte)} \end{array}$$

## La saillance pour minimiser l'effort

- **Problème** : identification des facteurs de saillance
- **Approche** :  
appréhender simultanément saillance visuelle et saillance linguistique
- **Méthodologie** :
  - confrontation des diverses définitions de la saillance
  - confrontation des diverses caractérisations de la saillance
  - identification de critères génériques de saillance
  - caractérisations des saillances visuelle et linguistique
  - quantification des saillances visuelle et linguistique
  - exploitation de cette quantification

## Définition de la saillance

- **Angle d'attaque perceptif** :  
saillance = degré d'attention accordé aux entités perçues.  
  
Capte l'attention, arrive en premier à l'esprit. Facteurs de saillance : perceptivité (identifiabilité) ; simplicité naturelle ; clarté ; stabilité ; originalité (entité nouvelle) [Osgood & Bock 87], [Alquier 98]
- **Angle d'attaque cognitif** :  
saillance = degré d'attention accordé aux entités représentées mentalement.  
  
Reste activée dans la mémoire à court terme. Facteurs de saillance : intention communicative ; focus d'attention ; familiarité ; accessibilité ; aspects perturbateur, curieux, énigmatique (entité qui pose un problème) [Loftus & Mackworth 78], [Stevenson 02]



## Caractérisation générique de la saillance

Facteur de saillance	Exemple visuel	Exemple linguistique
critère intrinsèque aux unités mise en évidence conventionnelle placement à un endroit stratégique isolement répétition symétrie rupture dans une continuité infraction à une règle implicite exploitation d'une norme structuration du message	grande taille brillance tiers du cadre objet isolé des autres régularité symétrie disposition linéaire disposition incongrue composition classique ligne directrice	nom propre présentatif, accent début de l'énoncé apposition répétition chiasme rythme d'élocution raté, erreur construction usuelle focus, topic

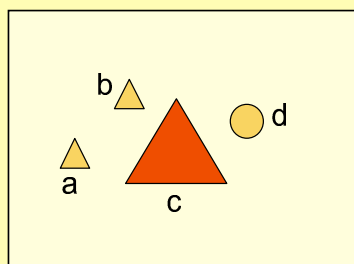
## Caractérisation de la saillance visuelle

- Des critères physiologiques :
  - la saillance liée à la fovea
  - la saillance liée aux temps de latence
- Des critères liés à la structure de la scène visuelle :
  - la saillance liée aux zones qui attirent le regard
  - la saillance liée à la perspective
  - la saillance liée aux symétries ou aux équilibres
- Des critères liés au rendu visuel des objets :
  - la saillance due à une mise en évidence explicite
  - la saillance liée à la catégorie et aux caractéristiques physiques
  - la saillance liée à la localisation dans la scène
  - la saillance liée à l'incongruité ou à la dynamique

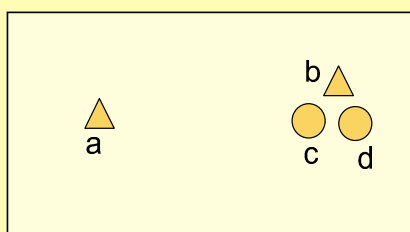
# Caractérisation de la saillance linguistique

- Des critères liés à la forme de l'énoncé :
  - la saillance intrinsèque au mot
  - la saillance par une mise en avant explicite lors de l'énonciation
  - la saillance par une construction syntaxique dédiée
  - la saillance syntaxique liée à l'ordre des mots
  - la saillance liée aux fonctions grammaticales
  
- Des critères liés au sens de l'énoncé :
  - la saillance liée au rôle thématique
  - la saillance liée au thème
  - la saillance liée aux inférences pour l'identification de l'implicite
  - la saillance liée au sujet de la conversation

## Scores pour la saillance visuelle



	catégorie	taille	couleur	disposition	
a:	0	0	0	0	→ 0
b:	0	0	0	0	→ 0
c:	0	1	1	0	→ <b>0.5</b>
d:	1	0	0	0	→ 0.25



a:	0	0	0	1	→ <b>0.25</b>
b:	0	0	0	0	→ 0
c:	0	0	0	0	→ 0
d:	0	0	0	0	→ 0

## Scores pour la saillance linguistique

A. « le triangle rouge se met à côté du bleu »

Facteur	triangle rouge	triange bleu
construction syntaxique dédiée	0	0
place stratégique en début d'énoncé	1	0
fonction grammaticale sujet	1	0
rôle thématique privilégié	1	0
thème	1	0
<b>total de saillance ramené entre 0 et 1</b>	<b>0.8</b>	<b>0</b>

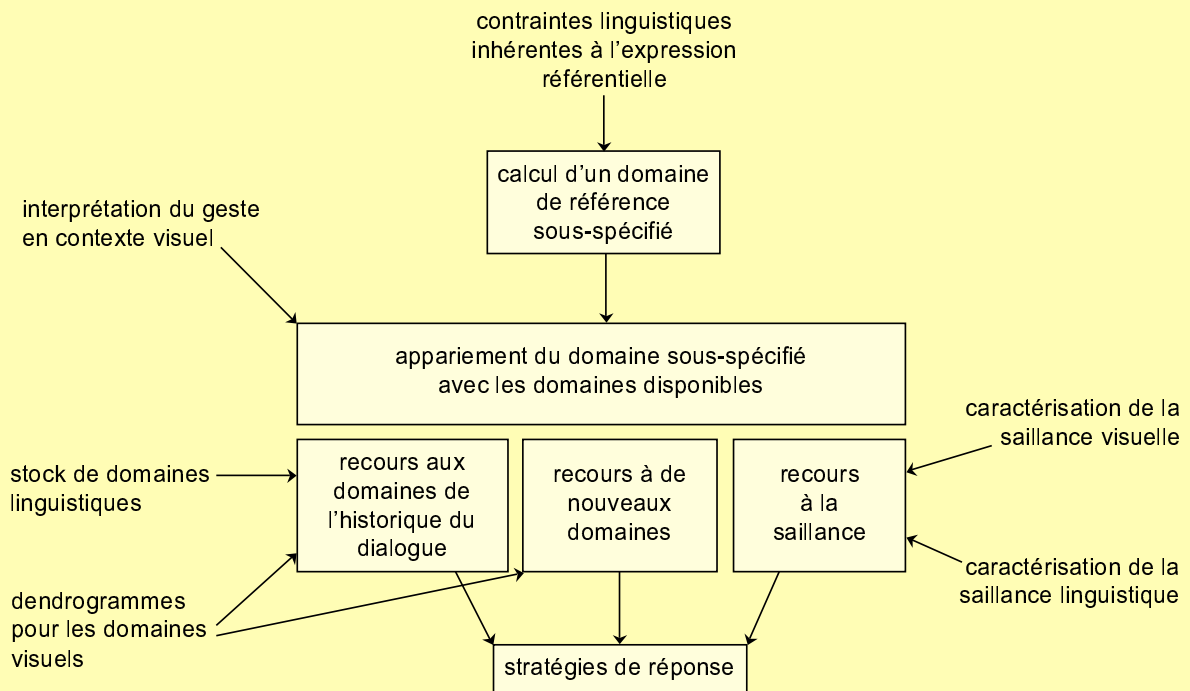
B. « c'est le triangle rouge que tu dois mettre à côté du bleu »

construction syntaxique dédiée	1	0
place stratégique en début d'énoncé	1	0
fonction grammaticale sujet	0	0
rôle thématique privilégié	1	0
thème	1	0
<b>total de saillance ramené entre 0 et 1</b>	<b>0.8</b>	<b>0</b>

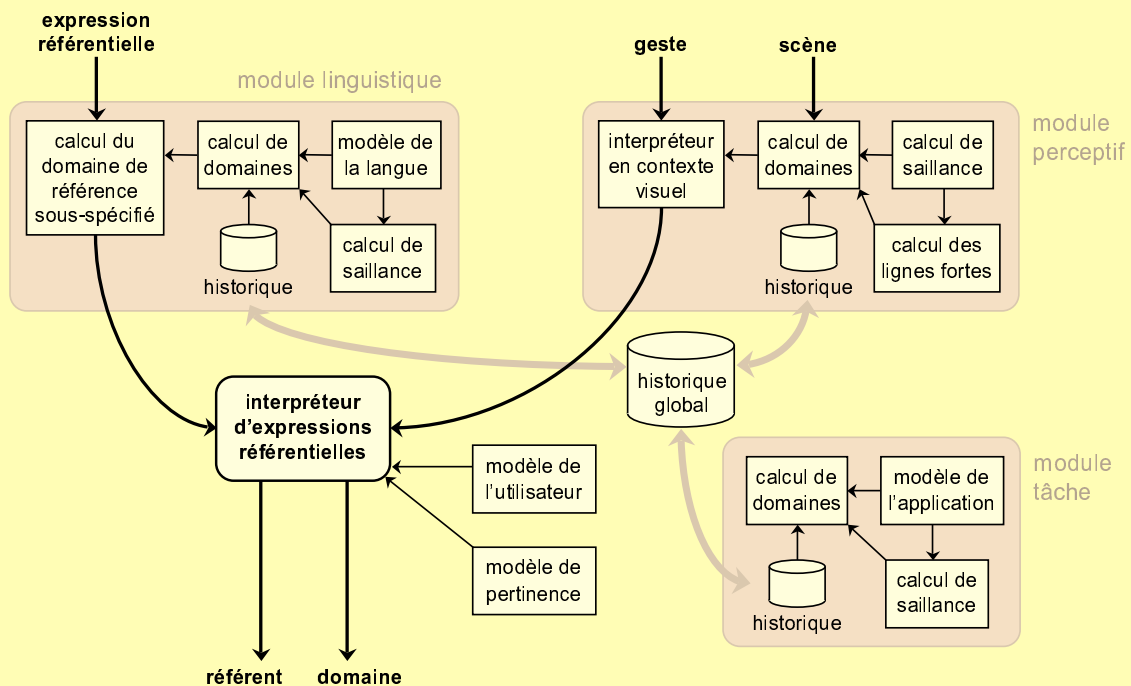
## Exploitation de ces scores

- Interpréter une expression référentielle langagière ambiguë compte tenu du contexte :
  - « le N » dans une scène visuelle comprenant plusieurs N, un seul d'entre eux étant saillant
  - « le N » dans un contexte d'interaction où un N vient d'être manipulé ou vient d'être mentionné (anaphore)
  - « le N » compris comme « le N suivant dans la succession des tâches » dans un contexte applicatif particulier
- Interpréter une expression référentielle multimodale ambiguë compte tenu du contexte
- Prédire sur quelle entité l'utilisateur va se focaliser et va construire son énoncé

# Résolution des références



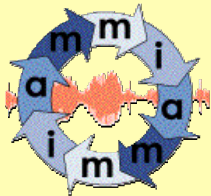
# Architecture modulaire



## Validation de certains aspects du modèle

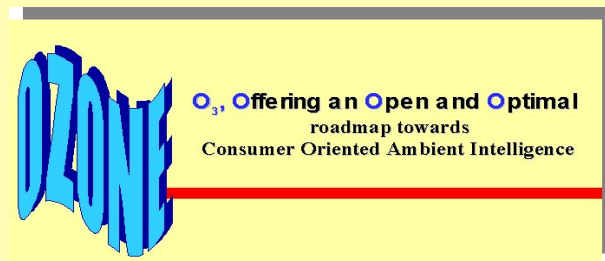


Le projet *ACTS COVEN* :  
Collaborative Virtual ENvironments



Le projet *IST MIAMM* :  
Multidimensional Information Access  
using Multiple Modalities

Le projet *IST OZONE*



## Validation de certains aspects du modèle

- Validation de l'existence des domaines de référence
  - COVEN** : groupement selon le critère de proximité
  - MIAMM** : application à la modalité haptique
  - OZONE** : domaines liés à la tâche applicative
- Validation de l'intervention de la saillance
  - COVEN** : ajustement des critères de saillance ; adaptation à la 3D
  - MIAMM** : application à la modalité haptique
  - OZONE** : application à des scènes complexes (cartes géographiques)
- Validation de la méthode de résolution de la référence
  - COVEN** : prise en compte de la saillance
  - MIAMM** : application à la modalité haptique ; aspects architecturaux
  - OZONE** : mise en place d'un démonstrateur

## Conclusions

- **Intérêt d'une approche pluridisciplinaire :**
  - confronter les apports de chaque discipline sur un même sujet
  - constituer un modèle plus large au niveau des phénomènes traités
  - apporter des perspectives de formalisation
  
- **Intérêt des notions proposées :**
  - prendre en compte plusieurs domaines permet d'identifier les ambiguïtés dues au langage et à la multimodalité
  - prendre en compte un critère de pertinence permet de ne garder que les ambiguïtés pertinentes, de fournir au système un critère de décision
  - prendre en compte la saillance permet de proposer une solution dans une situation ambiguë
  
- **Adéquation du modèle avec la Théorie de la Pertinence**

## Perspectives

- **Extension du modèle**
  - meilleure formalisation de la Gestalt et de la saillance
  - prise en compte des références spatiales
  - référence aux événements, actes de langage
  
- **Validation et implantation**
  - continuation de la boucle : simulation/modélisation
  - expérimentations psycholinguistiques (saillance, pertinence)
  - implantation d'une architecture de type « carnet d'esquisses »
  
- **Vers un nouveau modèle**
  - vers un modèle combinant compréhension et génération automatique
  - vers un modèle cognitif de la communication multimodale
  - vers une quantification de la pertinence