

Théorie de l'Esprit & Multimodalité

Gérard Bailly, GIPSA-Lab

Sylvie Pesty, LIG-Lab

2016-2017

Introduction

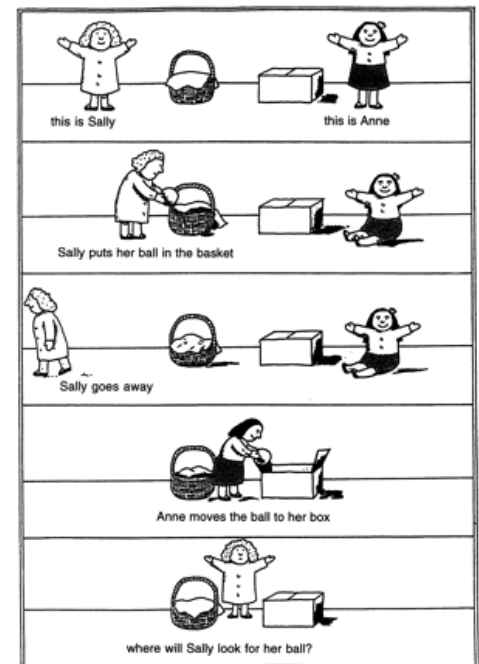
- Les actions d'autrui sont plus que des formes spatiotemporelles ; elles reflètent , et permettent de prévoir, leurs états internes

Comment infère-t-on les causes des actions d'autrui?

Psychologie sociale:
Théorie de l'attribution

Psychologie développementale:
Théorie de l'esprit

- Théorie de l'attribution: généraliser, affecter des propriétés *stables* aux observations
 - Objets, agents, situations
 - « Jean rit de bon cœur à cette histoire »
 - Est-ce que tout le monde rit à cette histoire ?
 - Jean rit-il de bon cœur à toutes les histoires?
 - Jean rit-il toujours de bon cœur à cette histoire ?
- Théorie de l'esprit: prédire l'action
 - Différencier le monde sensible et ses représentations mentales
 - Jeux de théorie de l'esprit (cf. Sally-Anne)



Théorie de l'attribution

- Définitions

- Attribution causale (Heider, 1944): recherche des causes d'un évènement, « décrire les processus par lesquels le sujet naïf construit ses explications, comment ses explications influencent en retour ses réactions émotionnelles ou cognitives vis-à-vis de l'évènement ou des personnes qui en sont à l'origine, mais aussi quelles difficultés il rencontre face à la complexité même du phénomène d'explication »
- Modèle attributionnel (Weiner, 1986): causalité interne/externe, in/stable, in/contrôlable
- Analyse SWOT d'une personne, d'une situation...

Attribution	INTERNE		EXTERNE	
	STABLE	MODIFIABLE	STABLE	MODIFIABLE
CONTRÔLABLE	Méthode de travail, organisation, Stratégie d'apprentissage	Travail accompli, concentration	Culture, milieu social et familial	Perception de soi par les autres
INCONTROLABLE	Intelligence, don,...	Maladie, Trac, surmenage	Difficulté de l'activité	Chance

- Implications

- Cause « stable »: permanente et susceptible de se maintenir dans le temps: trait de personnalité, niveau d'intelligence -> modèle BDI (croyances, désirs, intentions...)
- Schème (Kant): représentation mentale qui joue le rôle d'intermédiaire entre les catégories de l'entendement et les phénomènes sensibles; "noyau" de savoir-faire (Piaget)
- Traitement/inférence
 - Modèle de la covariation (Kelley, 1967): « l'effet (l'évènement observé) est attribué à la cause qui est présente quand l'effet est présent, et absente quand l'effet est absent »
 - Théorie des inférences correspondantes (Jones et Davis, 1965) & schémas de causalité (Kelley, 1972) : structure de connaissance qui dépassent l'observation: de « Richard demande à son oncle de lui prêter de l'argent », on infère que celui-ci veut réaliser un projet pour lequel il n'a pas suffisamment d'argent, que l'oncle a beaucoup d'argent et qu'il a l'habitude de prêter à Richard...

Modèles d'attribution sociale

- Heider (1920)
 - Motivations, intentions, sentiments... causes du comportement
 - Distinction entre causes internes (endogènes) & externes (exogènes)

- Wroom (1932)
 - Théorie des attentes
 - Motivation



- Théorie de la covariation
 - L'individu s'appuie sur trois sources d'informations pour expliquer/anticiper un comportement
 - Consistance : comportement habituel de la personne dans mêmes situations et à d'autres moments
 - Consensus : comparaison avec le comportement d'autres individus dans la même situation
 - Différentiation : comparaison du comportement de la personne dans d'autres situations
 - Statistiques
 - Préjugés, identité sociale, etc

Théorie de l'esprit

- Définitions

- Esprit: croyances, désirs, émotions, perceptions & intentions
- Théorie de l'esprit: capacité à attribuer ces états mentaux à soi et à autrui afin d'interpréter et prédire les comportements

- Implications

- Implique de différencier le monde et ses représentations mentales
- La lect. d'esprit (mind reading) s'appuie sur la perception de l'état mental d'autrui à travers ses (ré)actions -> modules bas-niveau
 - Distinction agents/objets
 - Objets: lois de comportements/affordances, etc.
 - Agents: regard, expressions faciales, etc
- Développement
 - Apprentissage : Sally-Anne vers 4-5 ans
 - Autisme: lecture d'esprit défaillante -> défaillance d'apprentissage de la conduite sociale

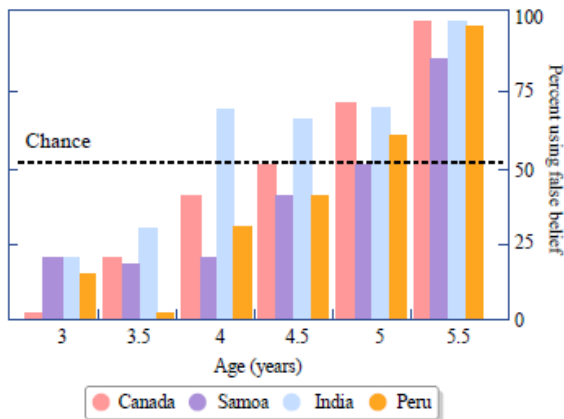
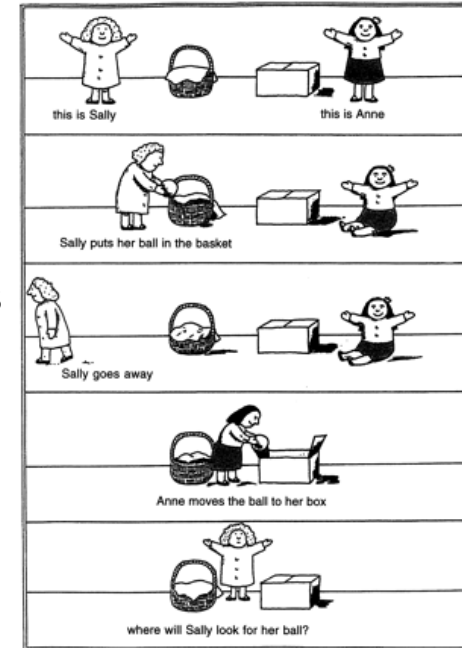


TABLE 1. Approximate developmental timeline of some aspects of theory of mind, with illustrative references.

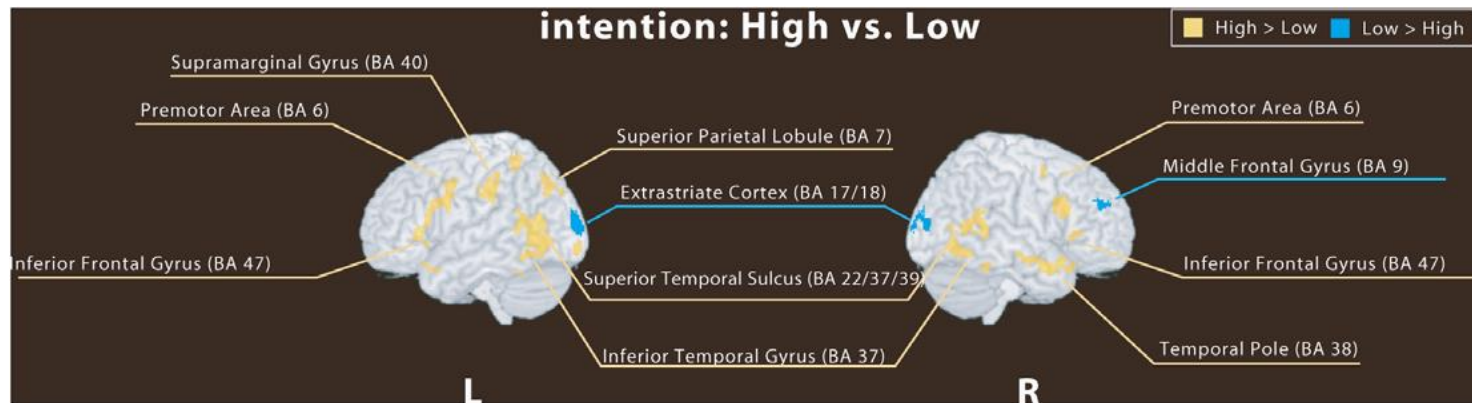
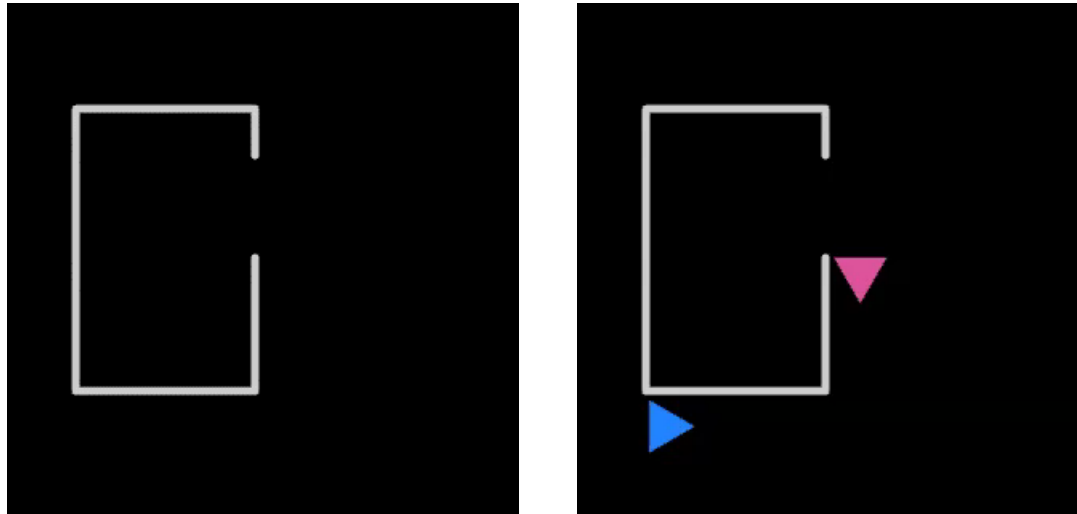
Age	Aspects of theory of mind
6–12 months	<ul style="list-style-type: none"> • Joint attention, including gaze and point following, and alternation of gaze between person and object (Bruinsma et al., 2004; Carpenter et al., 1998) • First words (Tomasello, 1995)
13–24 months	<ul style="list-style-type: none"> • Recognize intentionality in others as demonstrated in word use (Tomasello, 1995) • Recognize that others have desires different from one's own (Repacholi & Gopnik, 1997) • Early pretend play (Leslie, 1987)
30–36 months	<ul style="list-style-type: none"> • Begin to use mental state terms with truly mentalistic functions (Bartsch & Wellman, 1995) • Increasingly sophisticated pretend play (Youngblade & Dunn, 1995)
37–48 months	<ul style="list-style-type: none"> • Increasing ability to understand how things look from another's perspective (Flavell et al., 1981) • Begin to understand sentence complements (de Villiers & Pyers, 2002)
49–60 months	<ul style="list-style-type: none"> • Consistently pass false belief and appearance-reality tasks (Wellman et al., 2001)

Métacognition et théorie de l'esprit

- Théorie de l'esprit de l'enfant
 - 3 ans: son point de vue: un chien est heureux de trouver une voiture dans sa niche mais pas écuelle
 - 4 ans: perspective taking: dessin à l'envers ou à l'endroit [Melot, 2001]
 - 5 ans: expliquer les conduites erronées des autres non pas en fct de la réalité mais en fct de représentations erronées
- Fausses croyances
 - Capacité à prêter à une personne une pensée % à un evt objectif
 - Capacité à prêter à une personne une pensée en fct de la pensée d'une autre personne
 - « Mary pense que John pense que... »
- Métacognition: se représenter son propre esprit en train de représenter une info
 - une voiture rouge est passée ce matin
 - Conscience primaire: j'ai vu une voiture rouge
 - Mémoire: je me souviens que (j'ai vu une voiture rouge)
 - Méta-mémoire: je sais que (je me souviens que (j'ai vu une voiture rouge))
- Connaissance de soi et de l'autre se développent simultanément chez l'enfant
 - Biais du langage. Développement précoc: dès 15 mois [Onishi & Baillargeon, 2005]
 - Même réseau cortical. Deux alternatives [Dehaene, 2012]
 - soit nous disposons d'une représentation détaillée de nous-mêmes et utilisons ce réseau du « soi » pour **simuler** l'esprit des autres et tenter de le comprendre
 - soit nous ne disposons pas d'un système spécifique d'introspection: notre connaissance est fondée sur l'observation de « soi-même comme un autre »

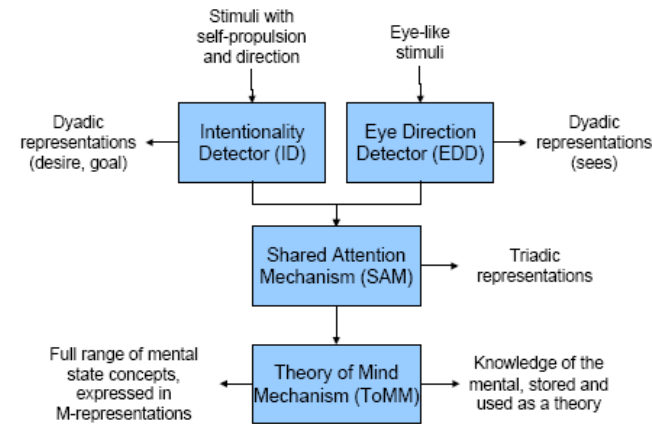
Interpréter toujours...

- Projeter des personnalités, des intentions
 - A des objets animés [Heider & Simmel, 1944; Osaka et al, 2012]



Modèles ToM

- Baron Cohen (1995)
 - ToBM (Theory of body motion)
 - lois physiques de causalité
 - distinction objet/sujet
 - EDD: Guidage de l'attention par tiers
 - ID: Répertoires
 - Implémenté par B. Scassellati (2001) pour robot



- Leslie (1994)
 - ToM et processus de sélection
 - Prédiction de l'action (que va faire l'autre?)
 - Croyance (que pense l'autre?)
 - Cognition sociale : lien avec modèles BDI (voir CM2: Architectures cognitives + Informatique affective)

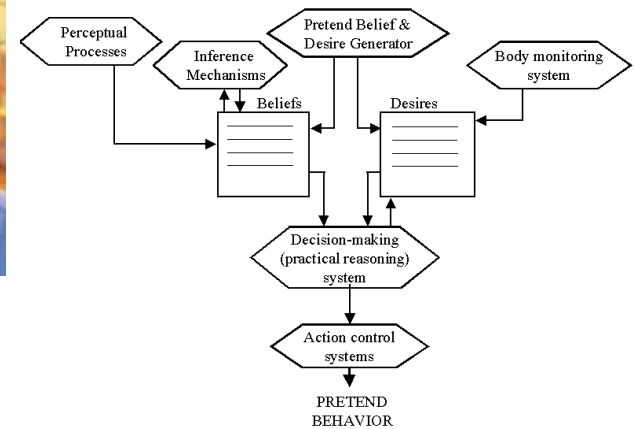
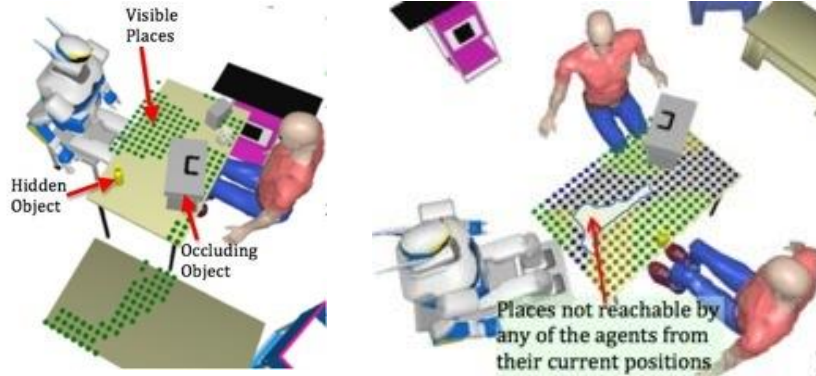


Figure 5

« Perspective taking »

- Plus largement, capacité de mettre à la place de quelqu'un
 - Champ de vision, espace de travail [Alami, 2007]



- Impact de la culture [Wu & Keysar, 2007]
- Quatre étapes (Winner, 2013)
 - Détection: M'a-t-il remarqué?
 - Motivation: pourquoi est-il là, que va-t-il faire?
 - Egocentré: qu'est-ce qu'il pense de moi?
 - Action: que puis-je faire pour influencer sur cet état mental?

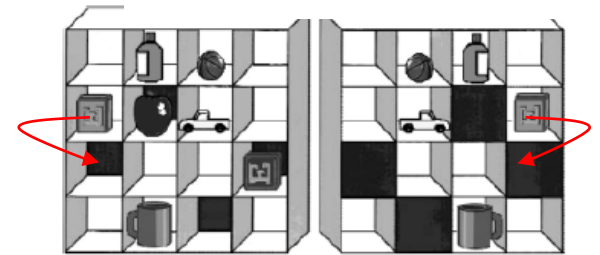
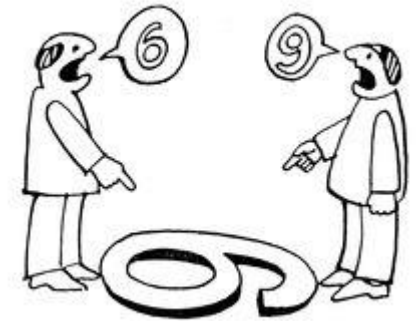


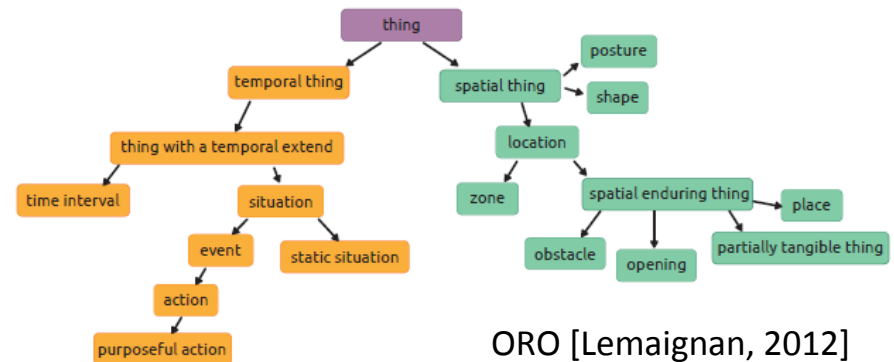
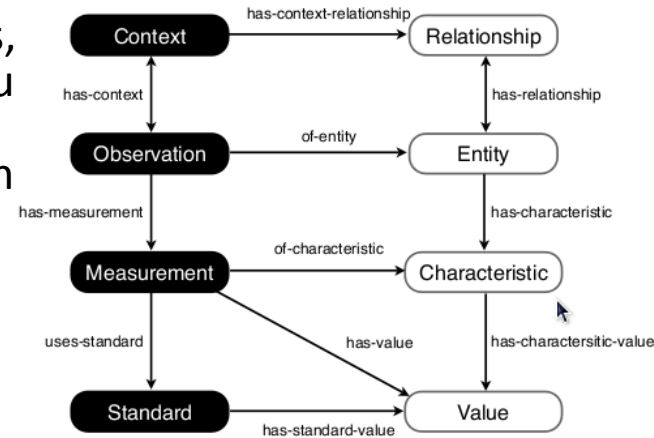
Fig. 1. Example of an array from the perspective of the director (right) and the perspective of the subject (left). The target is the mutually visible block (second row), and the competitor is the block (third row) visible only to the subject.

Connaissances - ontologies

- **Un ontologie** est l'ensemble structuré des termes et concepts représentant le sens d'un champ d'informations, que ce soit par les métadonnées d'un espace de noms, ou les éléments d'un domaine de connaissances. L'ontologie constitue en soi un modèle de données représentatif d'un ensemble de concepts dans un domaine, ainsi que des relations entre ces concepts. Elle est employée pour raisonner à propos des objets du domaine concerné.
- Les ontologies décrivent généralement :
 1. individus : les objets/agents de base
 2. classes : ensembles, collections, ou types d'individus
 3. attributs : propriétés, fonctionnalités, caractéristiques ou paramètres que les individus peuvent posséder et partager
 4. relations : les liens que les individus peuvent avoir entre eux
 5. événements : changements subis par des attributs ou des relations

- Langages et éditeurs d'ontologie

- S-expressions (cf. LISP)
- KM, OWL...
- Swoop, Kmgén...



ORO [Lemaignan, 2012]

Signaux sociaux

- Les actions d'autrui sont d'abord des formes spatiotemporelles...

Comment produit et perçoit-on les intentions d'autrui?

Alignement/synchronisation

Signaux sociaux
Verbaux, co-verbaux & non
verbaux

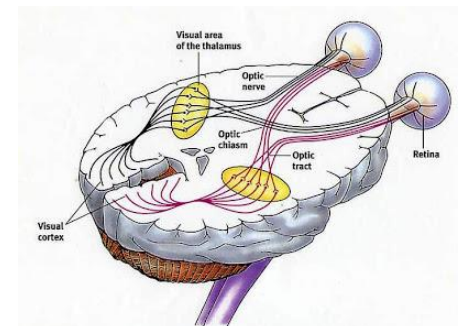
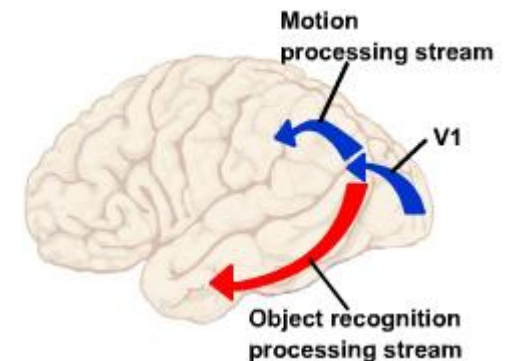
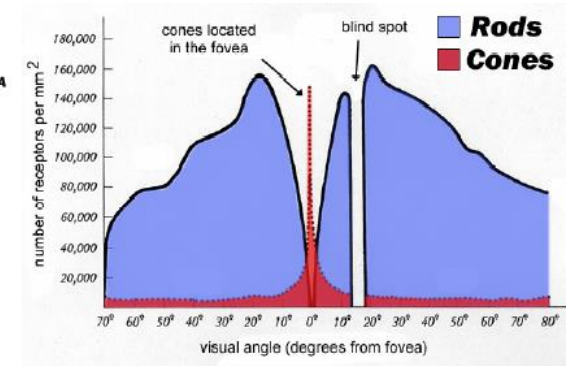
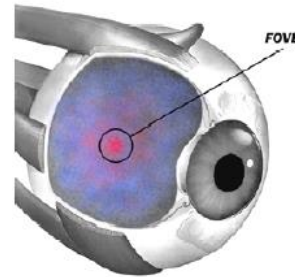
- **Signaux sociaux**
 - Parole: prosodie (dont respiration) mais aussi contenu linguistique/phonologique
 - Regard: tête/yeux, orientation de l'attention
 - Expressions faciales: émotions/attitudes
 - Gestes brachio-manuels: pour soi/l'autre
 - Posture: engagement du corps
- **Alignement/synchronisation**
 - Imitation, mimétisme
 - Lien/ciment social, systèmes couplés



Le regard

Vision (1)

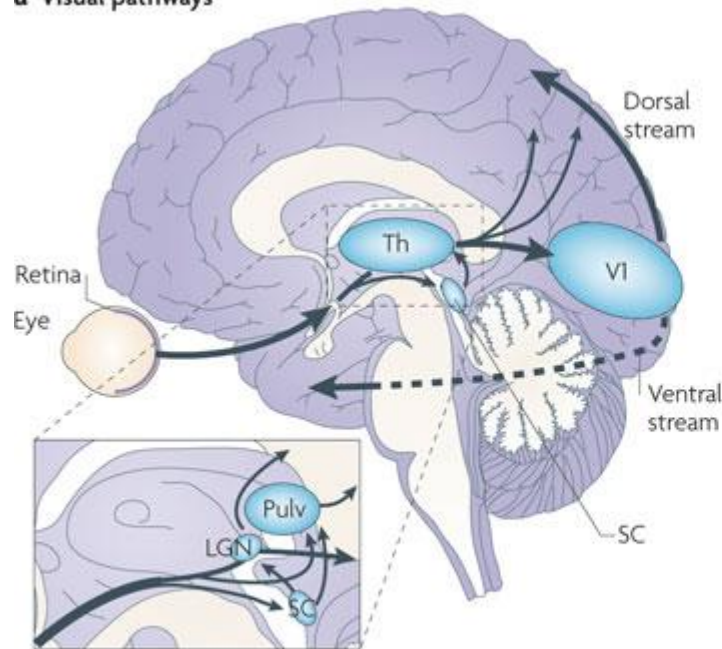
- Bâtonnets & cônes
 - Cônes: analyse colorimétrique
 - Bâtonnets: analyse cinématique
- Deux systèmes visuels
 - Voie ventrale, système rétino-géniculo-strié (RGS) - perception pour la compréhension: fonctions perceptives (apprentissage visuel, discrimination, reconnaissance des objets : formes, couleurs, estimation des distances, etc.)
 - Voie dorsale, système rétino-géniculo-tectal (RGT) - perception pour l'action: détection des déplacements afin de déclencher rapidement les mouvements orientés vers la cible détectée... qualifié par certains de « détecteur de mouche » chez les amphibiens (Barlow, 1953; Mach, 1959). Le RGT est donc impliqué dans le contrôle des habiletés motrices (par exemple : préhension, saisie, attraper etc.)
- Relai par corps géniculé latéral du thalamus; Perception non directe: 20% de l'information vient de la rétine
 - Partie du système somatosensoriel: Intégration vision/audition cf. localisation
 - Forte rétraction de V1; cf. attention



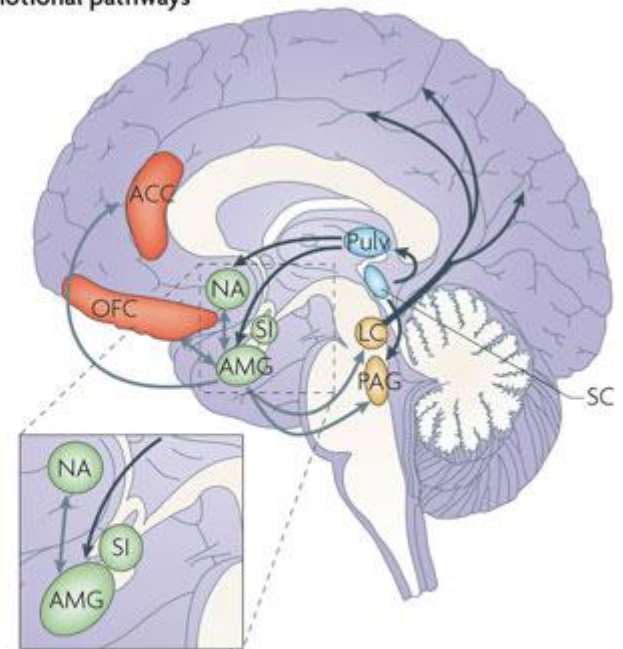
Vision (2)

- Vision & émotion
 - Des modules communs: Colliculus supérieur & Pulvinar
 - Traitement inconscient de stimuli émotionnel (deGelder, 2010)

a Visual pathways



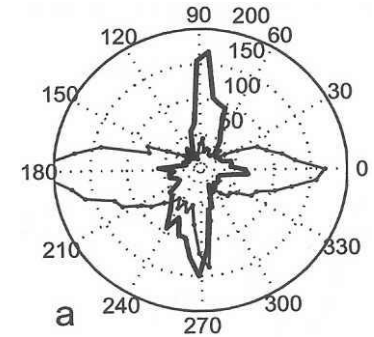
b Emotional pathways



Fixations & Saccades

- Saccades

- Attention: compétition entre exogènes (événements saillants de la scène, apparition abrupte..) et endogènes (buts fixés par l'observateur)
- 150,000 par jour [Abrams, 92], vitesse max: $500^{\circ}/s$, 30 à 120 ms, 1 à 40° ; exploration écran: 7 m/mn
- Coordination binoculaire: vergence

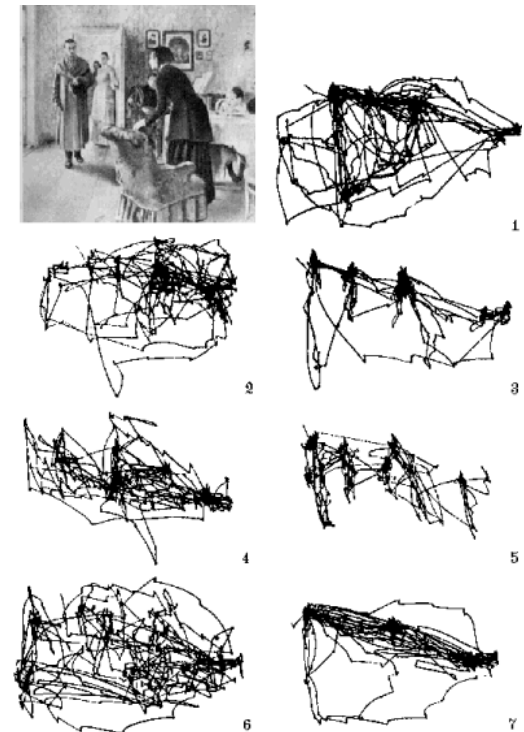


- Fixations

- Analyse de la cible de l'attention amenée dans la fovéa de la rétine
- 200-300 ms, 3/5 par seconde
- Mouvements miniatures: trémolo, dérive et micro-saccades (compensation à l'adaptation rétinienne)

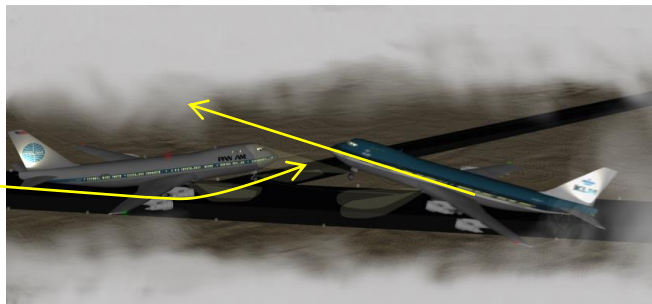
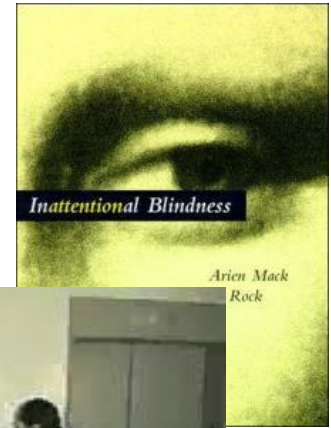
- Saccades exploratoires

- Dépendent de la demande cognitive (Yarbus, 1967: général, environnement, âge, activité, vêtements, positions, durée de l'absence)



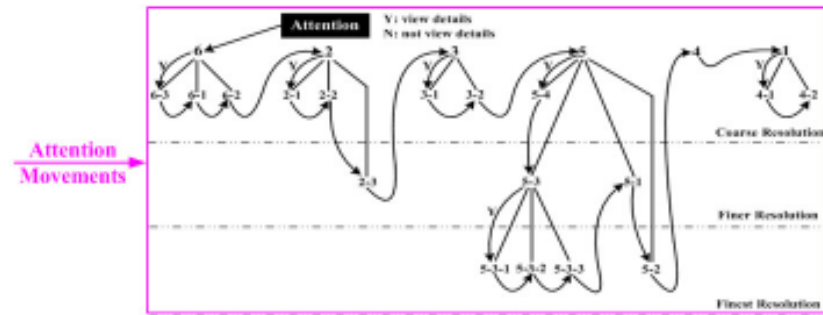
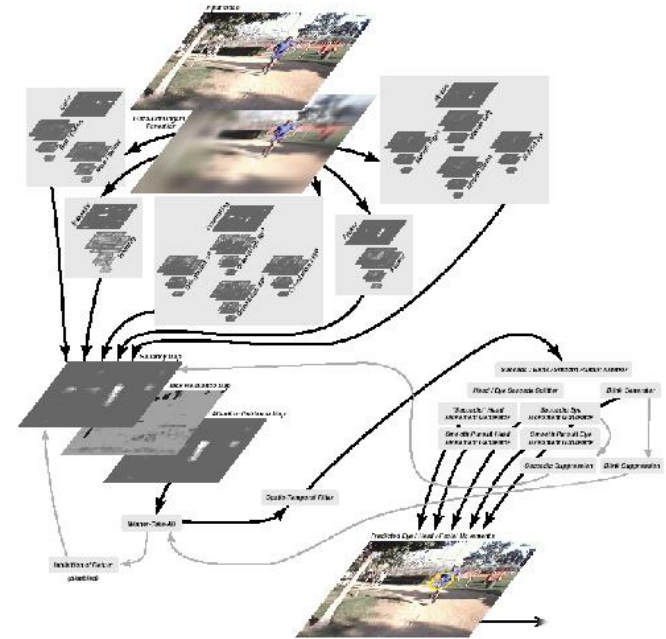
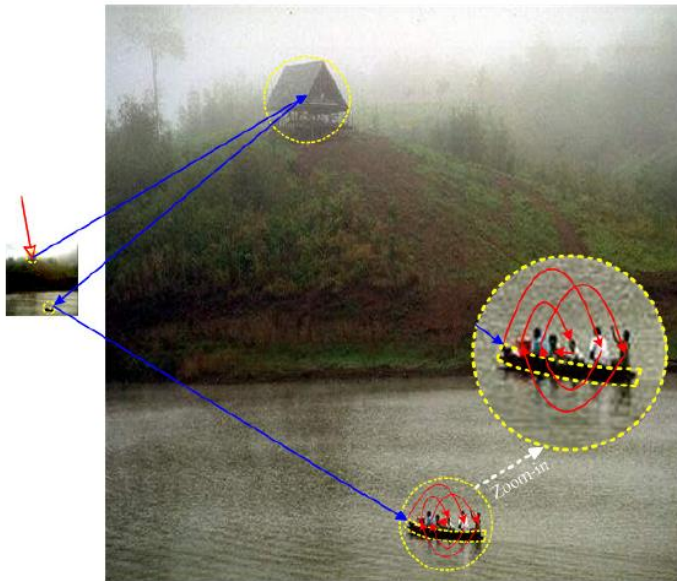
Cécité attentionnelle

- Echec à remarquer un stimulus pourtant parfaitement visible
 - Mack, A. and I. Rock (1998). Inattentional Blindness. MIT Press: 287 pages.
- Le gorille de Chabris
 - Invisible pour 60% des participants... malgré une fixation d'1s en moyenne!
- Plusieurs raisons
 - Saillance
 - Capacité (expertise, fatigue...)
 - Capture cognitive (attention, routine...)
 - Anticipation (attentes, défauts de l'expertise!)
- Bénéfices
 - Ressources cognitives minimales
 - Illusionnistes (bonneteau, ventriloques...)
- Mais graves conséquences...
 - Transport (cf. accident Tenerife 1977)



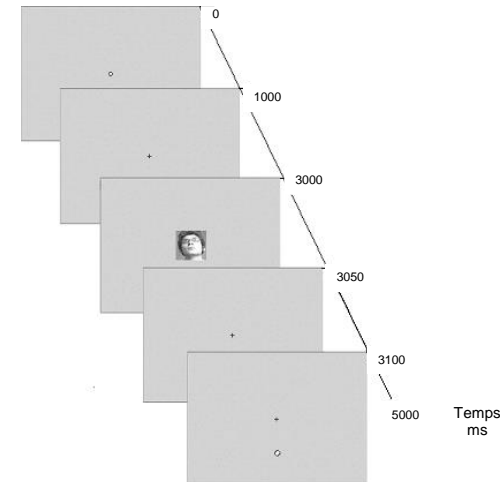
Analyse de scène

- Modèles d'attention visuelle
 - Modèle d'Itti et al [2003]
 - Couplage Perception/action
 - Saillance, pertinence (!), attention
 - Inhibition de retour
 - Analyse syntaxique [Sun, 2003]
 - Where/what [Treisman, 80]
 - Scrutation hiérarchique



Regard d'autrui

- Stimulus exogène particulier
 - Module de base de la Théorie de l'esprit (ToM)
 - Détecteurs d'intentionnalité et de la direction de l'œil / Mécanisme d'attention partagée / Module de théorie de l'esprit (Baron-Cohen, Leslie et al. 1985; Baron-Cohen 1995)
 - Signal d'alerte (me regarde)
 - Objet d'intérêt (champ vision commun)
 - Attire attention (signal tierce)
 - Base de l'apprentissage du monde et du langage
 - Where/what multimodal
- Biais stimulus exogène
 - Langton & Bruce
 - avantage de 20ms si lieux cohérents
 - + important en haut/bas vs. gauche/droite
 - Compétition



Œil coopératif

- Cooperative Eye Hypthesis (CEH: Kobayashi et al; 2002)
 - Evolution de la taille et couleur de la sclère pour faciliter la lecture de la direction de regard par l'autre
 - Width/height et the eye outline (WHR=A/B)
 - Exposed sclera size in the eye outline (SSI=C/D)
 - Testée en perception par Tomasello et al (MPI) sur incongruence tête/oeil: enfants intègrent beaucoup plus la direction du regard par rapport aux grands singes
 - Autres hypothèses: signe de bonne santé (sélection sexuelle), etc
- Impact sur systèmes
 - Voir/faire voir (Bailly et al, 2014): importance du respect des structures/proportions

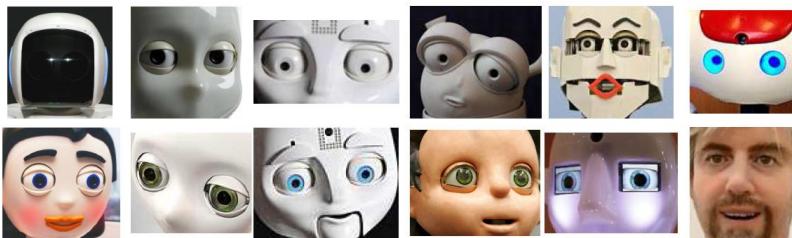
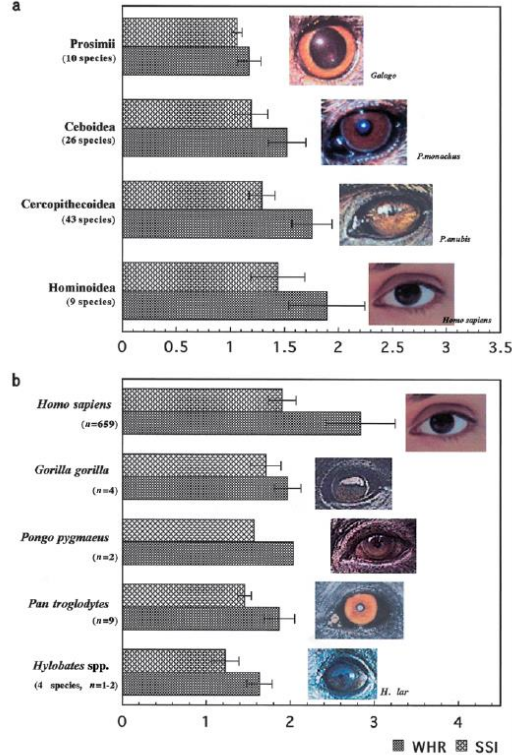
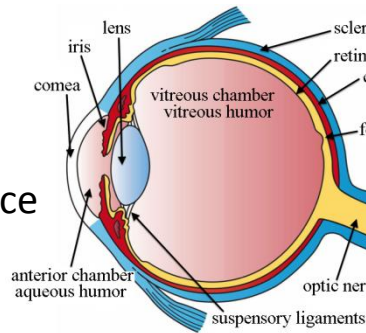
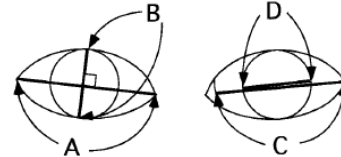


Figure 2: A gallery of robotic eyes. From left to right, top to bottom: hidden gaze of Asimo; original iCub, Nexi, Reeti and Kobian with no iris, Nao with no sclera; Flobby, iCub, Nexi and Zeno with sclera, iris and pupil; the Robothespian eye graphics vs. the hyper-realistic eyes of Geminoid DK. The panel of anatomies and pupil/iris/sclera proportions is very diverse.

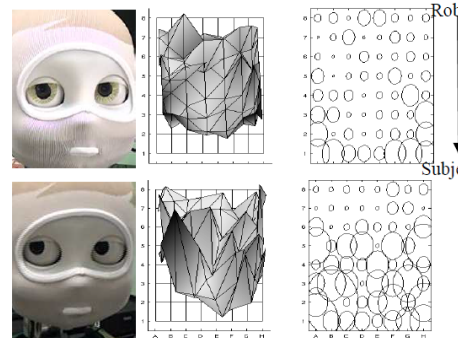


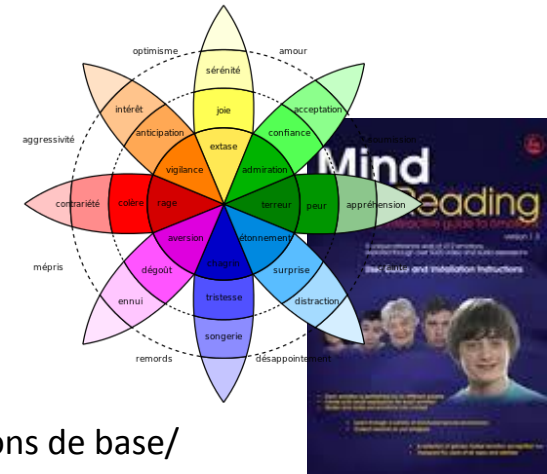
Figure 2. Variation of WHR (width/height ratio of the eye outline) and SSI (index of exposed sclera size in the eye outline) among the phylogenetic groups of primates (a), and in Hominoida (b).

Les expressions faciales

Taxonomies

- Emotions primaires

- Les « big six »: admiration, amour, haine, désir, joie & tristesse (Descartes, 1649) vs. joie, tristesse, peur, surprise, colère & dégoût (Ekman et al., 1969)... génératives
 - innées, universelles et communicatives (Darwin, 1872)
- Autres organisations
 - 2D (cf. Plutchik, 1980), 3D (modèle VAD, Osgood, 1992)
 - Hiérarchique: Mindreading (Baron-Cohen, 1995-2004): 24 émotions de base/ 417 émotions secondaires
 - Développement (Durand et al, 2007). Comment et quand...
 - 5-6 ans : reconnaissance de la joie et de la tristesse
 - 7 ans : reconnaissance de la peur,
 - 9 ans : reconnaissance de la colère
 - 13 ans : reconnaissance du dégoût



- Evaluation cognitive (*appraisal*)

- Evaluation en 4 dimensions (K. Scherer, 2001)
 - Est-ce que cet évènement affecte directement ma personne ou mon groupe social ? (pertinence)
 - A quel point les conséquences de cet évènement vont-elles affecter mon bien-être ou mes buts à court et long terme ? (implications)
 - À quel point suis-je capable de faire face à ces conséquences ? (potentiel de coping)
 - Quelle significativité a cet évènement par rapport à mes convictions personnelles ainsi que face aux normes et valeurs sociales ? (significativité normative)
- Réévaluation possible de l'évènement (Lazarus et Folkman, 1984)

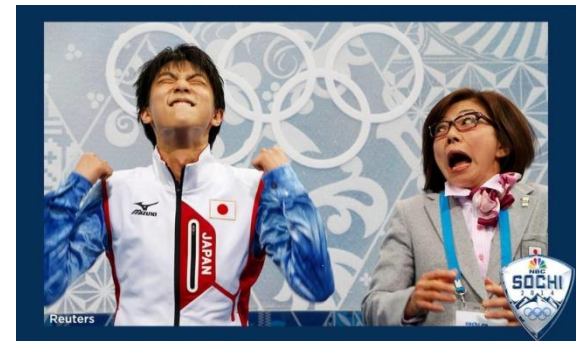
Dynamique

- Micro-expressions

- Expression faciale brève et involontaire que le visage humain exprime en fonction des émotions vécues. Elles apparaissent généralement lors de situations où les enjeux qui en découlent sont élevés
- Peuvent se produire en un temps très court, de l'ordre d'1/15 à 1/25^e de seconde
- Analyse du langage corporel (synergologie). D. Matsumoto (San Francisco State University & Humintell)
 - Production innée (judoka aveugles)
 - Perception apprise

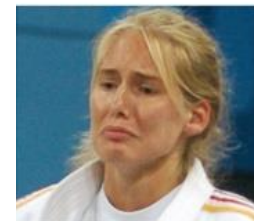
- Culture populaire

- Lie to Me, the Mentalist...



Blind athlete

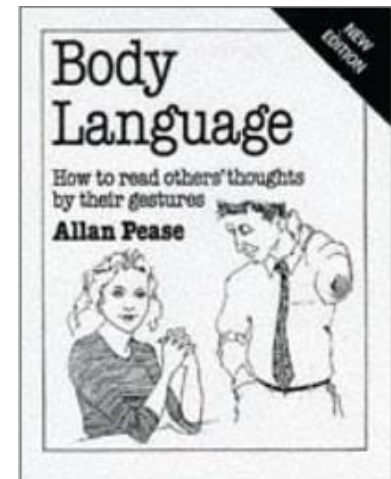
Sighted athlete



Les gestes braccio-manuels

Typologie des gestes

- Gestes communicatifs
 - Utilisés par animaux
 - Chimpanzés (indication d'épouillage)
 - Corbeaux (attirer attention sur objet [Pika & Bugnyar, 2011])
- Catégories: déictiques, iconiques, adaptatifs
 - Déictiques (associés ou non à mots spatiographiques: ici, là, démonstrateurs...)
 - Panto/pictomimiques (action/forme/qualité du référent...)
 - Emblèmes (ok, au revoir...)
 - Synchronisateurs de l'action (contacts, régulateurs, hochements, ...)
- Extracommunicatifs
 - Autocentrés (grattage, tapotement, bâillement), ludiques (manipulation d'objets), confort (croisement de bras, posture...)
- Echelle convention
 - Gesticulation > Gestuelle > Pantomime > Emblème > Langue des signes
 - Geste pour soi/l'autre (Golden-Meadow)



Typologie des gestes

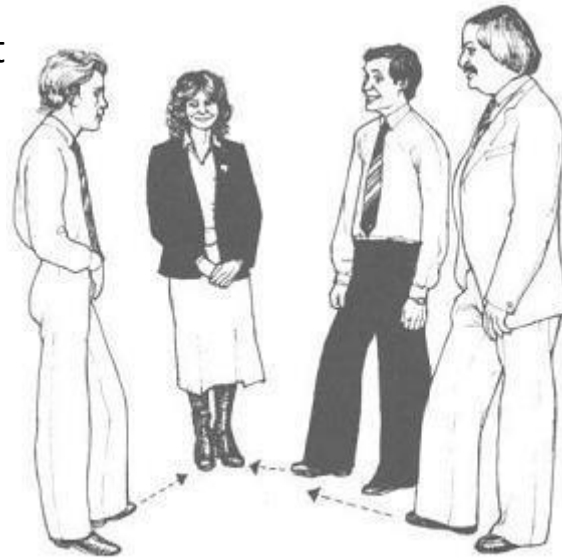
- Gestes exocentrés
 - Attirer l'attention & communiquer (pull)
 - Agents
- Gestes endocentrés
 - Signale état mental non intentionnellement (push)
 - Confusion, concentration, angoisse [Rozin & Cohen, 03]
 - Vers la face
 - Utilisés aussi en HCI/HRI. Mahmoud et al (2011): HCI (16%) vs. interactions dyadiques (25%)
- Métaphores temporelles & spatiales
 - Distance au corps
 - Avant/arrière: passé devant et futur derrière chez les Aymara, nation indigène des Andes [Nunez & Sweetser, 2006]



Posture

Relation à l'interlocuteur

- Langage du corps
 - Proxémique et entrée du territoire personnel
 - Torse vs regard: gestion du référent vs. affilié [Andrist, 2012]
 - Footing [Goffman, 1963]: qui participe à la conversation et à quelle hauteur
 - Tour de parole
 - Orientation des segment interlocuteurs
 - Torse
 - Pieds
- Ouverture/fermeture
 - Exposition du corps



Feet signalling what's on the mind



Referential Gaze



Gaze toward participant



Gaze toward map

Affiliative Gaze



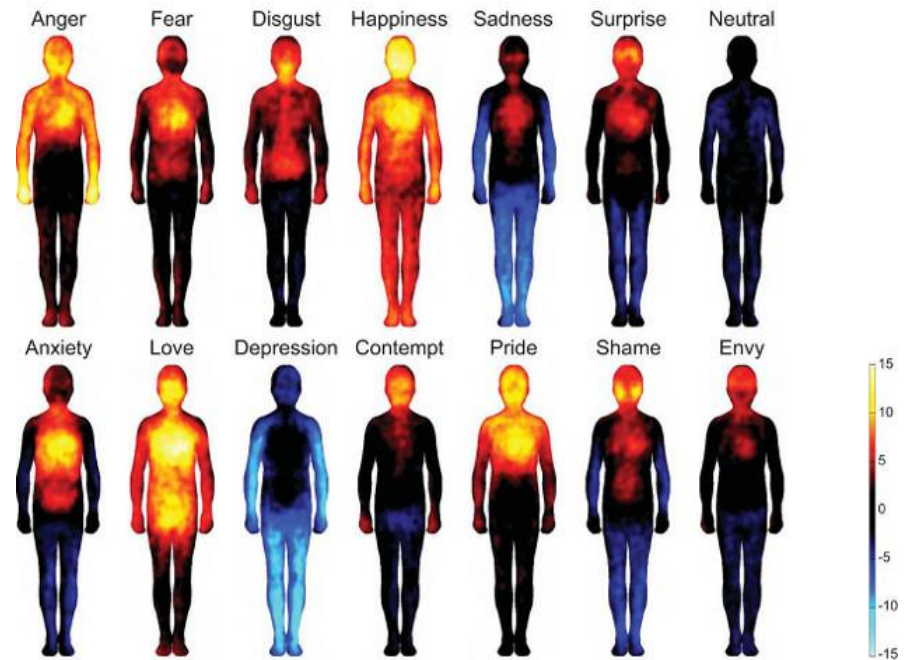
Gaze toward participant



Gaze toward map

Emotion et corps

- Ressenti topographique
 - Nummenmaa et al (2014) enquête sur 700 participants en Finlande,, Suède et Taiwan
 - Biomarqueurs émotionnels
 - Changements physiologiques torse (respiration & rythme cardiaque), ventre (système digestif) & face (action musculature faciale, température peau, glandes lacrymales...)

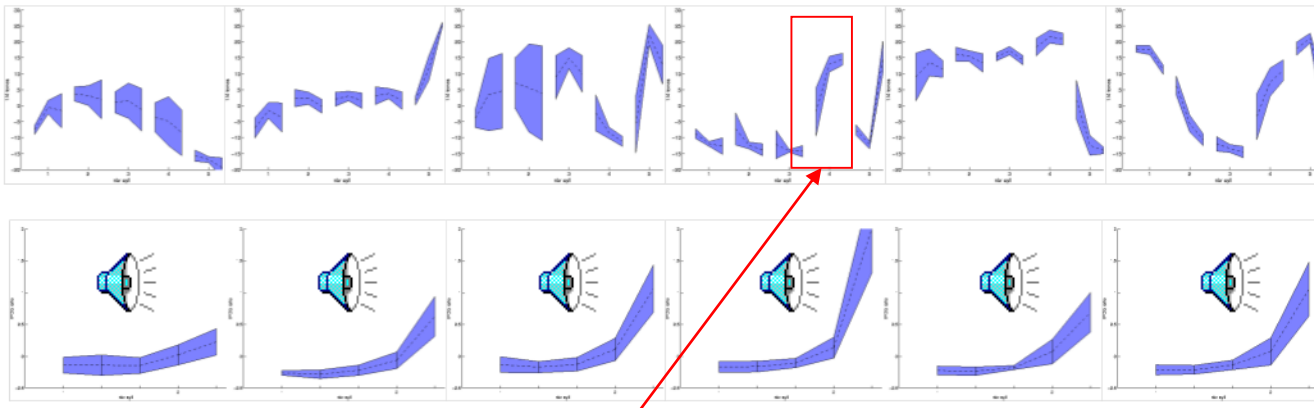


Contenu verbal

Et la parole

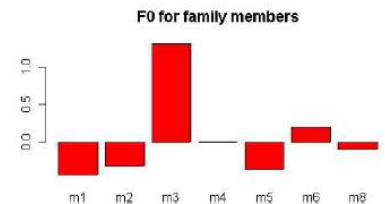
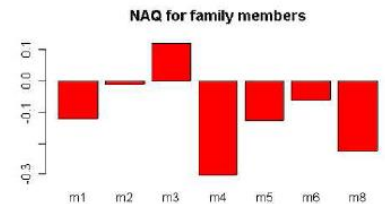
Intonation

- Fonctions attitudinales & émotionnelles
 - Intonation and its uses (Bolinger, 1989): "Intonation [is] a non arbitrary, sound-symbolic system with intimate ties to facial expression and bodily gesture, and conveying, underneath it all, emotions and attitudes... [it is] primarily a symptom of how we feel about what we say, or how we feel when we say".
- Modalité, affect, attitude
- Modèle Gestaltiste
 - Contours spécifiques: déclaration, question, exclamation, incrédulité, ironie de soupçon, évidence...



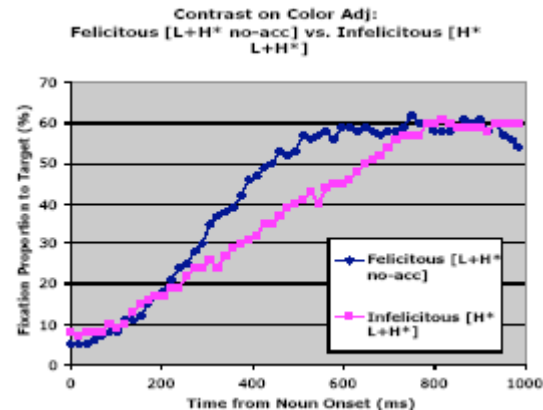
Forme fait sens: événement saillant intégré dans le contour global

- Bien d'autres paramètres: timbre (ex: douceur de voix), etc



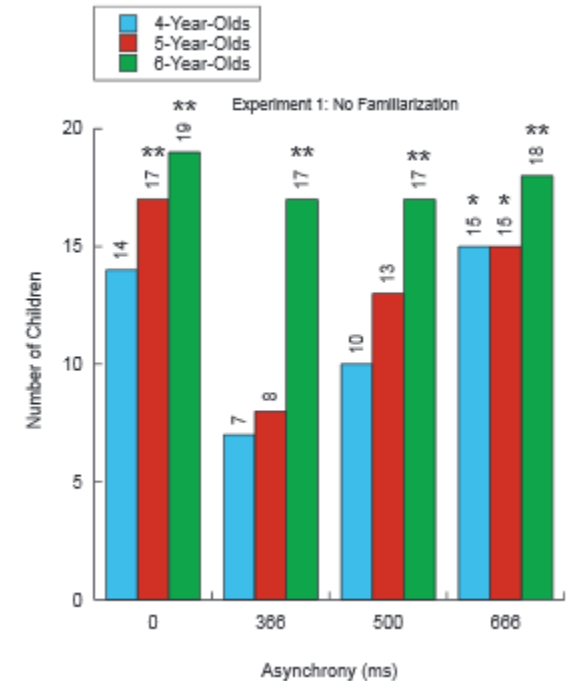
Deixis prosodique & information

- Séquences d'actions
- Gestion du rhème/thème
 - L+H* correct (e.g. First, hang the green drum. Next, hang the ORANGE drum.)
 - L+H* incorrect (e.g. Next, hang the orange DRUM)
 - Bénéfice de 200ms pour emphase correcte
 - Ito & Speer (2005) Ohio U.



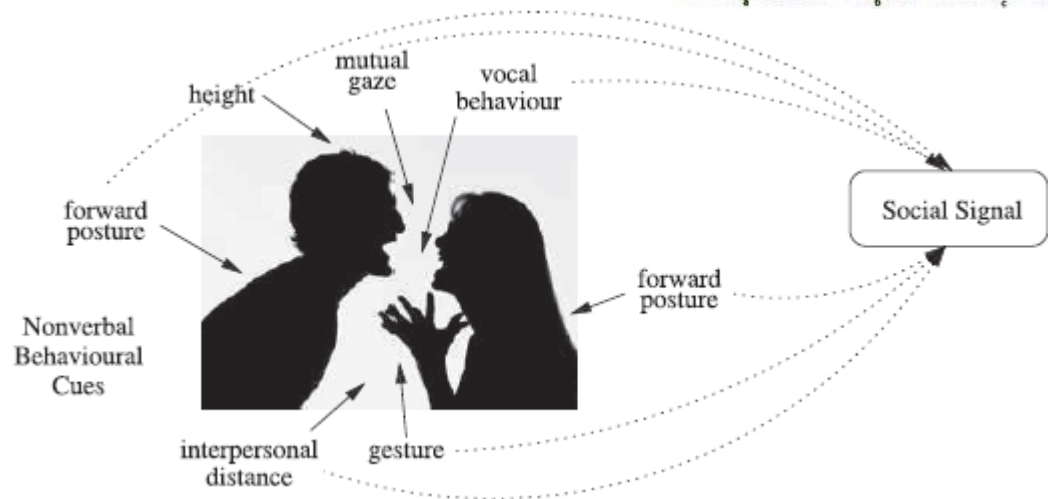
Multimodalité

- Redondance
 - Corrélations positives
 - Emotions dans la voix et sur la face (ex: (sou)rire, surprise...)
 - Mélodie vs sourcils & mvts de tête...
 - Pointage vs. regard
- Complémentarité
 - Mouvements faciaux, gestes et regard
 - Verbal et co-verbal
 - Parole vs. gestes brachio-manuels
 - Ex qualitatif & quantitatif: taille, poids, etc
 - Acquisition de la sémantique lexicale
 - Compensation chez les enfants dysphasiques
- Intégration
 - Rétrécissement de la fenêtre d'intégration [Lewkowicz 2012]



Traitement des signaux sociaux

- A. Pentland (MIT, Medialab): speed dating
 - Préd. succès d'une rencontre sur la base des signaux échangés
 - HMM couplés
- A. Vinciarelli (IDIAP & U. Glasgow): multimodal SS
 - Reconnaissance des relations sociales: dominance, rôle...
- Gestion de la relation interpersonnelle
 - Théorie de l'adaptation (CAT): imitation/mimétisme
 - Distance: proxémie
- Contagion
 - Rire



Adaptation à l'autre/envt

- Décodage/encodage des intentions
 - Boucle perception/action
 - Représentations sensori-motrices
- Influences culturelles et sociales
 - Comparaison interculturelles (ex. sarcasme Australie vs. Iran [Shahaeian 2014], politesse au Japon, etc)
- Apprentissage
 - Démonstration, observation
 - Développement, curiosité
- Alignement des comportements et des représentations
 - Lien/ciment social (social glue)
 - Contagion
- Cf CM4 Adaptation mutuelle



Conférences, challenges

- Journaux
 - *Journal of Personality and Social Psychology*
 - *International Journal of Human-Computer Studies*
 - *Journal on Multimodal User Interfaces*
 - *Journal of Nonverbal Behavior*,
 - *IEEE Signal Processing Magazine*
 - *IEEE Transactions on Affective Computing*,
 - *IEEE Transactions on Multimedia*
- Conférences
 - *Intelligent Virtual Agents (IVA)*
 - *International Conference on Affective Computing and Intelligent Interaction (ACII)*
 - *Human Behavior Understanding (HBU)*
 - *ACM International Conference on Multimodal Interaction (ICMI)*
- Challenges
 - [Mapping Personality Traits Challenge & Workshop \(MAPTRAITS\)](#)
 - *Computational Paralinguistics Challenge (ComParE), Interspeech*