



Interagir grâce à l'informatique émotionnelle

CNRS – ISIR – Sorbonne Université

catherine.pelachaud@isir.upmc.fr



Embodied Conversational Agents : an introduction

- Affective computing (Picard, 1997):
 - recognize emotions
 - express emotions (facial expressions, speech...)
 - « have » emotions
- Artificial Intelligence
 - reason, decide
 - has mental state
 - Build mental state of others

Embodied Conversational Agents: an introduction

Create virtual characters that can:

- simulate cognitive and expressive human capabilities
- communicate using verbal and nonverbal means
- display a wide range of socio-emotional behaviors
- be socially aware and emotionally competent
- be capable of holding multi-modal social interactions

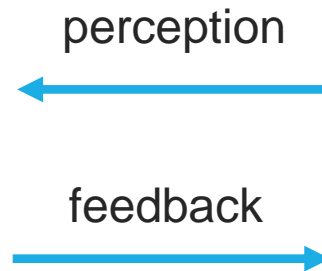
Embodied Conversational Agents: an introduction

General overview

- cognitive processes
- reactive processes



ECA's decision



Embodied Conversational Agents: an introduction

General overview : example

- cognitive processes:
« *human shows me
her friendship* »
- reactive processes :
mirror smile



ECA's decision:
smile



ECA perceives
human is smiling



ECA smiles to
human



Embodied Conversational Agents: Applications

Serious games:

- Training: job interview, intercultural communication
- Social ability: bullying, shyness



 TARDIS

Sabouret et al



Embodied Conversational Agents: Applications

Health & wellbeing

- Depression
- Post-traumatic stress disorder (PTSD)
- Autism
- Addiction (tobacco, alcohol)



Pierre Philip et al

Embodied Conversational Agents: Applications

Education: virtual tutors, peers,
mentors

Companions

Elderly care



photo credit: Justine Cassell

Communication

Definition of communication (Allwood, 06)

- not simply a transfer function
- sharing anything between arbitrary entities where all entities are active, interacting with each other and within a social and interrelational context

Involve different processes: sense, perceive, adapt, generate, plan

Non-Verbal Behaviors

Multimodal, dynamic, transient

- Facial expression
- Gesture
- Gaze
- Posture
- Head movement
- Voice
- Touch
- ...

Functions of Non-Verbal Behaviors

Various communicative functions: convey information about

- Semantic content, intention
- Conversation management, turn-taking
- Relationships, attitude, emotion
- Engagement, involvement
- ...

Smiles

Important social signals

Bear many meanings: more than 50 functions (Ekman; Hess)

- Positive/negative emotion (joy, embarrassment)
- Affiliation (social bonding, rapport)
- Attitude (dominance, irony)
- Greeting
- Politeness
- ...
- Not all related to emotions

Differences in expression:

- Change in facial and body cues
- Symmetry
- Dynamic

Embodied Conversational Agents where we started

Cassell, Pelachaud, Badler,
Steedman... Suggraph'94

Autonomous agents designed to be:

- Speaker

- Listener

With multimodal behaviors:

- Gesture

- Facial expression

- Gaze

- Lip movement



Gilbert and George at the Bank (Upenn)

Cassell, Pelachaud, Badler, Steedman... Suggraph'94



Socio-emotional agent

Non-verbal behavior model:

- display great variety of expressive and emotional behaviors
- go beyond the six basic expressions of emotions
- communicative behaviors

Methodology:

- theories from psychology, linguistic...
- data annotation and analysis
- data-driven

Corpora

NoXi (Cafaro et al)



Tardis (Chollet et al)



Ilhaire (Ding et al)



Emilya (Fourati et al)



Synthesized corpora

Toolbox

- Tools to create audiovisual corpus of virtual agent's non-verbal behaviors
- The framework to develop and distribute the tool online

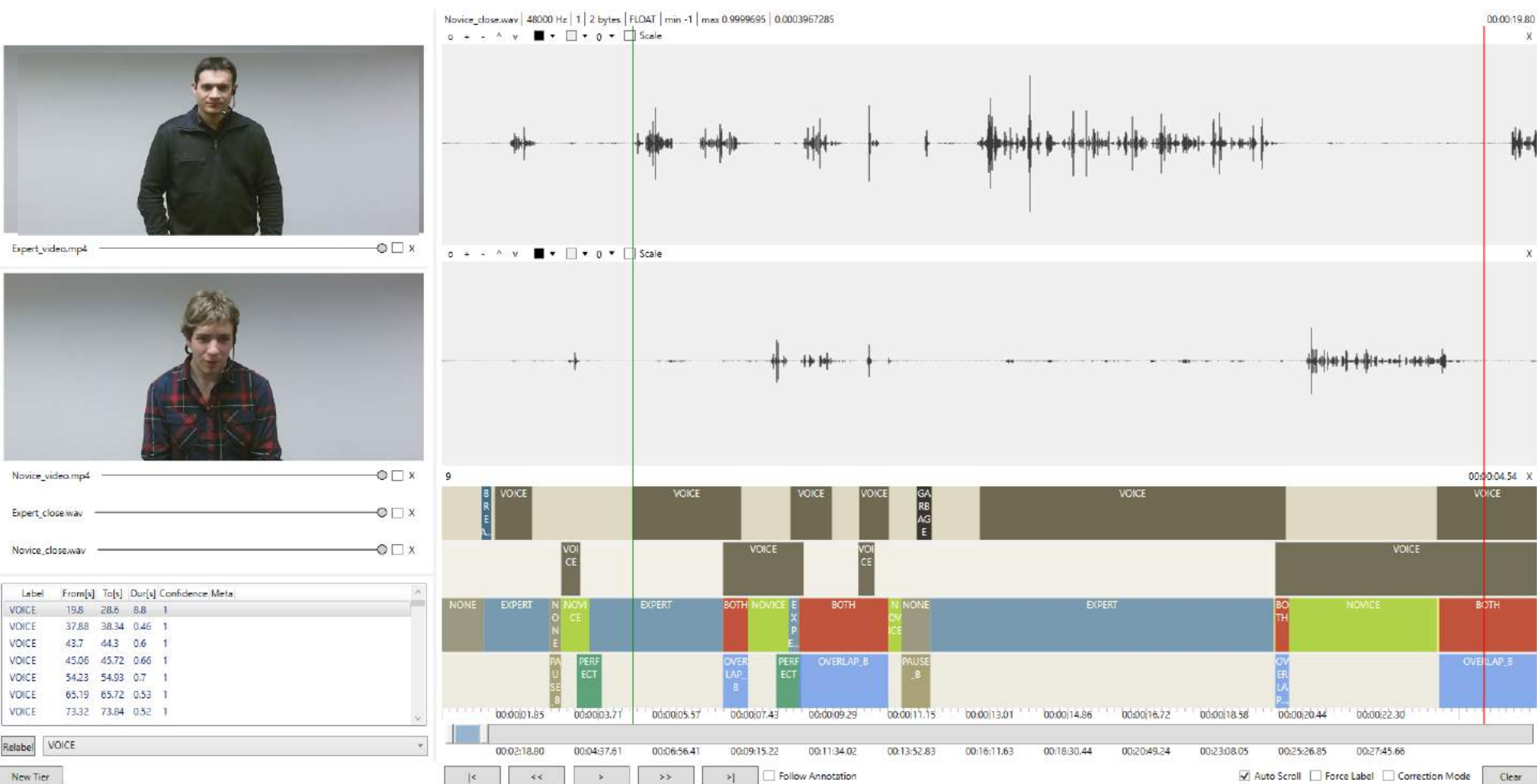
E-Smiles Creator (Ochs et al)



GenAttitude (Ravenet et al)



Corpus Annotation: Automatic/Manual

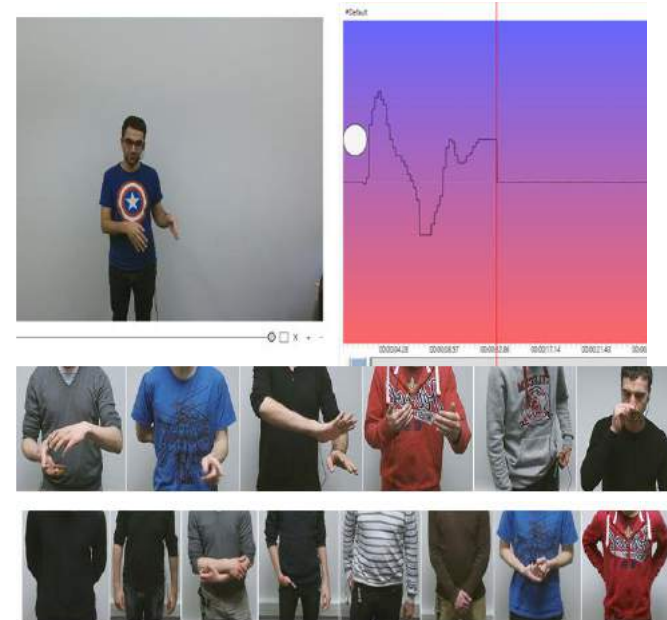
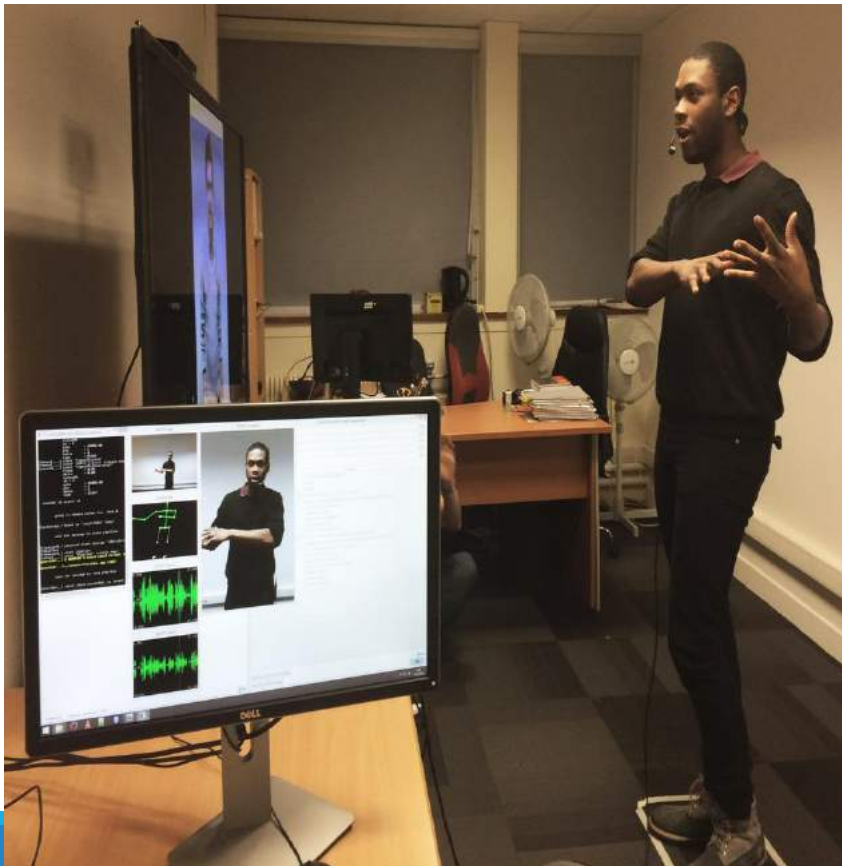


Corpus Analysis

What are the NVB cues?

Which social attitudes are displayed

What is the level of engagement?



(Biancardi et al)

Modeling the Expression of Interpersonal Attitudes

What makes a person appear more dominant or less friendly?

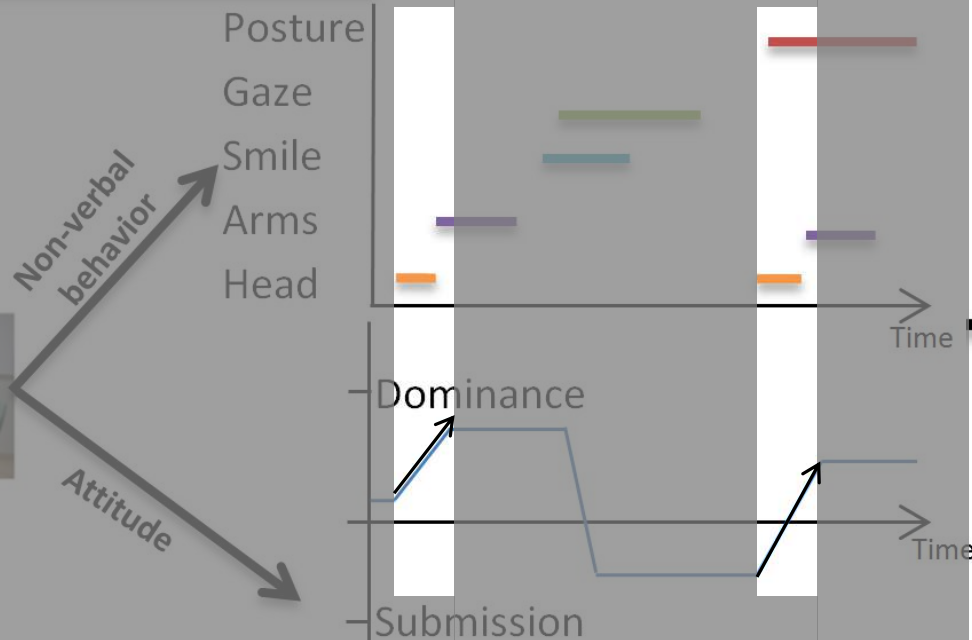
Our aim

1. Study the variation of social attitudes
2. Find out the sequences of multimodal behaviors that trigger this change in perception

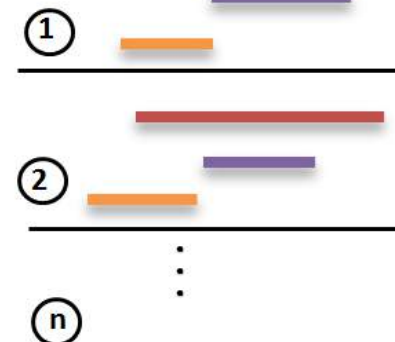


Extraction of sequence of non-verbal signals expressing **attitude variation** (increase or decrease): **temporal sequence mining**

Step 1 Annotation



Sequence Dataset for Dominance Increase



Step 2 Temporal Sequence Mining

Order

1

2

Arms crossed

Head shake

Duration

Staring time

Time



Step 3 Patterns Integration in ECA

Temporal Patterns for Dominance Increase

Example (Dermouche et al)



Dominance Increase



Dominance Decrease

Speech driven approach

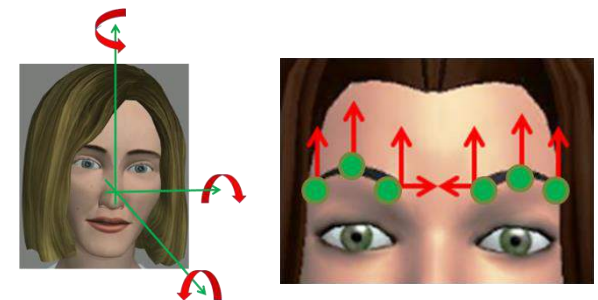
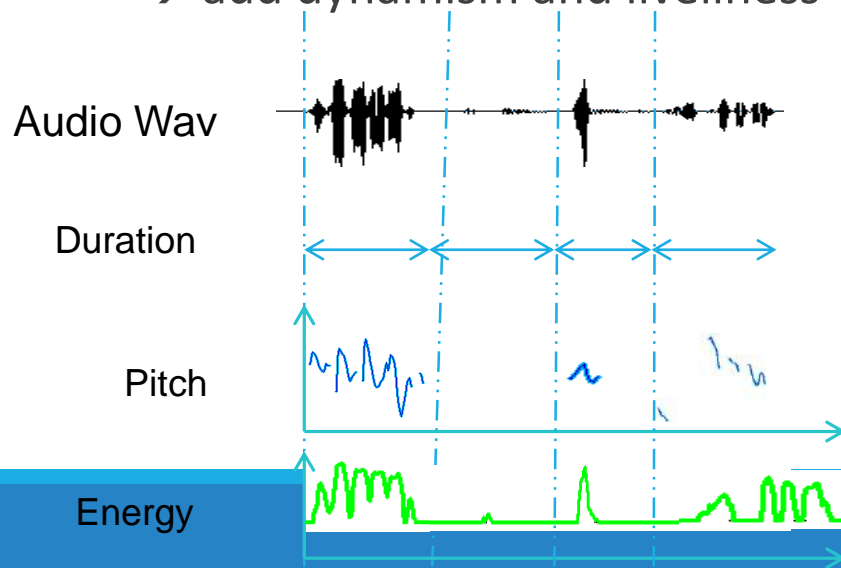
Input: speech features

Output: 3D animation parameters

→ infer lip movement, facial expression, head movements, gesture from audio cues

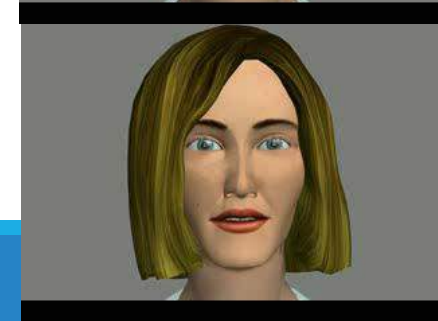
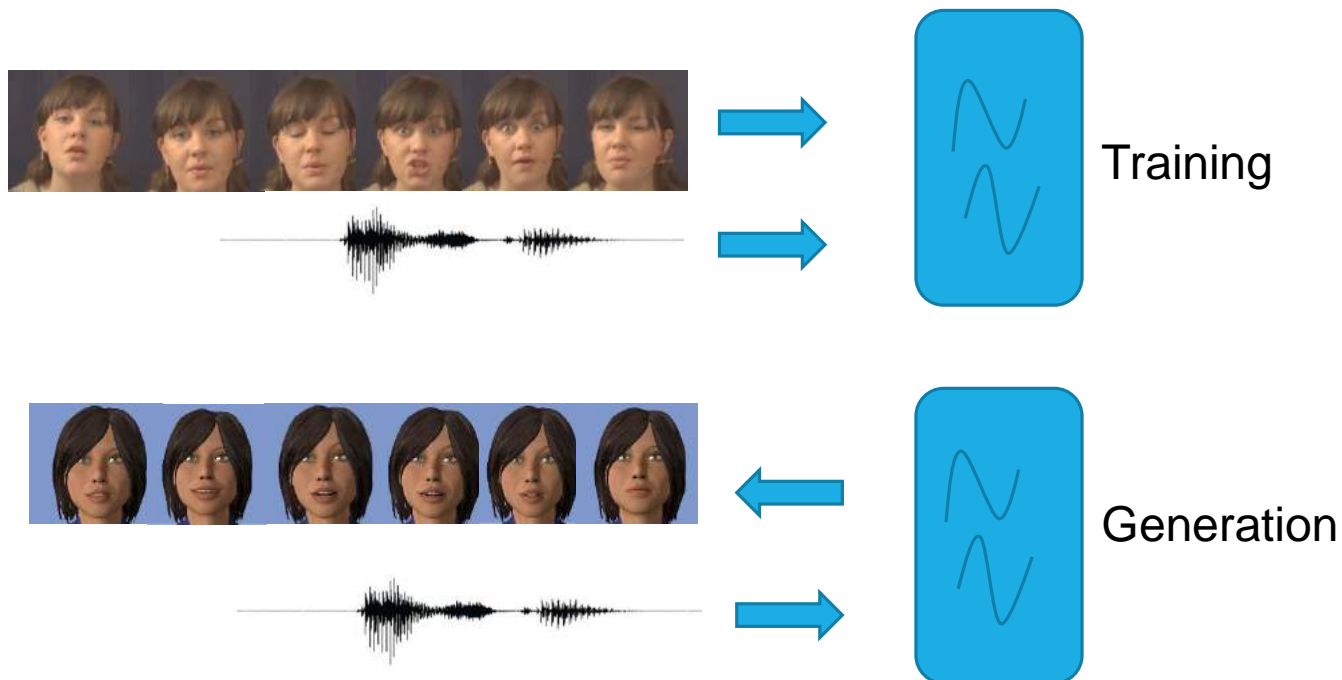
→ capture link between features

→ add dynamism and liveliness



Statistical Models

Contextual HMM: capture temporal relationship between speech data and multimodal movements taking into account context



Interaction

Interaction is by no means a one way communication channel between parties.

Speaker and listener adapt their behaviors to each other

- Speaker monitors addressees attention and interest in what she has to say
- Listener selects feedback behaviors to show the speaker that he is paying attention, agreeing, understanding, etc

Tight dynamic coupling between both interactants

What we have achieved

Real-time interaction:

- perceive and interpret multimodal signals
- build rapport and show engagement
- develop adaptation mechanisms
- impression management

Our aims

During conversations listeners display *backchannel signals* to [Heylen '07]

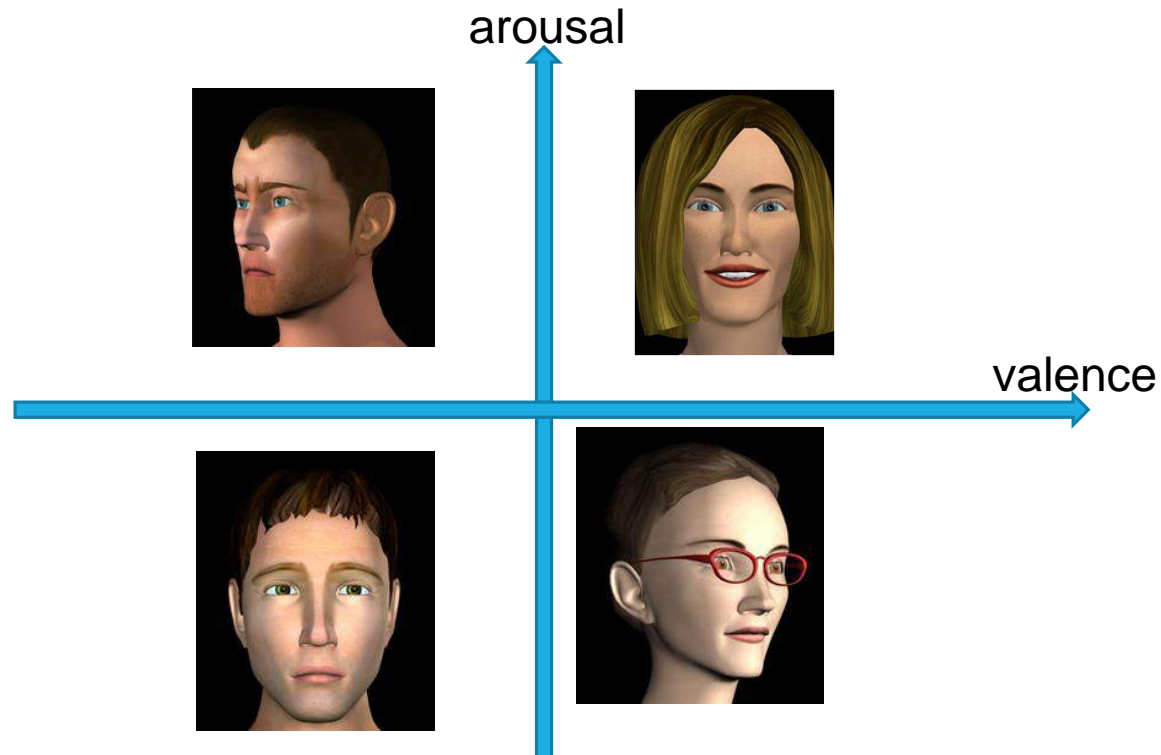
- put the interaction forward
- give information on perception and understanding of speaker's signals
- communicate one's attitudinal reactions

Build a real-time system

- ◆ Integrate components in a real-time architecture
 - audio and video analysis
 - low-level feature extractors
 - classifiers for epistemic-affective states
 - ASR
 - action planning
 - turn-taking
 - generate natural language utterances
 - verbal and non-verbal backchannels
 - continuous face/head expressivity – may involve mimicry
 - system behaviour
 - head and face gestures, TTS

Characters with Emotional Traits

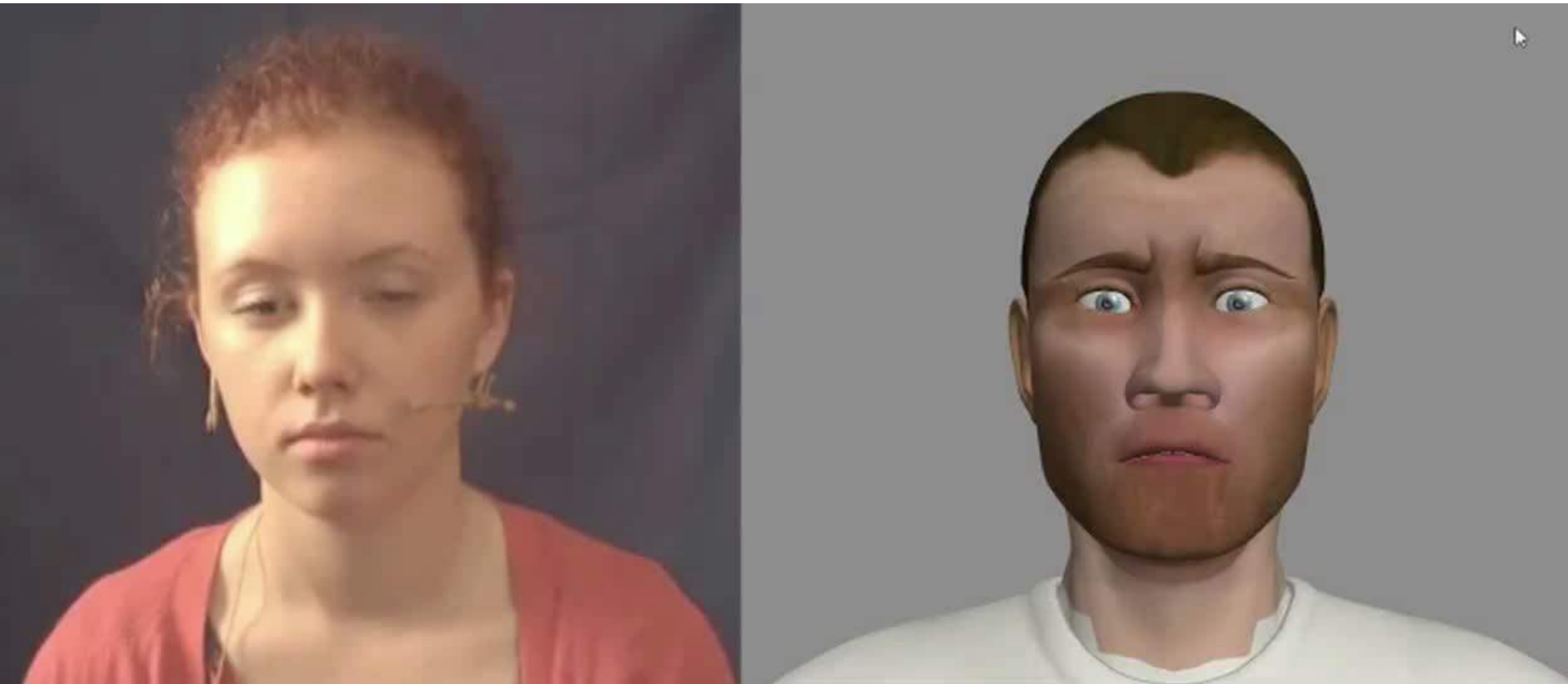
(Bevacqua et al)



Obadiah

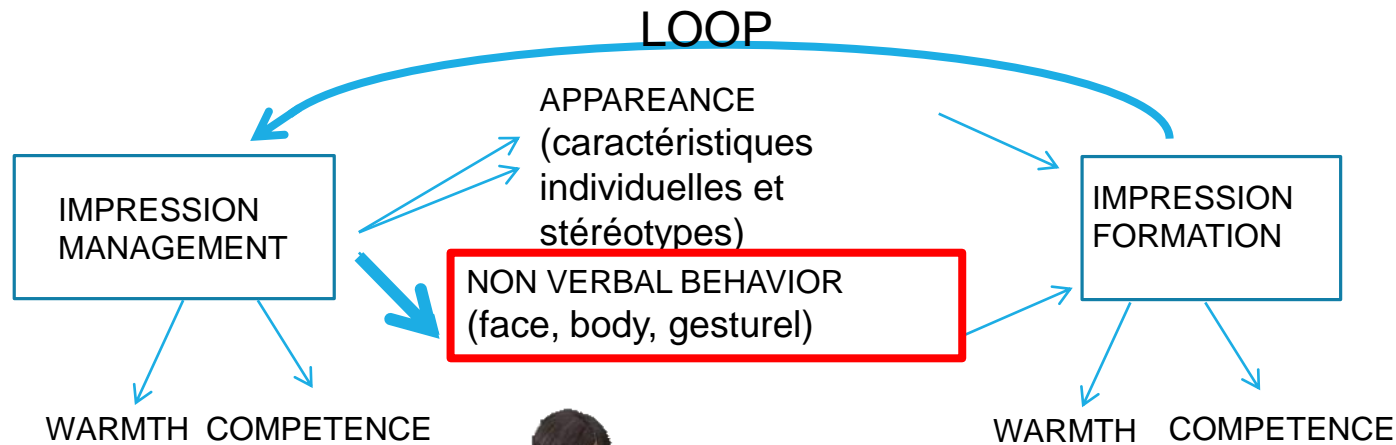


EXAMPLE - SPIKE

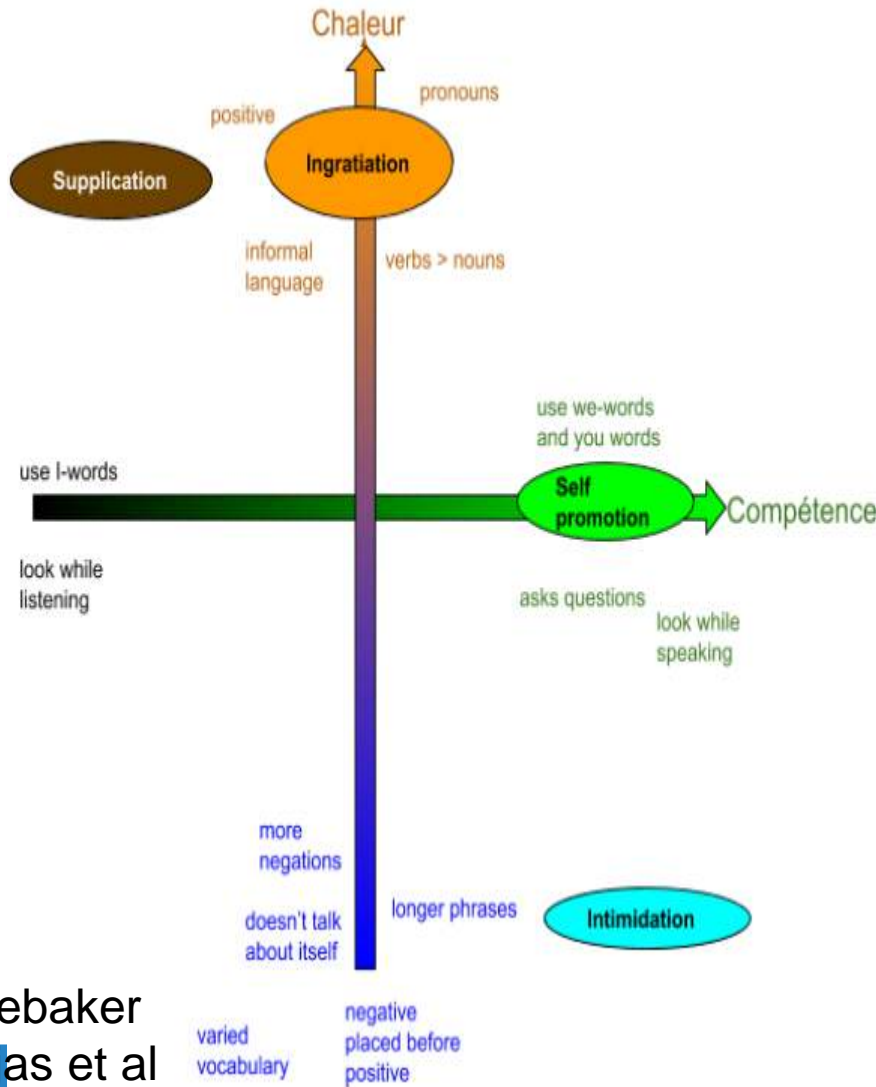


Impression Management

(Biancardi et al)



Impression Management of Warmth and Competence (Biancardi et al)



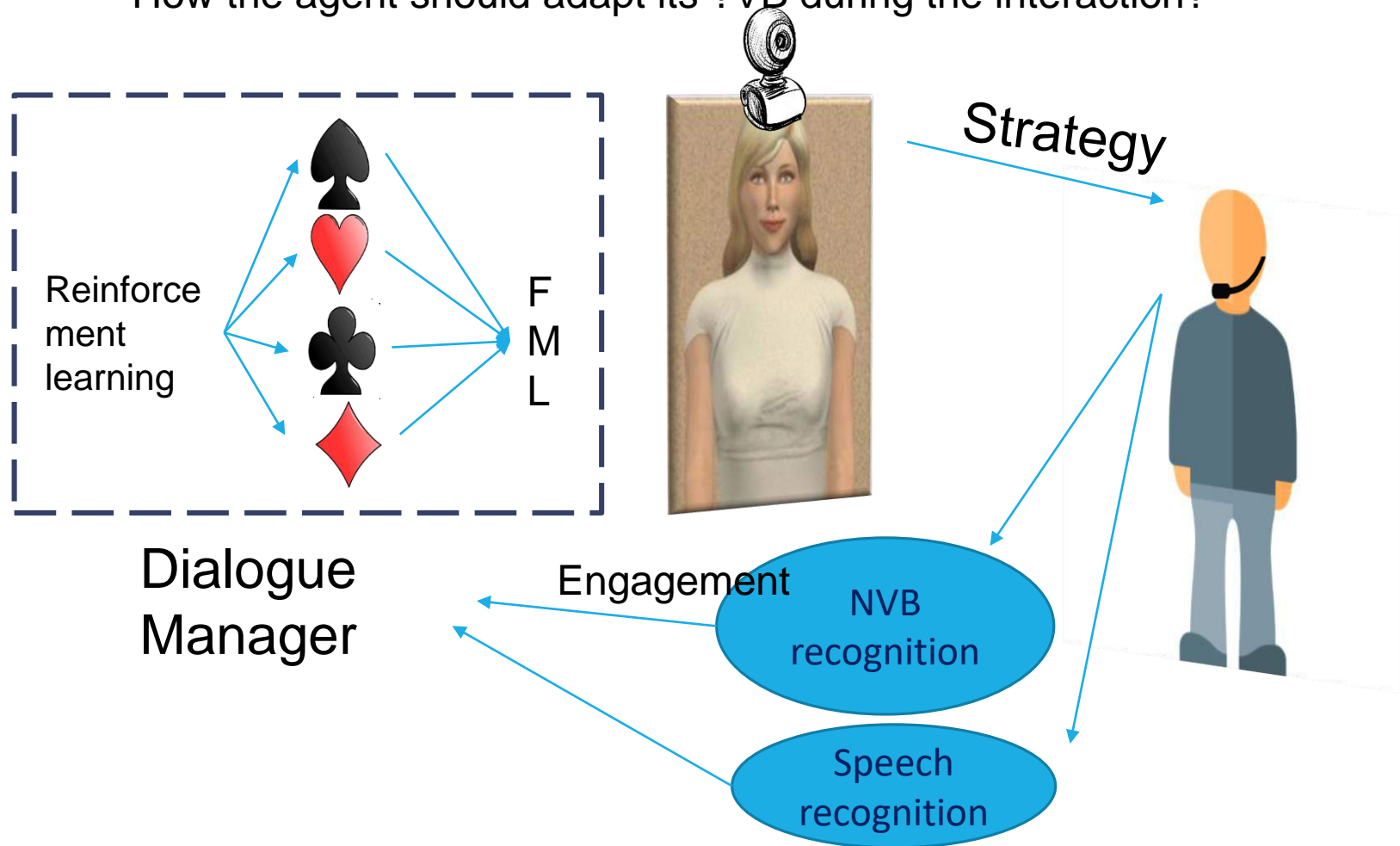
Strategies :

- ♠ Ingratiation **Ch +; Co N/A**
- ♥ Supplication **Co ↘ ; Ch ↗**
- ♣ Self-promotion **Co +; Ch N/A**
- ♦ Intimidation **Ch ↘ ; Co ↗**

Jones & Pittman, 82

Impression Management of Warmth and Competence (Biancardi et al)

How the agent should adapt its ?VB during the interaction?

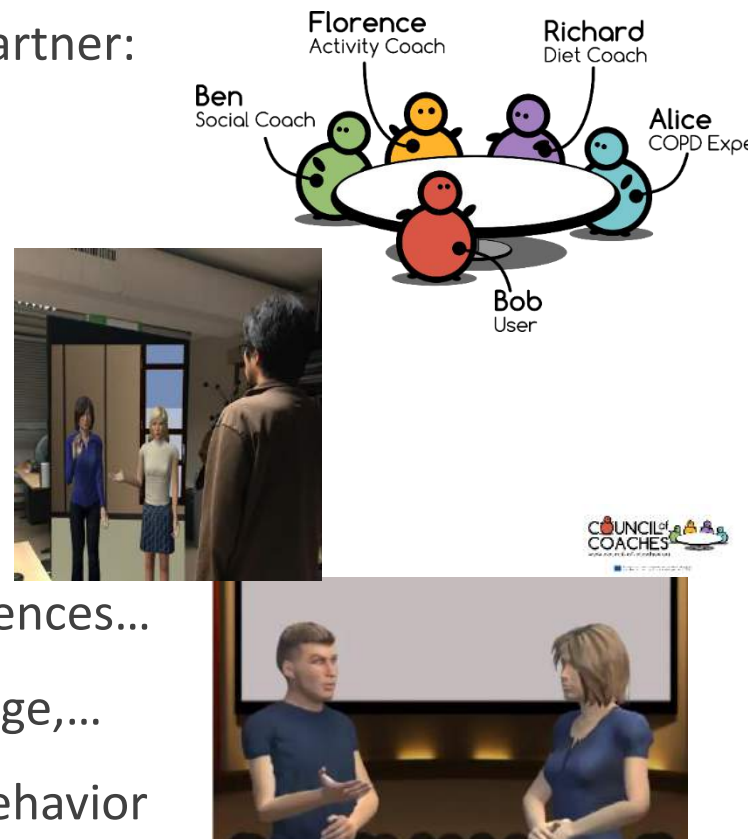


What is missing?

A lot before being a great social interactive partner:

- Conduct multi-party interaction
- Talk and act together
- Manage conversation flow
- Maintain long-term interaction
- Perceive context
- Adapt to users, in term of alignment, preferences...
- All agents follows a generic culture, genre, age,...
- Agents do not have identity, no individual behavior

Ownership of interaction outcome



Any questions?
