Sensitive Artificial Listeners make the user talk natural because today's machines lack the "soft skills" needed to maintain a conversation.

**Research question:**
What does it take for a dialogue with an artificial agent to "feel" like a real conversation?

**Method:**
1) Study human-to-human conversations
2) Implement key non-verbal capabilities in a real-time multimodal interactive system
3) Test which system configurations "work"

**Scenario:**
Sensitive Artificial Listeners make the user talk and, through their reactions, try to induce a certain emotion in the user. They have good non-verbal but very limited verbal skills.

**Data collection & annotation**

- How do humans maintain a conversation?
  - very careful high-quality recording setup
  - multiple cameras and microphones recorded in sync

**Solid SAL**

- human-to-human conversation in the SAL scenario (human operator pretends to be SAL)
  - covers both speaker and listener behaviour
  - 20 subjects, 15:20 hours
  - rich annotations of emotion, epistemic states, interaction
  - available for research from semaine-db.eu

**Wizard-of-Oz SAL**

- operator selects SAL agent responses
- data closer to fully automatic system, including failures
- recordings ongoing

**System building**

**SEMAINE system integrates components for**

- detecting emotions, interest, gender etc. from the voice
- detecting faces, nods/shakes, emotions from the face
- interpreting analysed information into a user state model
- determining what to say and what listener behaviour to show
- synthesising expressive speech and non-verbal vocalisations
- generating expressive 3d head and face animations

**Distributed component integration framework**

- based on Message-oriented middleware ActiveMQ
- standard W3C interface formats: SSML, EMMA, ...
- Java/C++; Windows, Linux, Mac OS X

**Iterative cycle of improving and tuning:**

- initially train analysis components on preliminary data, formulate interpretation and generation rules intuitively
- as data is becoming available, retrain and update rules

**System is publicly available**

- most components open source
- download from www.semaine-project.eu

**System evaluation**

- identify determinants of a successful interaction in the SAL scenario
  - test Wizard-of-Oz and automatic human-machine dialogues
  - give human operators limited information to provoke failures
  - no speech, filtered speech, full speech setups: how important is the prosody vs. verbal content?

- develop measures of conversation breakdown
  - give users a "yuk" button they can press when the conversation feels anomalous
  - include "integrated evaluations" into the setup, where users are queried about the quality of interaction within the system
  - annotation of interaction breakdown in recorded dialogues
  - relate these to objective measures: learn to identify situations of conversation breakdown to trigger repair strategies

**Future plans**

- Sensitive Artificial Listeners have the potential to become "experts" of a user's characteristics
  - the meaning of a given user's non-verbal expressions
  - their enduring preferences

- To be able to do that, SALs will need memory and adaptation+learning capabilities

- Their expertise can potentially make human-machine interaction more robust and natural in a broad range of multimodal interaction scenarios

www.semaine-project.eu