

## Multilingual Processing Systems as Mediators in Human Communication

November 22, 2006

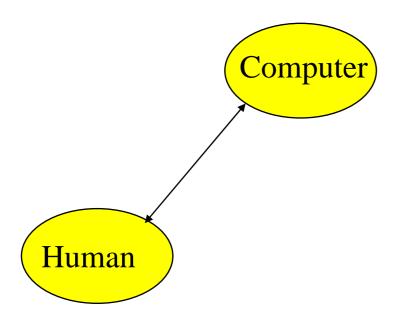
Alex Waibel
Interactive Systems Laboratories
Carnegie Mellon University
University of Karlsruhe
<a href="http://www.interact.cs.cmu.edu">http://www.interact.cs.cmu.edu</a>







## **Interact** Different Roles for Humans and Computer









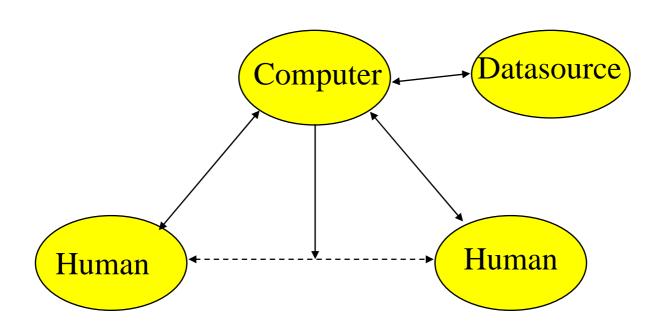
## interact Human-Human Interaction







## interact Different Roles for Humans and Computer







## The CHIL Project

### The CHIL Team:





Universität Karlsruhe (TH)

### Fraunhofer

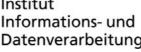
DaimlerChrysler

Institut Informations- und Datenverarbeitung













Centre de Tecnologies i Aplicacions del Llenguatge i la Parla

INFORMATION TECHNOLOGY

UNIVERSITAT POLITÈCNICA DE GATALUNYA









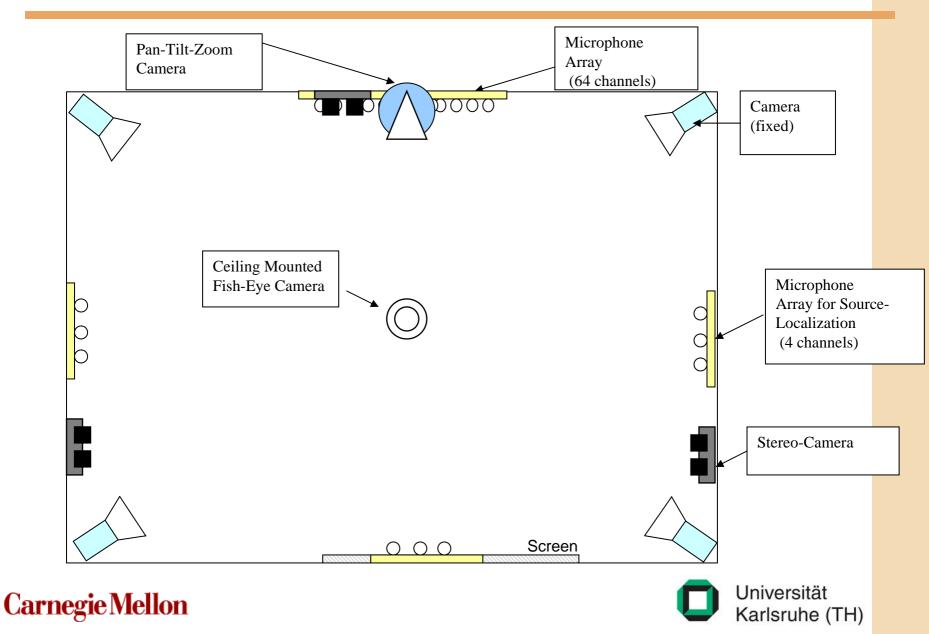






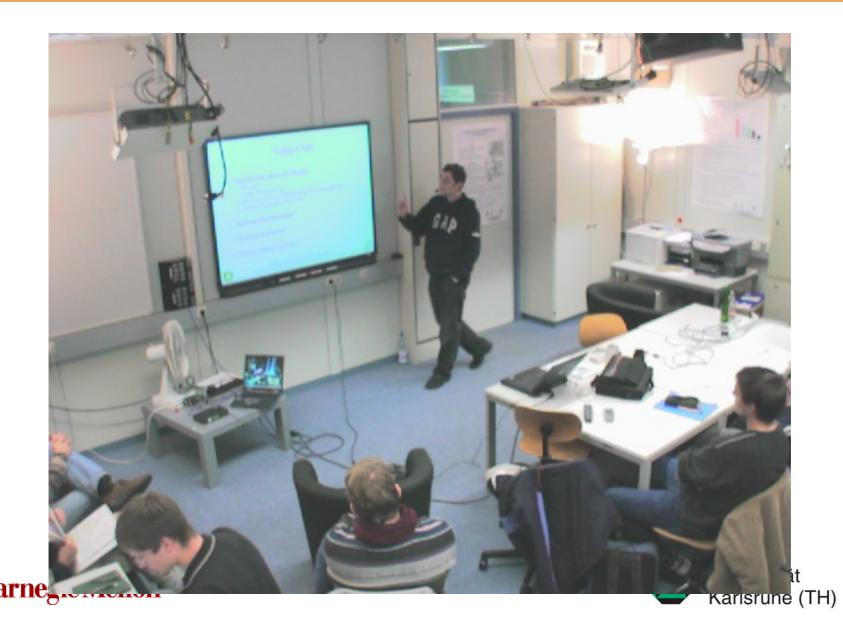


### Sensors in the CHIL Room



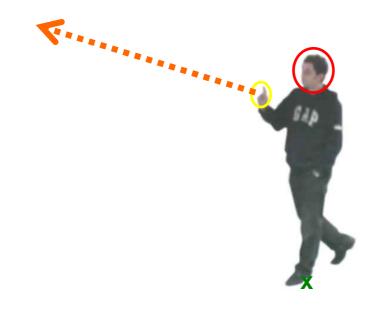


# **Interact** Describing Human Activities





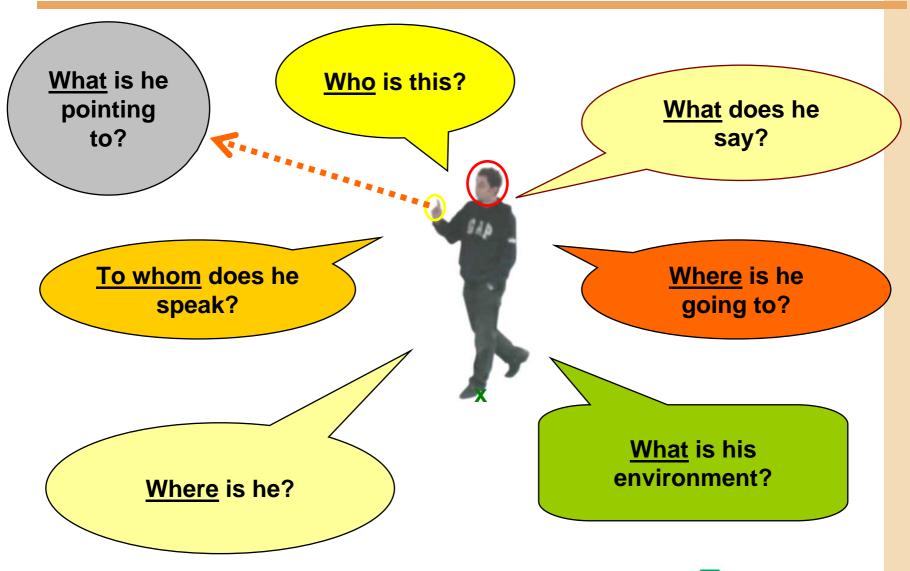
## Describing Human Activities







### Technologies/Functionalities



Carnegie Mellon





Carnegie Mellon

### Results, June 2004

### **Face Recognition (7 subjects) Speech Recognition** • 76% with manual alignment • Close talking: 37% WER •15% fully automatic • Far-field: 65% WER **Speech Detection** • 9% Mismatch rate (CTM) • 12.5% far field **Head Detection:** • 78% correct (error < 15 pixel) **Head Orientation:** Source Localization: Mean error ca. 10° • 11° root mean square error **Hand Tracking:** • 73% correct **Speaker ID: 3D Pointing Gestures:** • 100% correct, after 30s • 75% Recall • 77% Precision **Body Tracking:** Accoustic event classification • 80,7% correct (error < 30 cm) (25 classes) • mean error: 22 cm • 38,4% error

Universität

Karlsruhe (TH)



## Phone Calls During Meetings







## Memory Jog

....What was his name? ...Where did I meet him? ...What happened at the last meeting?





....What was his name? ...Where did I meet him? ...What did we discuss last time?





## Language Support

....what is he saying?







## **Speech Translation**

### History:

- Domain Limited, Clear Speaking Style (late 80's-91)
  - Janus (first European&US speech-to-speech system)
  - ATT, NEC, ATR
- Domain Limited, Spontaneous ('91-'00)
  - Janus II/III (work on 20 languages),
     Verbmobil, Nespole, Enthusiast,
     C-STAR, ATR, ETRI, NLPR,...
- Fieldable, Domain Limited, Spontaneous (current)
  - Transtac, Babylon, Phraselator, Thailator, ....
- Domain-Independent Speech Translation (current)
  - TC-STAR, STR-DUST, GALE







## **interact** Fieldeable Domain Limited Speech Translation

# Fieldable Systems: PDA Speech Translators

- Tourism
  - Conferences
  - Business
  - Olympics
- Humanitarian
  - Refugee Registration
  - First Responder
  - Healthcare
    - USA, Latino Population
    - Europe, Expansion
    - Third World
- Government
  - Peace Keeping, Police





# interact Pocket Interpreter (Mobile Technologies)





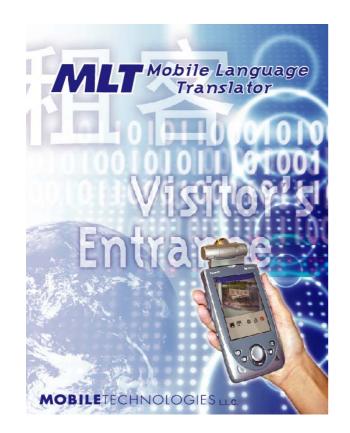
## interact Multilingual Communication







## **interact** Sign Translator







Pocket Translator of Foreign Signs (Mobile Technologies, LLC Pittsburgh)





## Missing Science

# Problem 1: Domain Limitation cannot handle:

- TV/Radio Broadcast Translation
- Translation of Lectures and Speeches
- Parliamentary Speeches (UN, EU,..)
- Telephone Conversations
- Meeting Translation









## Translation of Speeches (TC-STAR)







# interact Lecture Translator









## Main R&D Challenges

- Performance, Robustness
  - Spontaneous and Conversational Speech
  - Noise Channel, Environment, Distance from Mic, Stress/Emotion
  - Coverage, Named Entities, Plausibility
- Integration/Human Factors
  - Platform, Footprint, Comp. Requirements
  - Human Factors, Appropriate User Interface
  - Interaction H-H & H-M
- Domain, Scope, Coverage
  - Scope: Phraselator → Domain Limited Two-Way → Domain Unlimited
  - New Domains Vocabularies, Grammars (Speech/NLP)
  - Interactive Tools for Adaptation in the Field
- Multilinguality, Portability
  - Problems: Morphology, Lexica, Scripts, Foreign Accents, Names
  - 6000 Langauges:
    - Language Portability Language Independence and Adaptation
    - Data Sparseness



