# Using Non-Verbal Information to Augment Designs of Language-based Interactions

Hao-Chuan Wang . 王浩全

*Department of Computer Science*

*Institute of Information Systems and Applications*

*National Tsing Hua University*

http://www.cs.nthu.edu.tw/~haochuan


*Dec 18, 2015 @ 2015 IR Workshop, Academia Sinica*

# About Me

**Hao-Chuan Wang** 王浩全

**Assistant Professor, NTHU** (Feb 2012 –)

國立清華大學 National Tsing Hua University

**PhD & Postdoc, Cornell** (3.5 years)

**PhD student, Carnegie Mellon**
(2 years; transferred to Cornell)

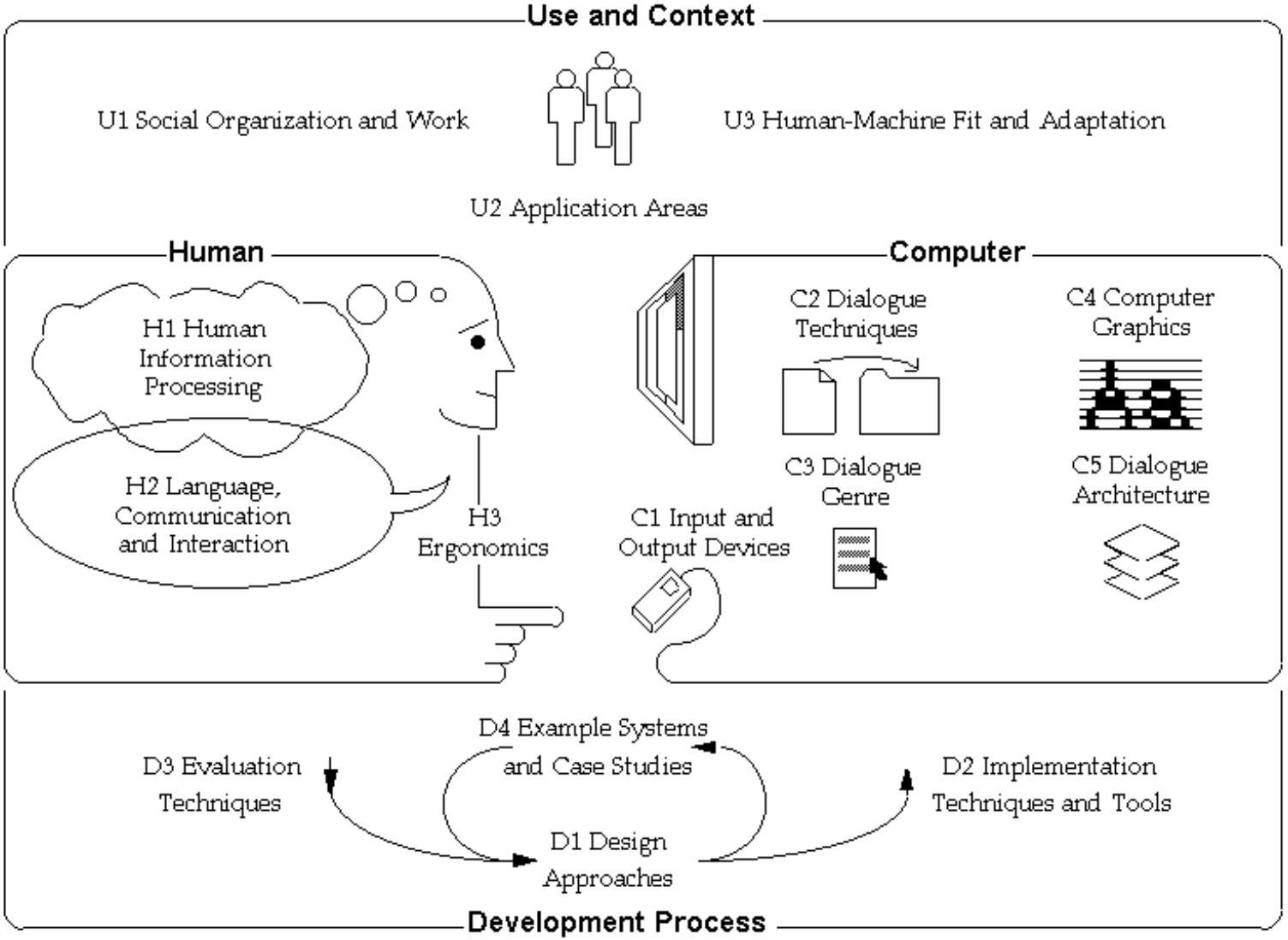
**Human-computer interaction w/ focus on social technology**

- ACM CSCW Steering Committee
- ACM CHI Associate Chairs (2014, 2015, 2016)
- ACM CSCW Associate Chairs (2013, 2015)
- Chinese CHI Steering Committee & Program Co-Chairs (2014, 2015)

SIGCHI special interest group computer human interaction

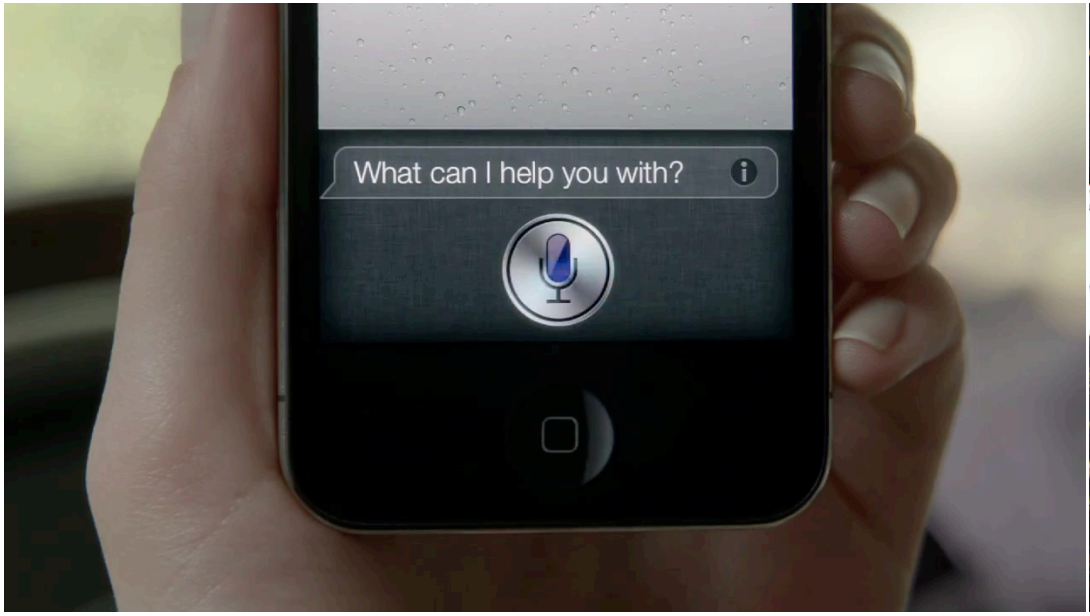
# *A Quick Overview of Human-Computer Interaction, HCI*

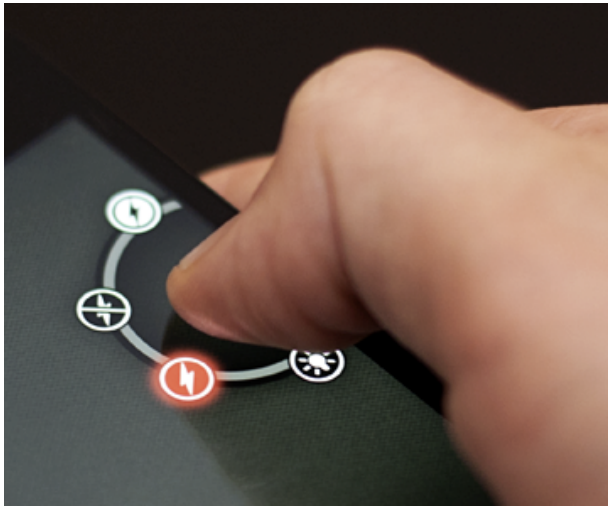
# HCI: Studying Existing and Possible Relationships between Computers and People




**ACM SIGCHI Curricula 1996**

# The two “senses” of Human-Computer Interaction: From interface …

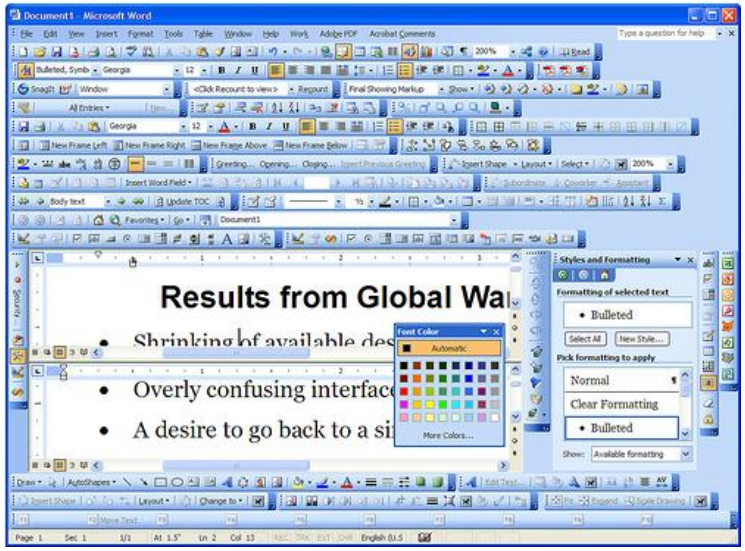


**HCI as a topic:** “Interaction” in the sense of input and output between machines and people.

# … to problem solving and value creation in the real world

**Identifying & fixing usability problems**




**Persuasive (behavioral change) computing**

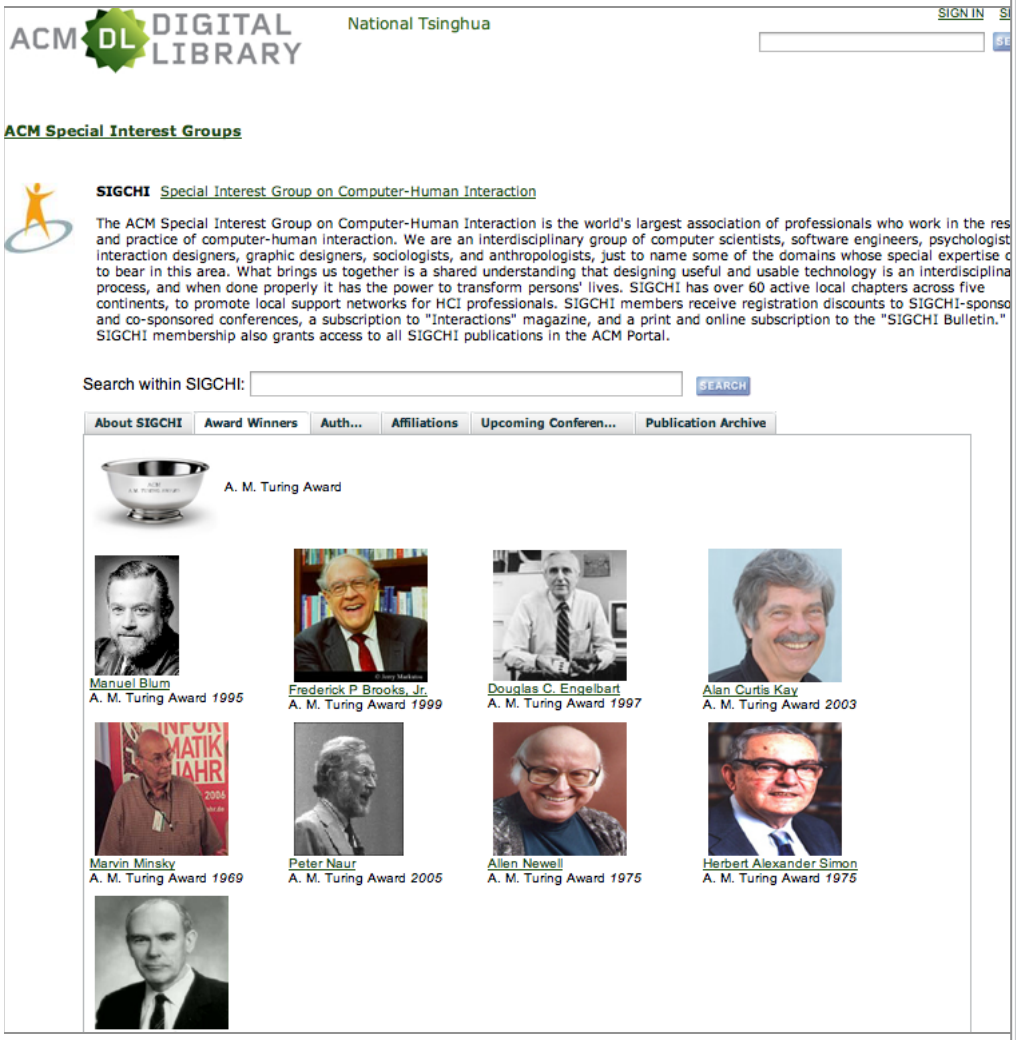
**Applications of HCI (not all the same as HCI applications)**

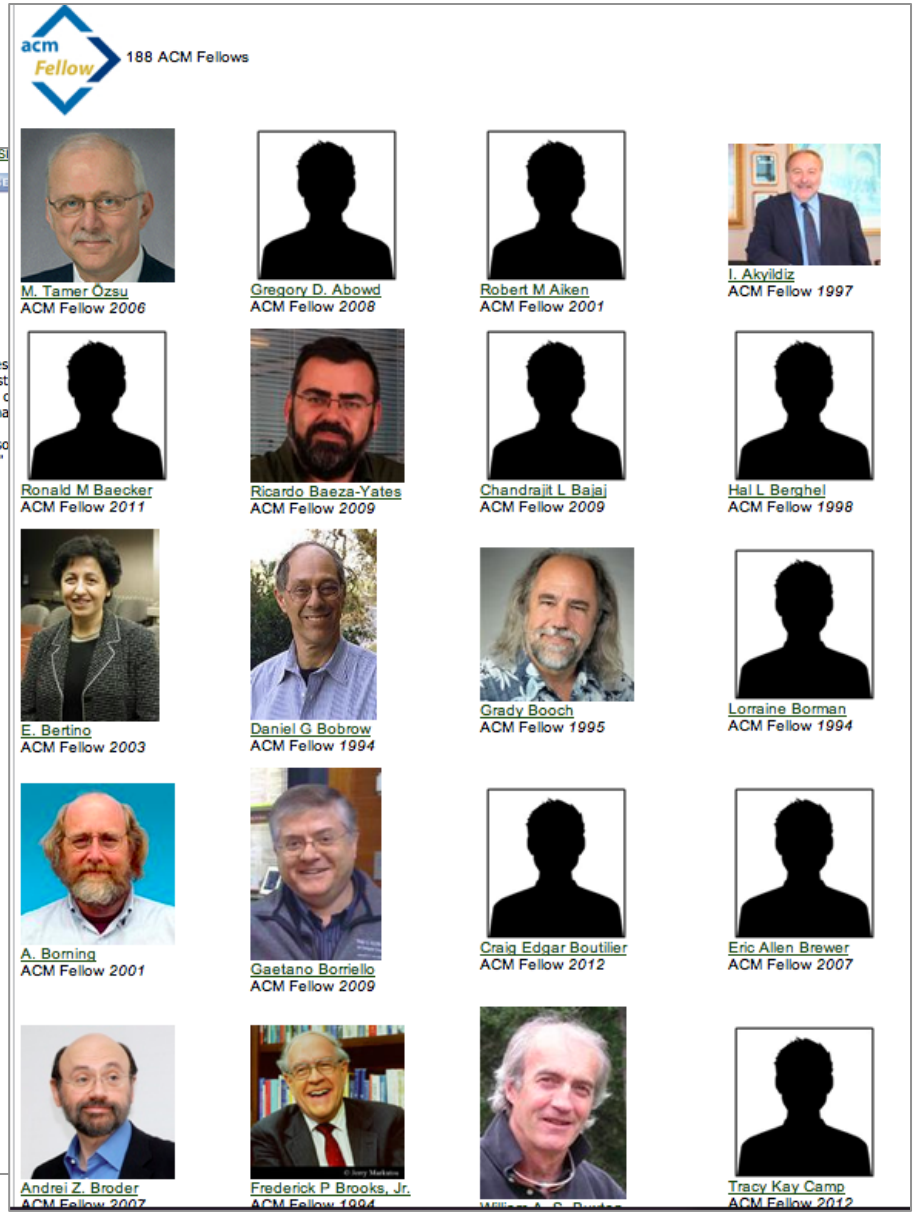
**HCI as a discipline:** "Interaction" in the sense of designing, appropriating technologies based on user needs, goals, constraints, and characteristics.

# 30+ Years of the HCI Community

**ACM SIGCHI:**

9 Turing Award Winners / 188 ACM Fellows




http://dl.acm.org/sig.cfm?id=SP923

# *Language-based Interactions as an HCI Problem*

# The Origin of Language-based Interactions

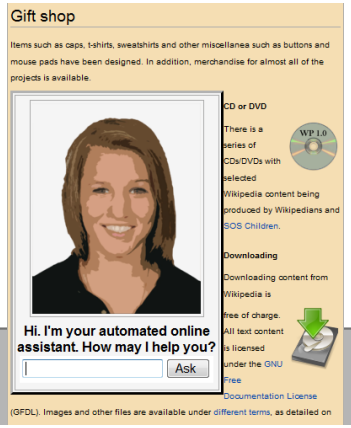
People use language to communicate as one common way of social, interpersonal interactions

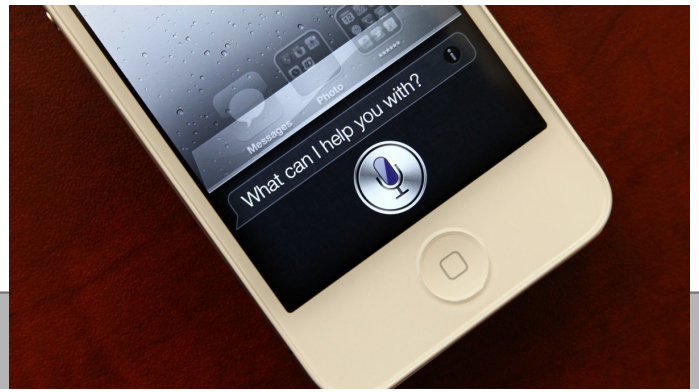
- A joint, collaborative effort on message exchange
- One states, another follows
- One asks, another responds
- An important basis for work, learning, relationship building etc.

# User Interface as Communication *Metaphor*

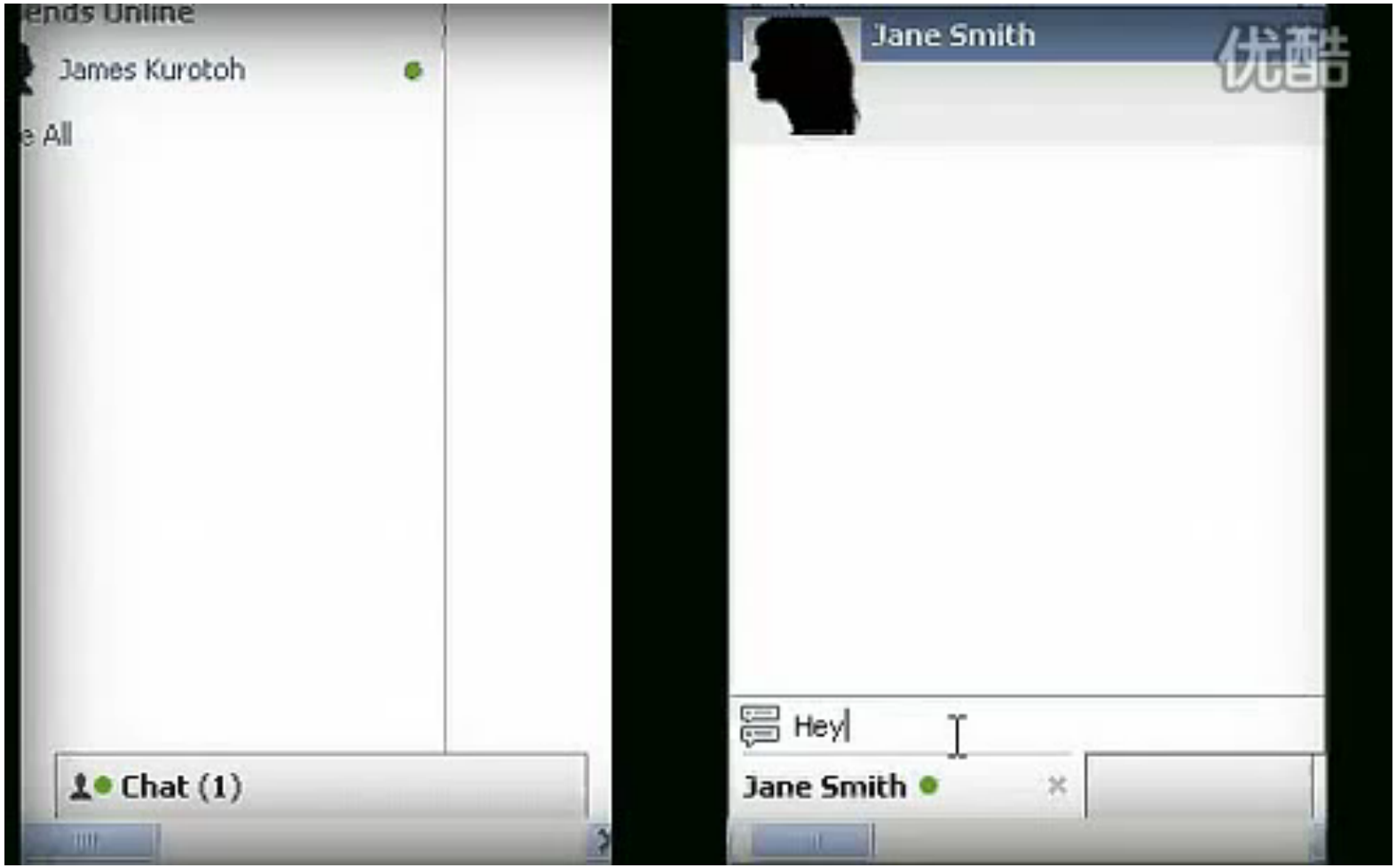
When considering the design of machine-user interface, it's useful to consider it as a type of *communication* (between human and system)

- Language as a model of design. Command-line interface, dialogue system, search user interface etc.
- But, numerous problems at the level of interaction and human factor can emerge…



What's the longest distance in the world?
世界上最遠的距離是什麼？


ends Online
James Kurotoh
e All
Chat (1)
Jane Smith
优酷
Hey
Jane Smith

# Observations from Human-Human Communication

**Human as both the producer and consumer of messages**

- ***Problem of overloading due to multitasking***

**Norm of conversational interaction**

- Ideally, we expect quality, quantity, relevance and manner in conversations (Grice's maxims)
- ***Problem of imperfect processing***
  Practically, either human or machine can be poor at any of these (can we always express clearly and efficiently, and stay relevant?).

**Requirement of linguistic/communication skills**

- Training/learning may be required
- Barriers between educational levels, professions, languages and cultures.

# So Maybe Just Avoid Language?

## Multimedia/multimodal approach

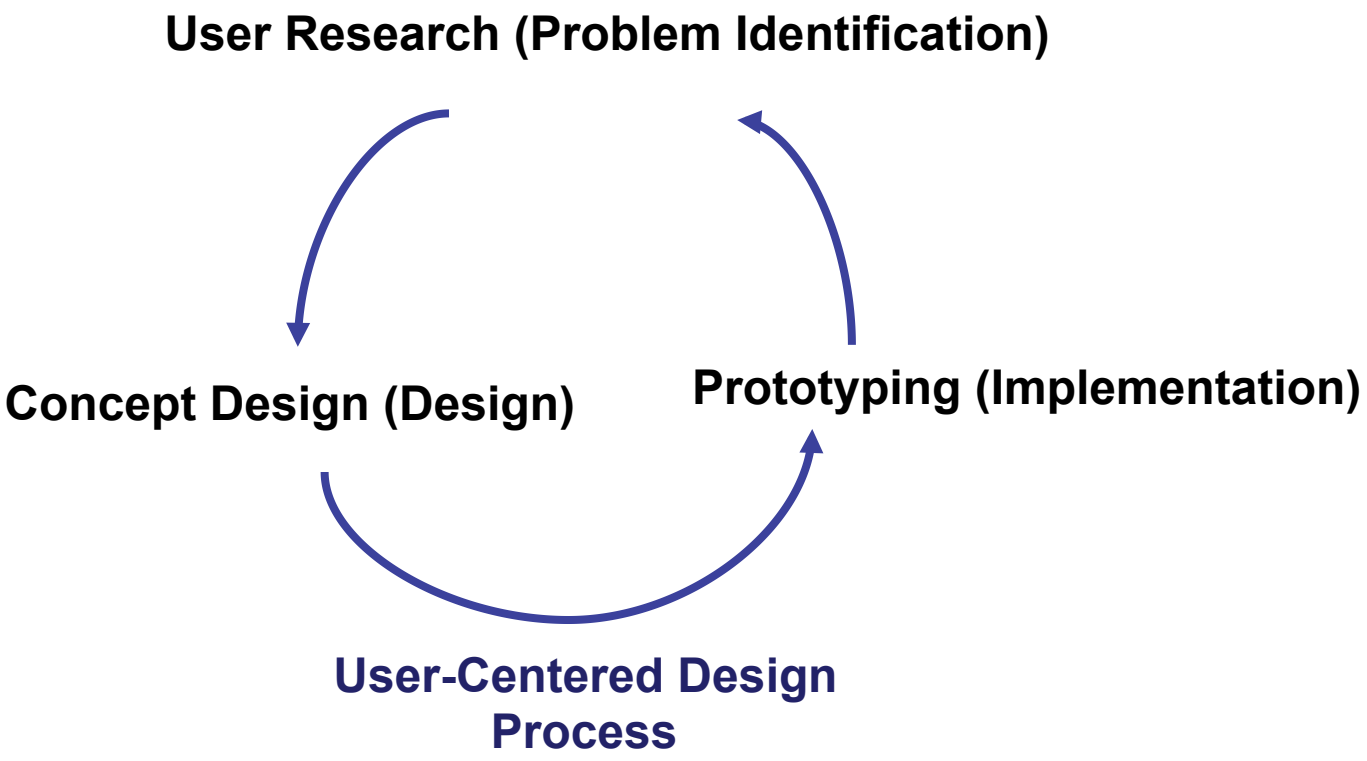
## Video conferencing

- Mimicking face-to-face communication
  - "The bigger, the better"
- Does it work? Sometimes.
  - Adoption is not guaranteed.
    Privacy and other social concerns
  - Video along cannot deliver meaning

## Using signs/gestures

- Signs as language (still need learning)
- Gestures can be ambiguous

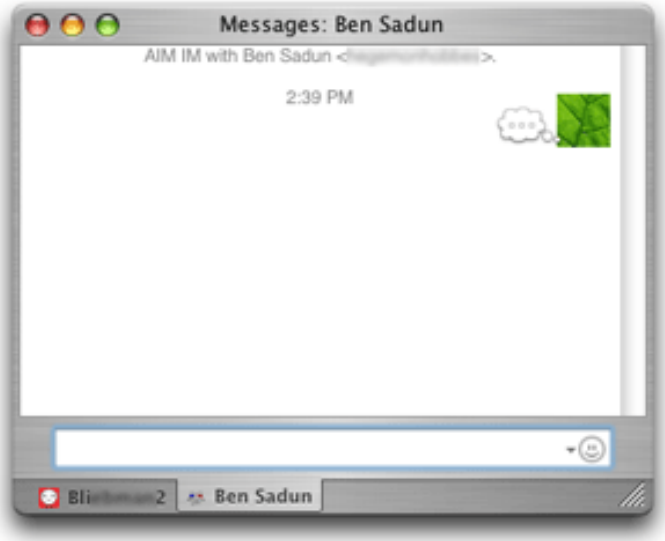
# Applying Design Thinking



# Design Approach: Small Tweak, Huge Impact

Designs of interaction can work better when features and constraints of human communication are understood and considered.

Ex. Non-verbal *awareness indicator* that makes "typing" visible in instant messaging.




**Identify key problems to solve. Introduce suitable non-verbal information in designs.**

# *Grounding Communication*

# How Would You Describe…

Where you live in Taipei?

Where you lived when you were in U.S.?

# My Answer

Where you live in Taipei?

In 南勢角.

Where you lived when you were in U.S.?

In Ithaca, a college town in the middle of New York state if you know where it is. It's where Cornell University is located.

# My Answer

Where you live in Taipei?

In 南勢角.

Where you lived when you were in U.S.?

In Ithaca, a college town in the middle of New York state if you know where it is. It's where Cornell University is located.

**Do you see the general difference? Why?**

# My Answer

Where you live in Taipei?

In 南勢角.

Where you lived when you were in U.S.?

In Ithaca, a college town in the middle of New York state if you know where it is. It's where Cornell University is located.

**Do you see the general difference? Why? The amount of knowledge that we shared.**

# Common Ground

Knowledge, beliefs, attitudes we share,

and know that we share,

and know that we know that we share,

influence how we use language to communicate.

**Grounding:** Interactive process

by which communicators exchange evidence of their understanding to arrive at the state of **common ground.**

Herbert Clark
Stanford University

# Evidence of Common Ground

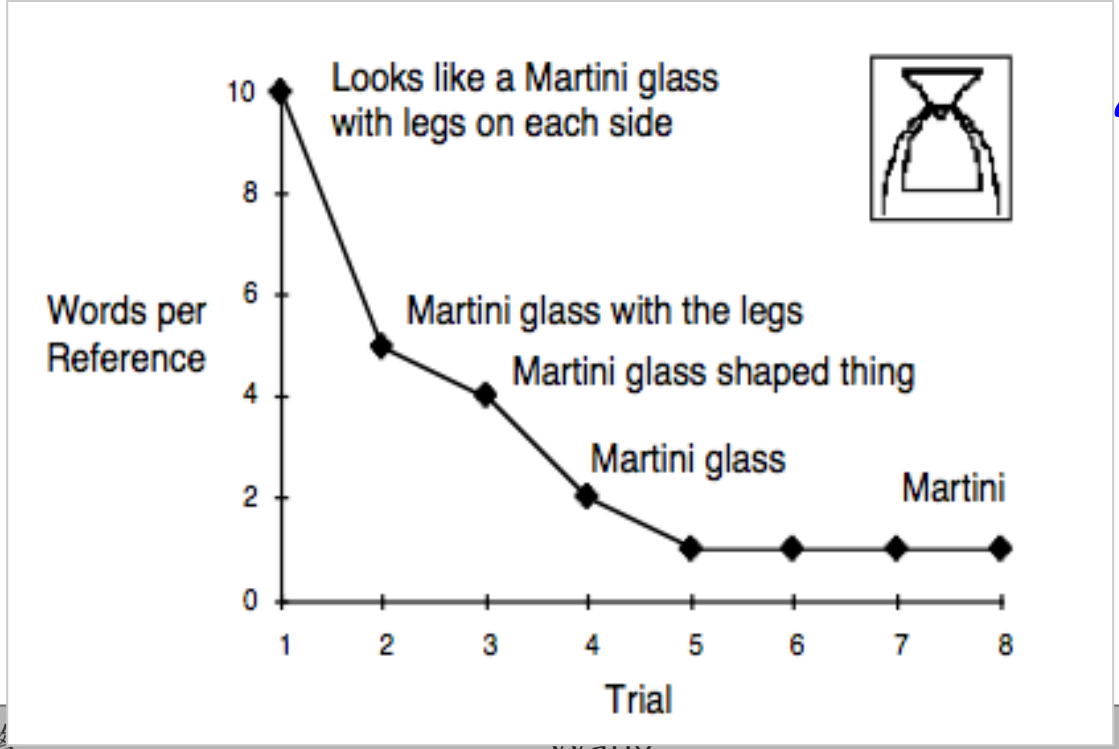
Physical co-presence (being co-located)

- "close **that door**"

Shared community membership

- "Let's meet at 小七"

Linguistic co-presence (can access same utterances)




**"What's this?"**

# The Role of Media: Affordances

An influential HCI-rooted concept, which roughly means "action-permitting properties" of objects that people see

- Chair **affords** sitting
- Door-knob **affords** door-opening
- Virtual keyboard **affords** typing (but is this trivial?)

Don Norman

# Affordances of Communication Media

| Affordance | Face-to-Face | Video Conf. | Phone | Email |
|---|---|---|---|---|
| Copresence | ++ | ? | -- | -- |
| Visibility | ++ | + | -- | -- |
| Audibility | ++ | ++ | ++ | -- |
| Cotemporality | ++ | + | ++ | -- |
| Simultaneity | ++ | + | ++ | -- |
| Sequentiality | ++ | ++ | ++ | -- |
| Reviewability | -- | -- | -- | ++ |
| Revisability | -- | -- | -- | ++ |

[Clark & Brennan, 1991]

# Technology Changes Grounding

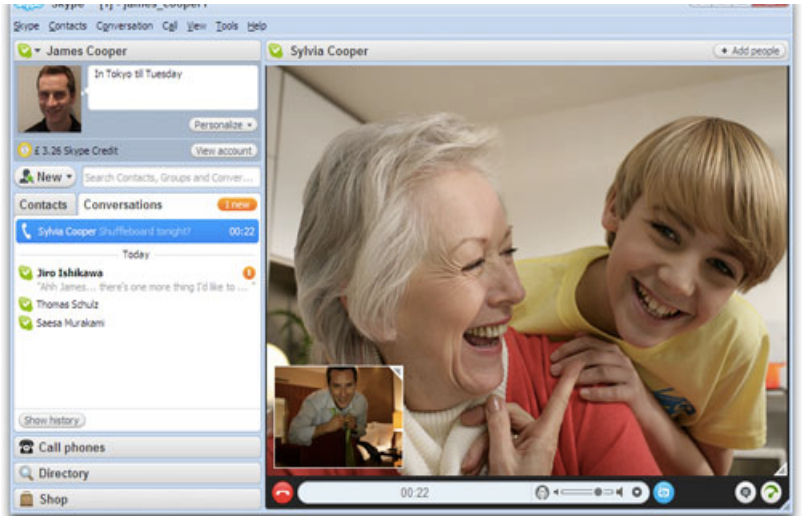
Affordances of media determine how people may interact with one another

- E.g., if no visibility, impossible to use head-nodding as a technique for grounding

People may learn to adapt their grounding behaviors (this happens. E.g., emojis in instant messaging)

Design *non-verbal features* to enable new ways of grounding in communication.


Gina
Green Tennis (6)
What are you doing?
Meeting Rebecca
Good


James Cooper
Sylvia Cooper
Contacts
Conversations
Call phones
Directory
Shop


twitter
A global community of friends and strangers answering one simple question: What are you doing? Answer on your phone, IM, or right here on the web!
Please Sign In!
Join for Free!

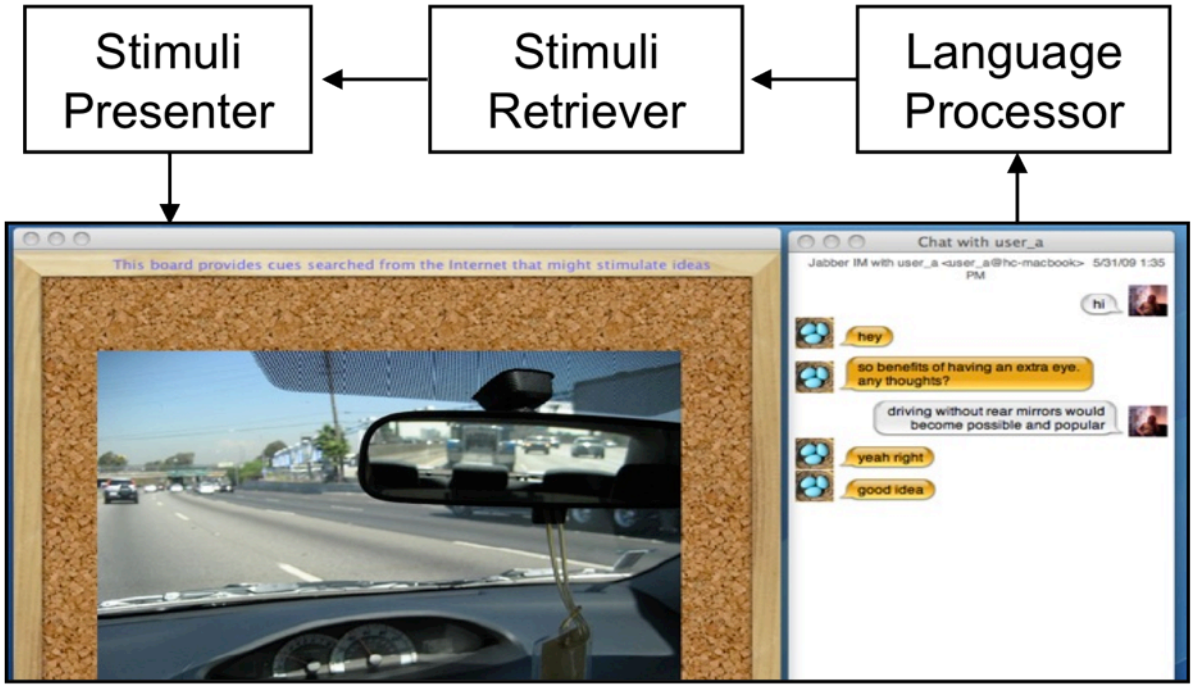

facebook
Haochuan Wang
Update Info
Activity Log
Friends 395
Photos 34
Map 11
Likes 54
What's on your mind?
Post
Activity
Haochuan is now friends with Yi-Ping Hung and 8 other people.
Haochuan liked Cornell Alumni Association.
Friends
395 Friends
Pei-Yao Hung
I attended Haochuan Wang's talk today. Learn some interesting perspectives he has toward HCI and CSCW research (and something else that might be even more valuable to me ^-^). The railway tickets do worth it, haha :)
Tsai Tung Jen
Interaction beyond the individual: A lecture on HCI-Oriented Collaborative and Social Computing Research –
Likes
July

# Sample Projects

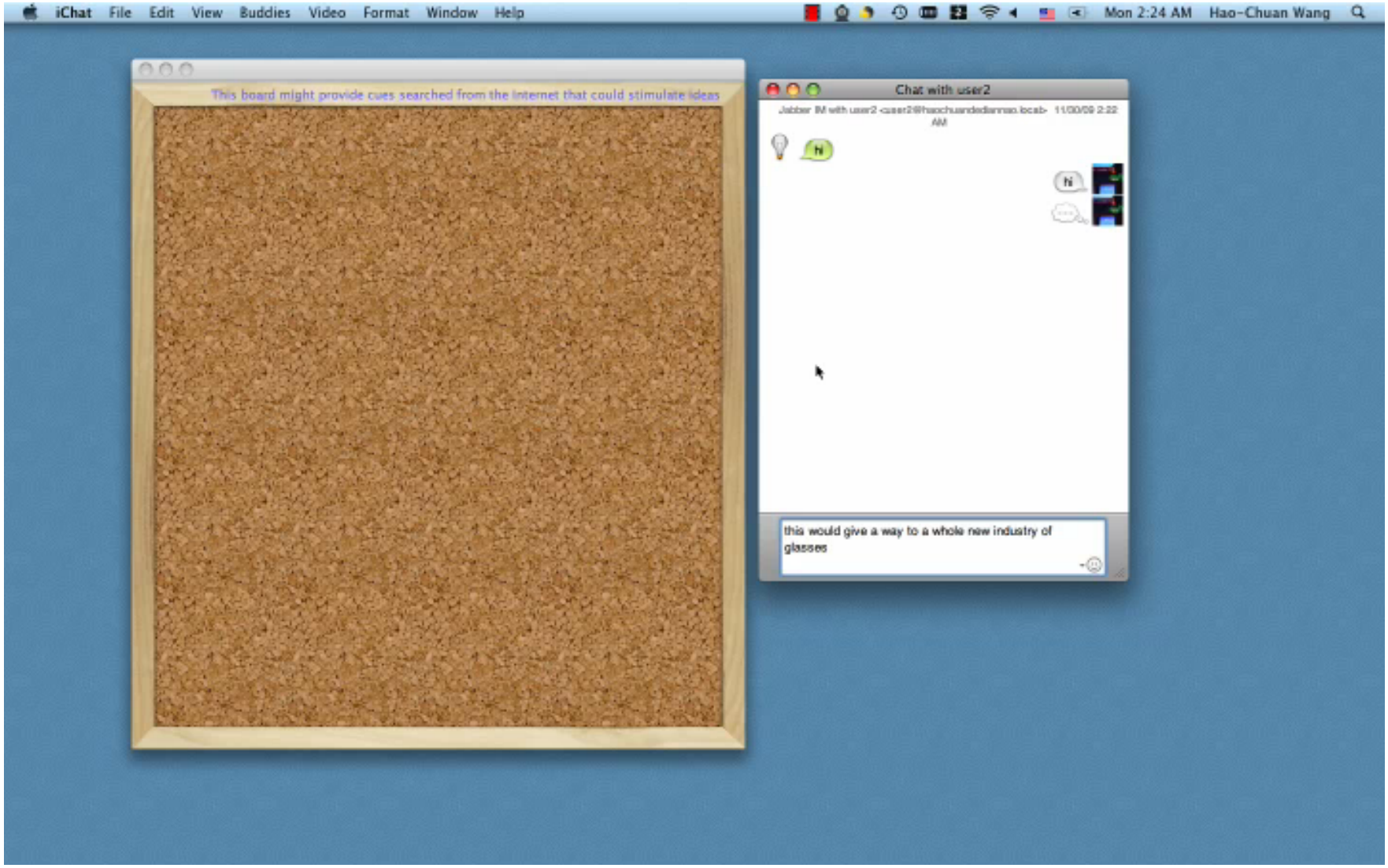
# IdeaExpander

**Augmenting Human Processing:** People can be bad at doing some work (e.g., creativity), and machines can possibly help out.

Ex. IdeaExpander- Supporting idea generation by visualizing ongoing conversations as relevant pictures.




[Wang et al., CSCW 2010, 2011] http://www.cs.cornell.edu/~haochuan/manuscripts/WangCosleyFussell_CSCW_10


iChat File Edit View Buddies Video Format Window Help
Mon 2:24 AM Hao-Chuan Wang
This board might provide cues searched from the Internet that could stimulate ideas
Chat with user2
hi
hi
this would give a way to a whole new industry of glasses

# Prototyping

**Language Processor**

SVM classifier, *f*: utterance -> (idea, ¬idea).

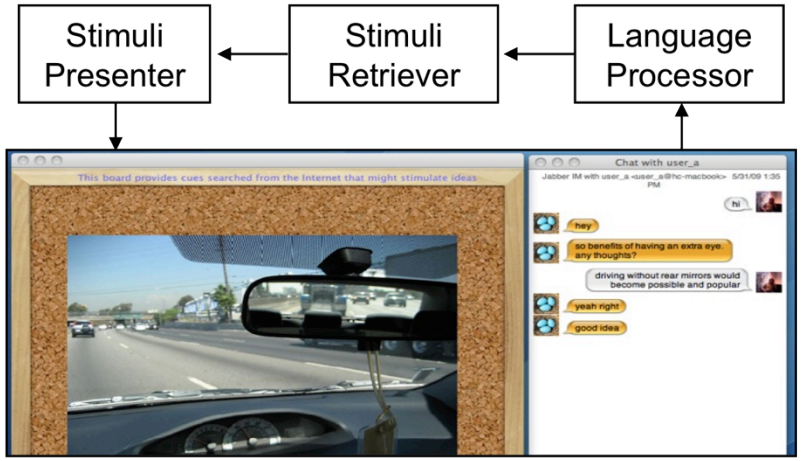
Feature sets: unigrams,

bigrams (markers of non-ideas, “I agreed”),
sentence length (ideas can have more words)

10 fold CV: 80% accuracy (.6 Kappa).

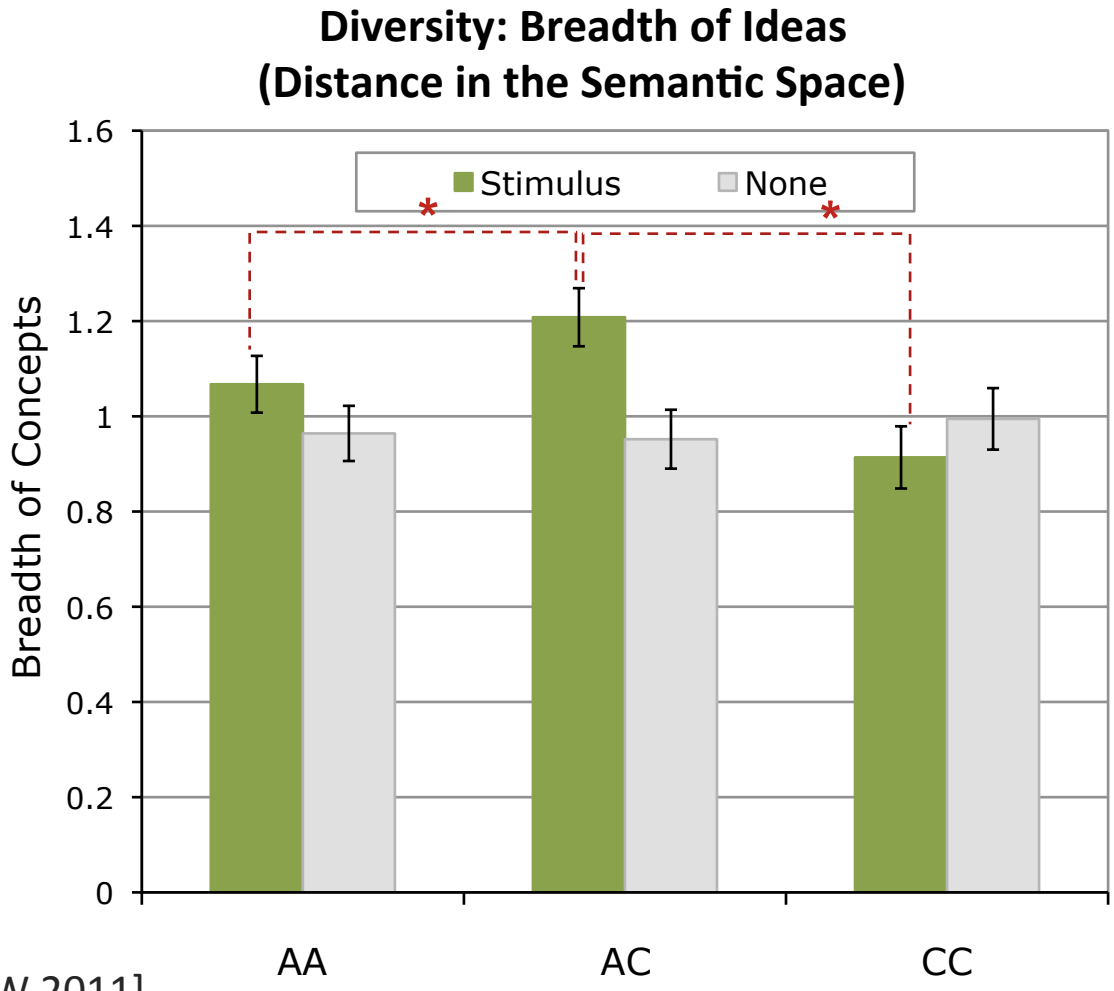
**Stimuli Retriever**

Using cosine similarity between input sentence and picture tags to rank candidate pictures.



# Impact on Group Brainstorming

With IdeaExpander, inter-cultural groups (AC) generated more ideas, and ideas with the greatest breadth.

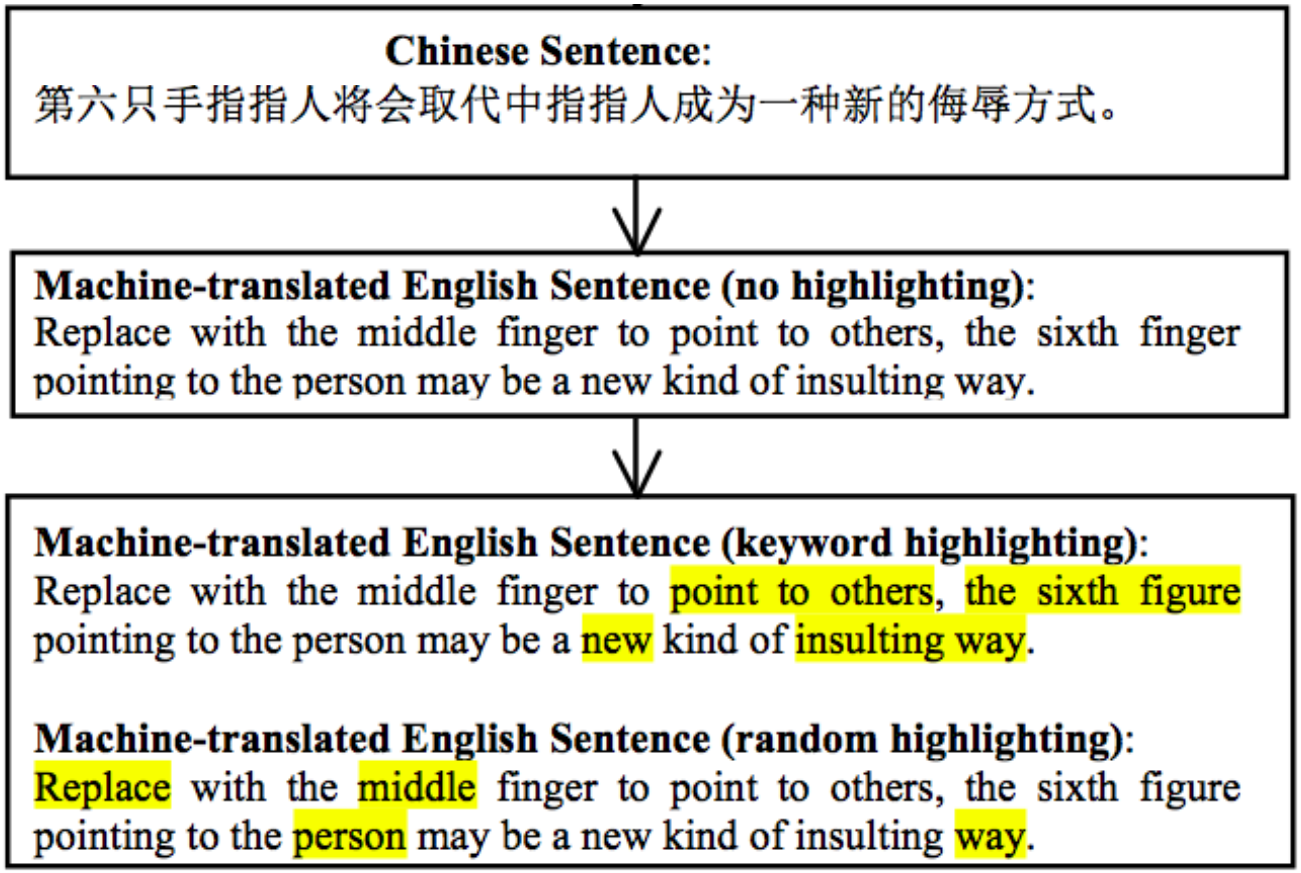



* p<.05

[Wang et al., CSCW 2011]

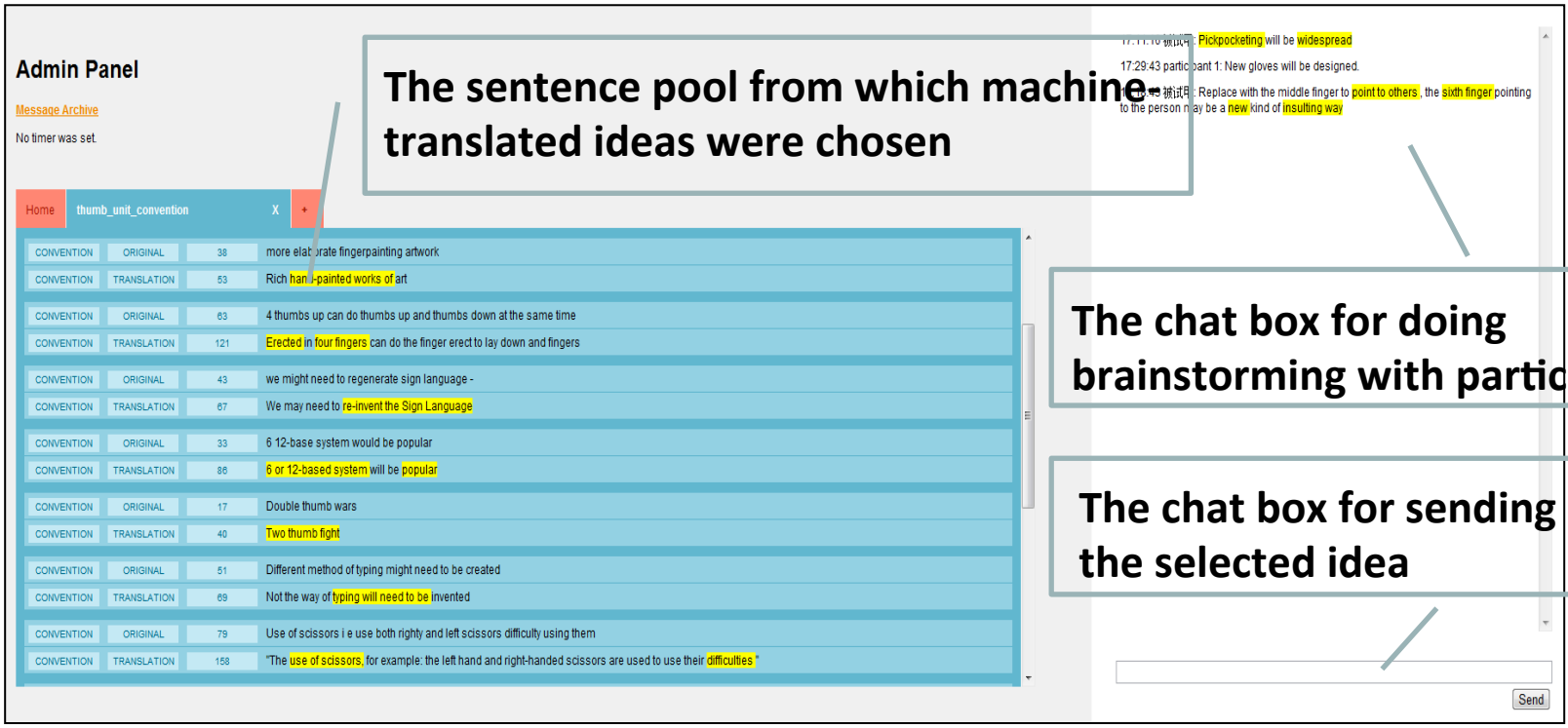
# Interfacing Machine Translation

**Enhanced machine translation** with keyword highlighting for cross-language communication (e.g., Chinese-English chat)




[Gao, Wang, Cosley, Fussell, CHI 2013]

# Wizard-of-Oz Study: Interface for the Confederate




[Gao, Wang, Cosley, Fussell, CHI 2013]

# Effects of Highlighting on MT Comprehension

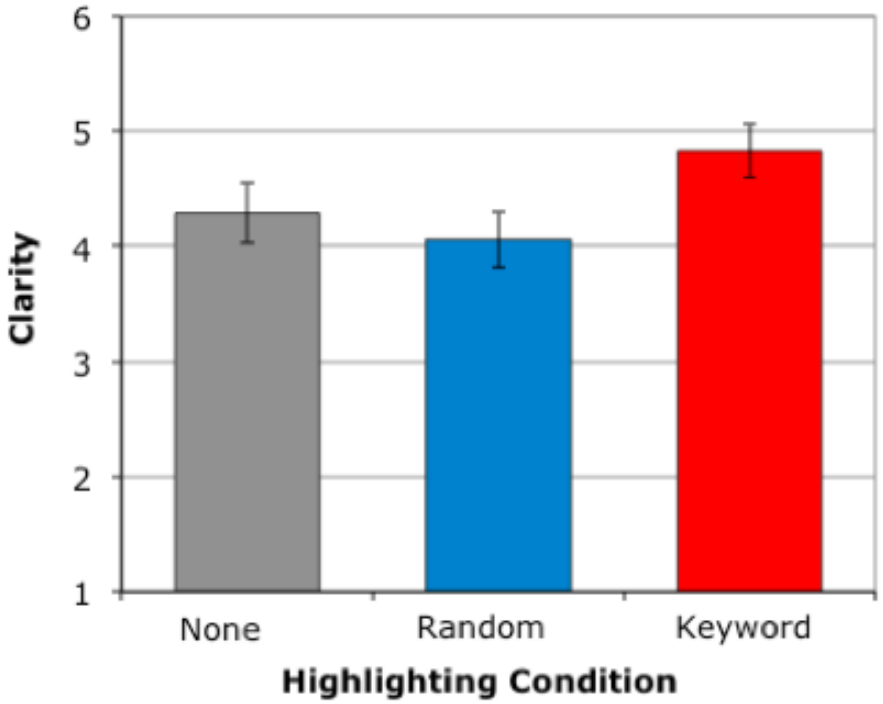



**Figure 4. Mean clarity ratings by highlighting condition (erro bars represent standard errors of the mean).**

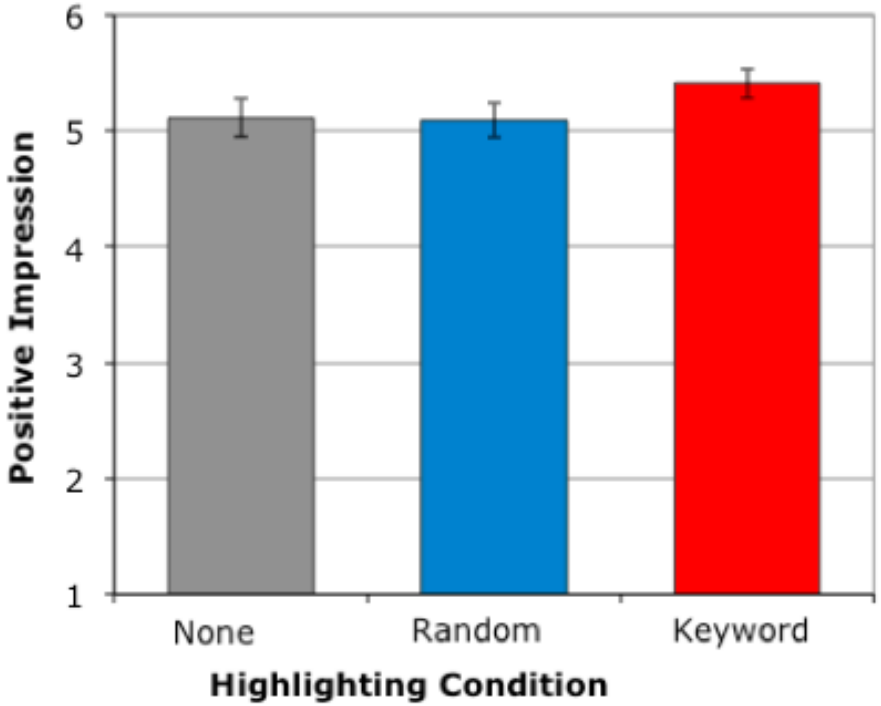



**Figure 6. Mean impressions of partner by highlighting condition (error bars represent standard errors of the mean).**

[Gao, Wang, Cosley, Fussell, CHI 2013]

# Effects of Interface Interactivity on Collecting Language Data to Power Dialogue Agents

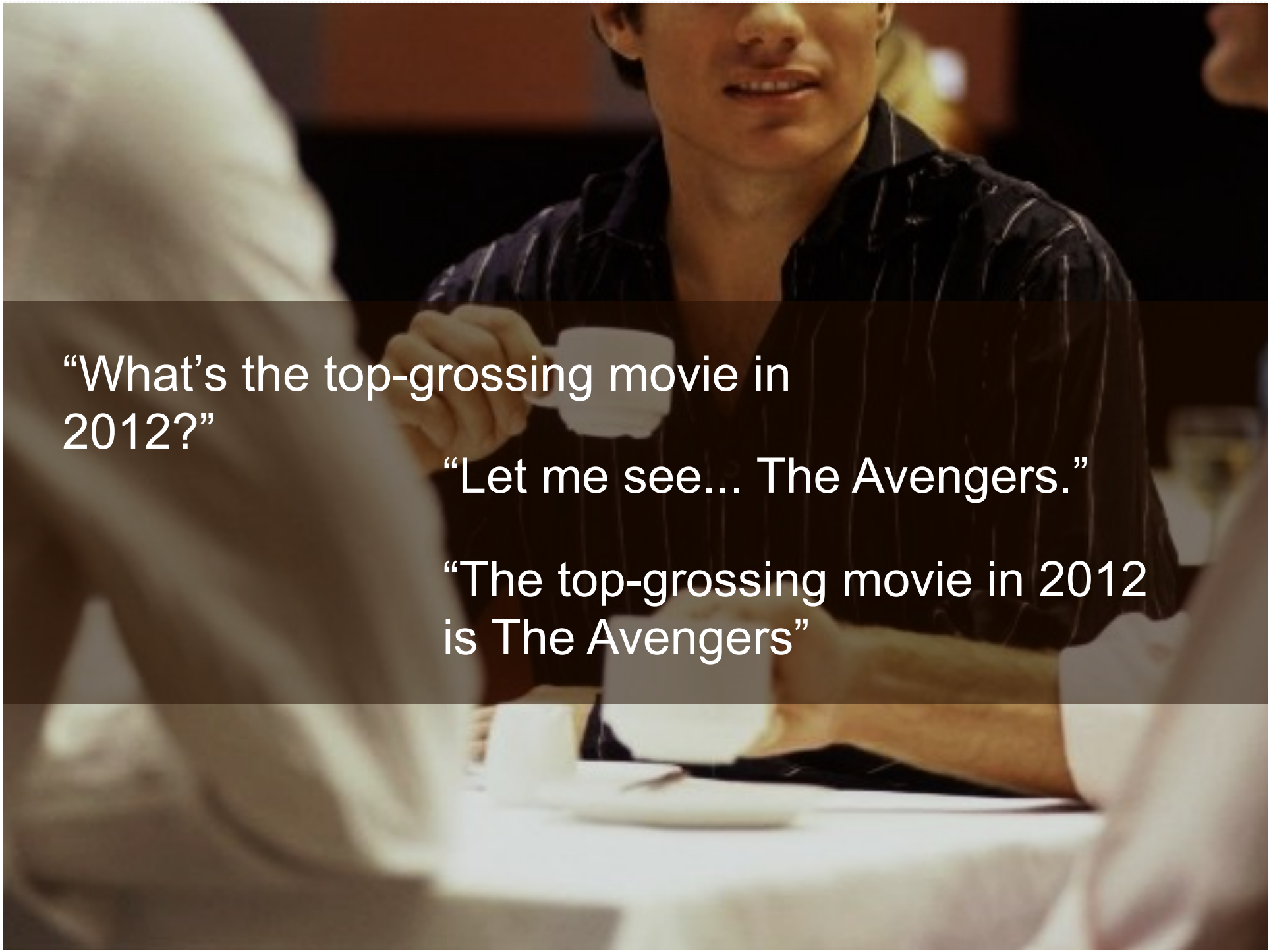
Hao-Chuan Wang, Tau-Heng Yeo, Hsin-Hui Lee, Ai-Ju Huang
**National Tsing Hua University, Hsinchu, Taiwan**

Jia-Jang Tu, Sen-Chia Chang
**Industrial Technology Research Institute, Hsinchu, Taiwan**

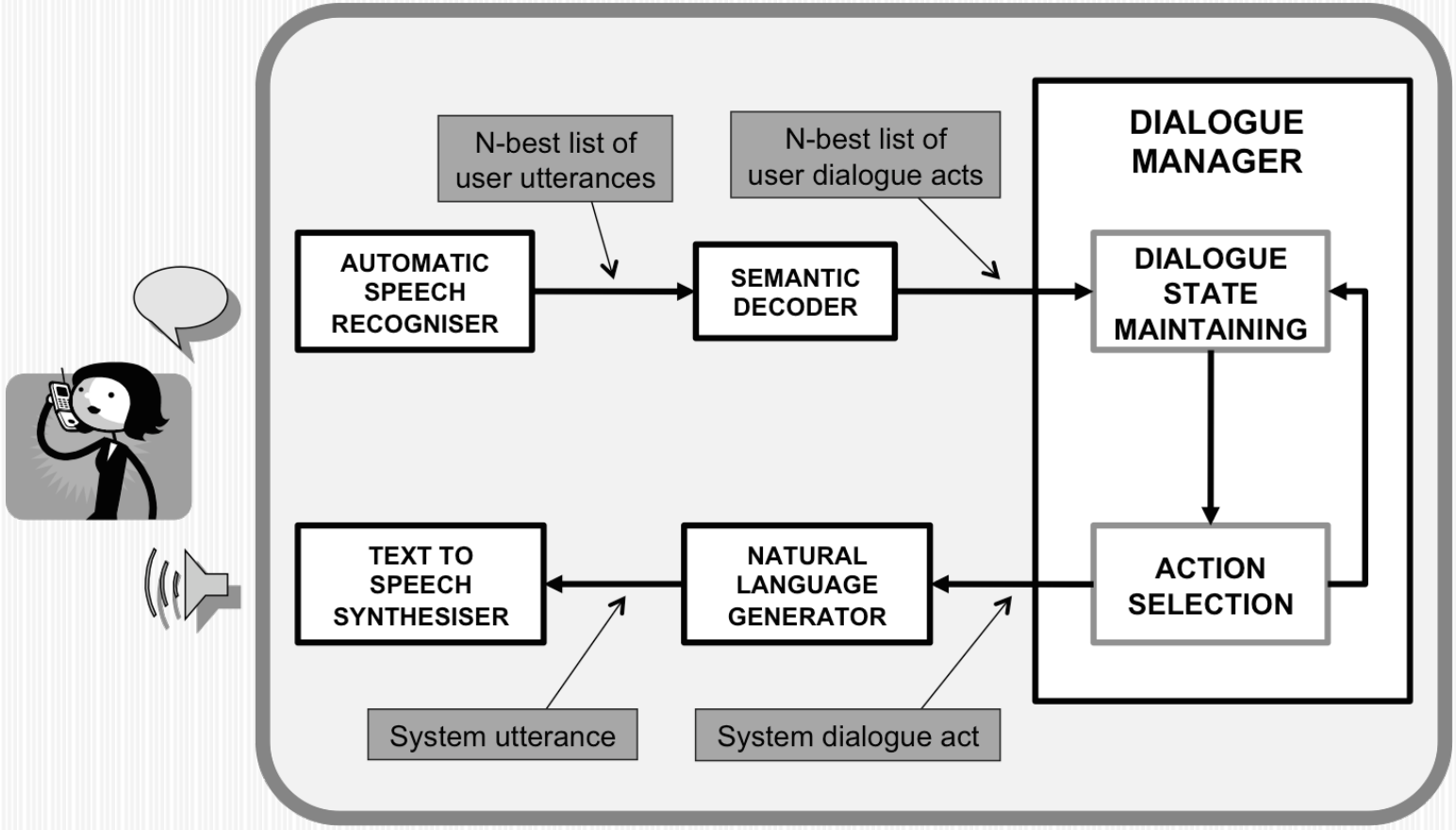


國立清華大學 NATIONAL TSING HUA UNIVERSITY

ITRI Industrial Technology Research Institute

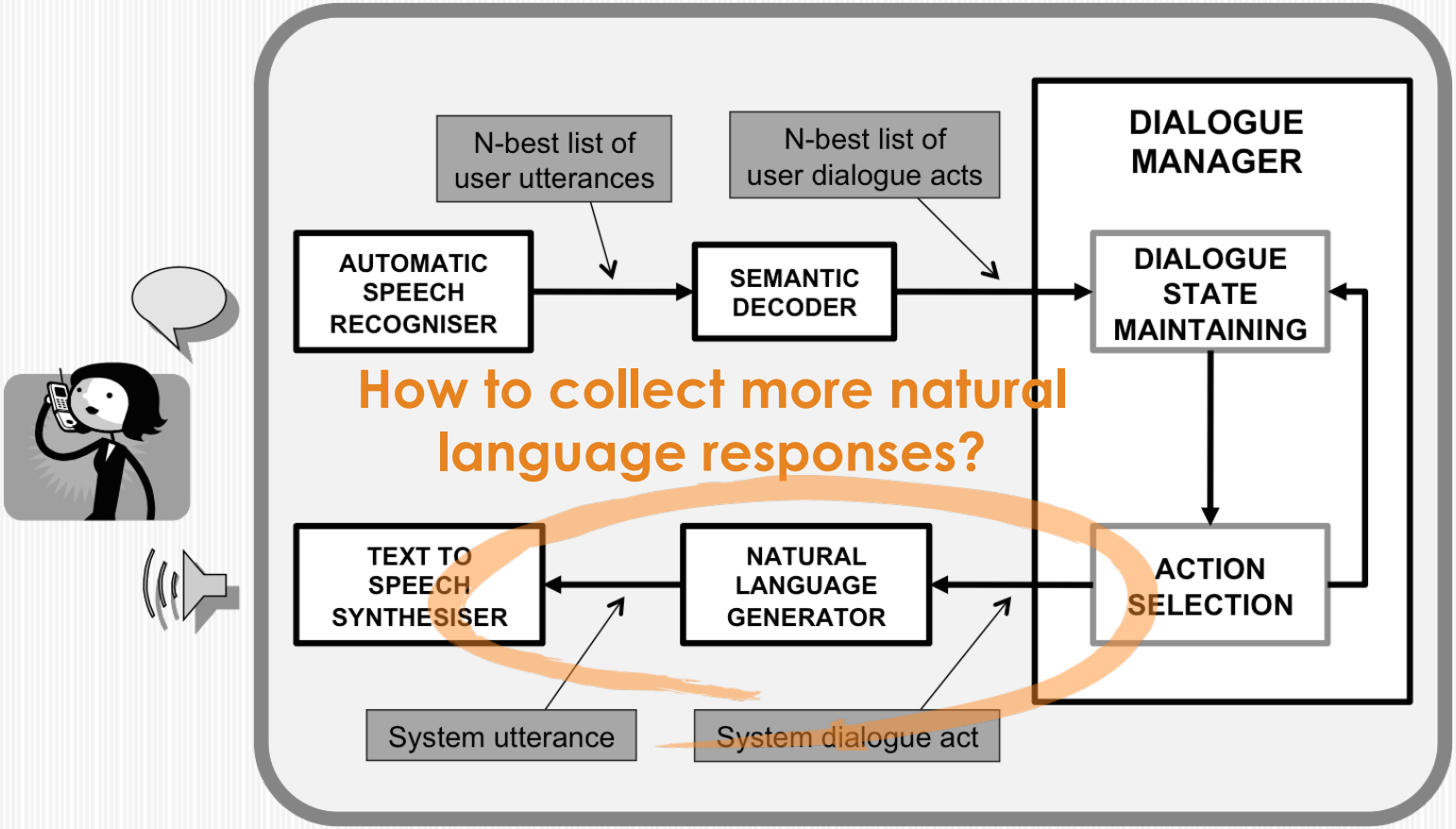

“What’s the top-grossing movie in 2012?”
“Let me see... The Avengers.”
“The top-grossing movie in 2012 is The Avengers”

# Spoken Dialogue Systems




Young, S., Keiser, S. & Gašić, M. Spoken Dialogue Management using Partially Observable Markov Decision Processes

# Language Generation Task



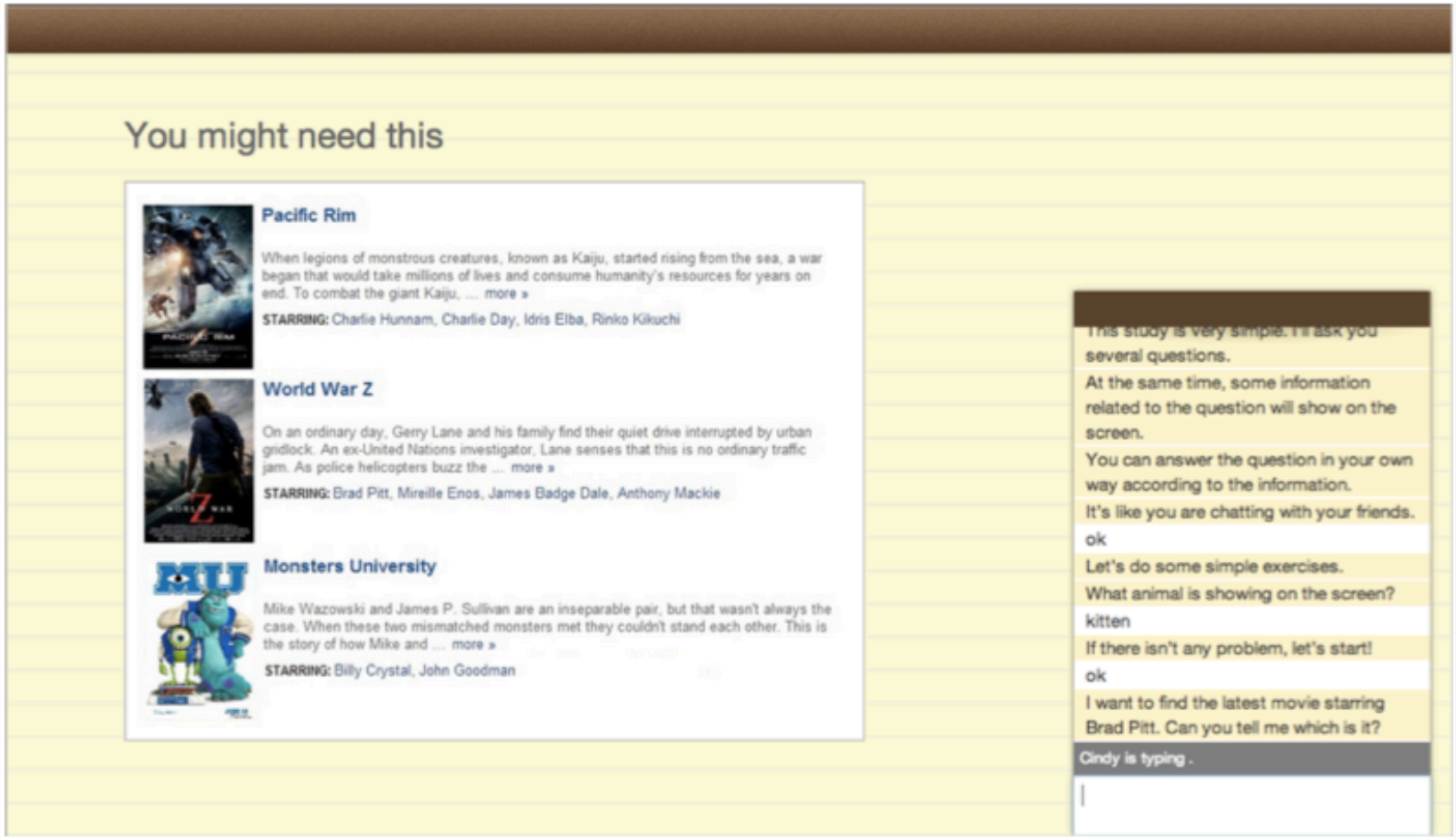

Young, S., Keiser, S. & Gašić, M. Spoken Dialogue Management using Partially Observable Markov Decision Processes

# Some Existing Methods

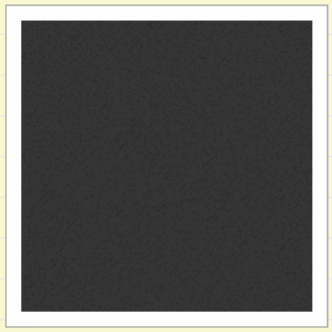
- One-on-one interviews to get the responses from people
  - Manual data collection.
  - Expensive.

- Using surveys with specific instructions, *"Imagine that you're answering people's questions …"*
  - Less expensive.
  - Non-interactive, "imagined interaction".

# Idea: Using an Interactive Chat Bot to Elicit Natural Responses



## You might need this

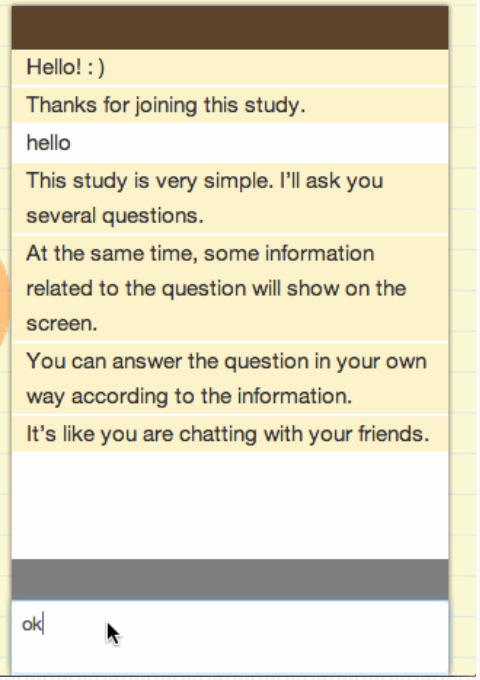


**Anthropomorphic features:**

✓Greet workers

✓Simulate human typing delays

✓Wait for response



# Static Interface

Hello! Thanks for answering this questionnaire. This questionnaire is very simple. I'll ask you several questions. At the same time, some information related to the question will show on the right of the screen. You can answer the questions in your own way according to the information.

What animal is showing on the right of the screen?

I kind of want to watch "Transformer". How are the reviews of this movie?

**Transformers** (2007) - User Reviews

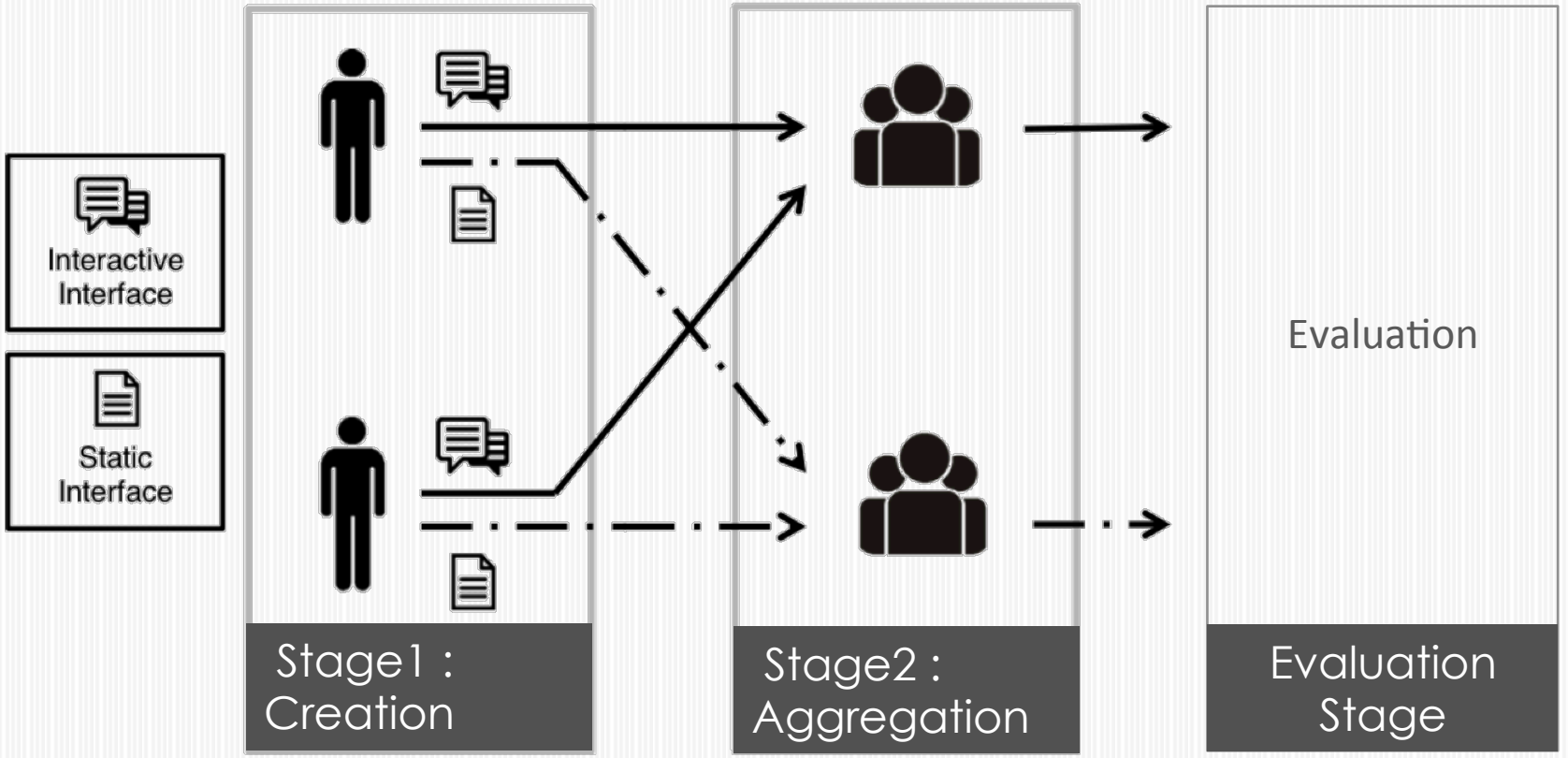
YAHOO! USERS:

★★★★½ 99,574 ratings

10520 USER REVIEWS

Transformers....A must see RATING: ★★★★½

# Crowdsourcing Answer Generation




Compare interactive and static interface

Crowdsourcing to select quality responses

Evaluate the results with end users

# Multilingual Crowdsourcing Study

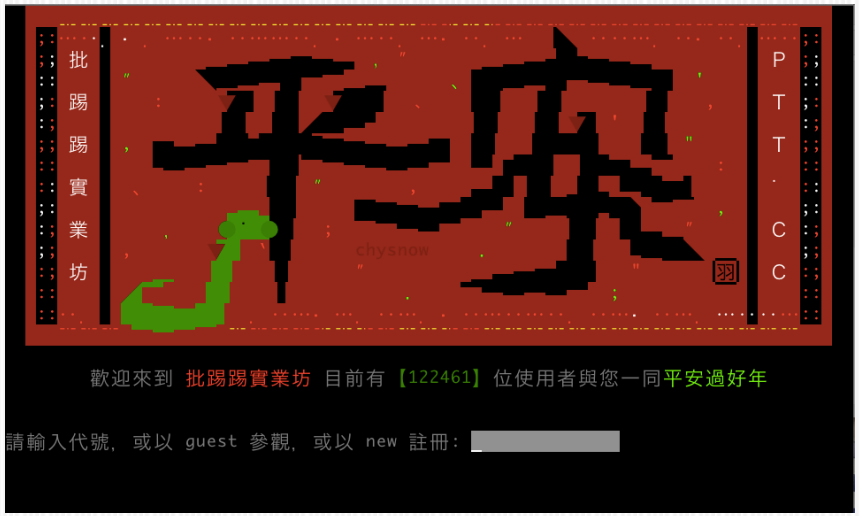
Chinese and English versions of ads and task instructions are prepared for crowdsourcing

MTurk

amazon mechanical turk beta

PTT

A BBS System and Online Community in Taiwan



## Stage 1 : Answer Creation

- 223 workers
  - 122 from MTurk
  - 101 from PTT

## Stage 2 : Answer Aggregation

- 222 workers

## Evaluation

- 165 workers
  98 from Mturk
  67 from PTT

# Key Results

# Interactive vs. Static Interface

- 73.6% of comments show preference for working with the interactive chat bot.
- Increasing the satisfaction of workers (Kittur, A., et al. 2013)

# Interactive vs. Static Interface

"Chat is much fun and more likely to make me think, while questionnaire is more standardized, like an exam."

"the chat interface is much better. it recognizes the text entered in real time and responds accordingly with artificial intelligence and recognition. very nice"

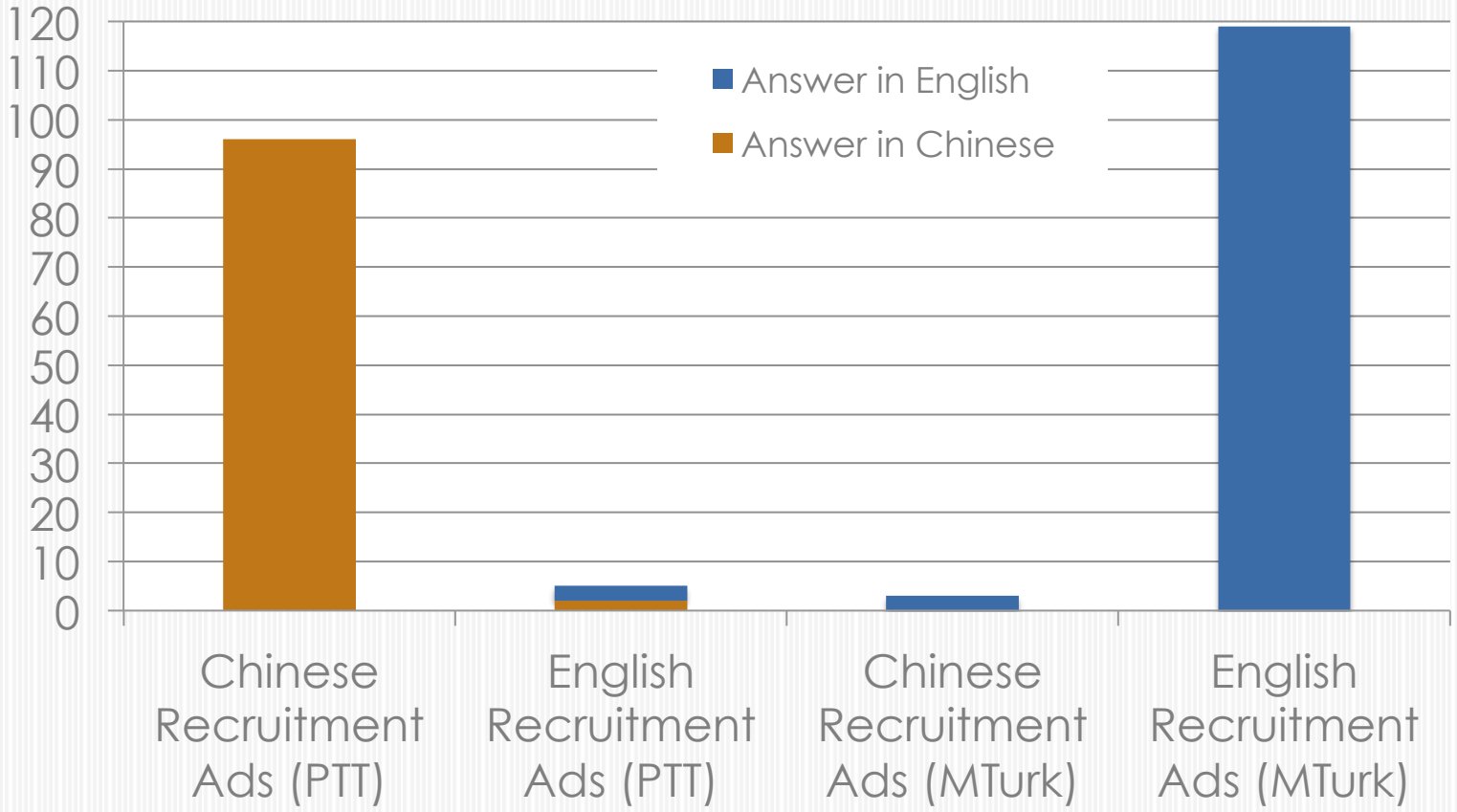
# Interactive vs. Static Interface

- 73.6% of comments show preference for working with the interactive chat bot.

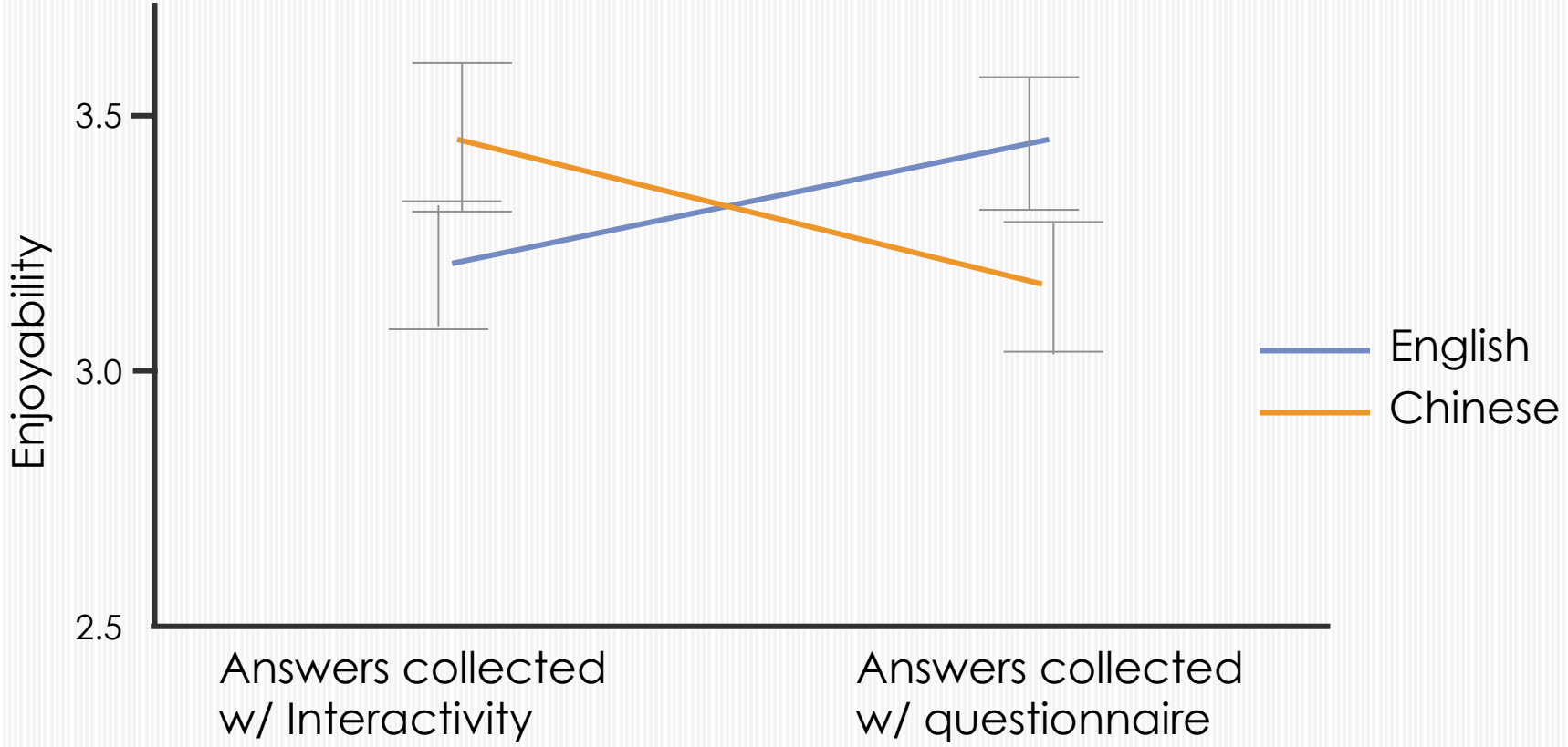
- Increasing the satisfaction of workers (Kittur, A., et al. 2013)

# Mturk vs. PTT : Language

- Two platforms are highly language-specific.



# Evaluation: Ultimate User Experience




- Cultural Differences. (Nisbett, R., 2003 & Hall, E. T.,1977).

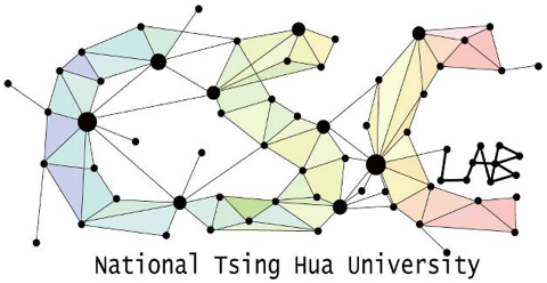
# Conclusion

- Present an interactive chat bot-based interface for crowdsourcing language generation tasks for building natural dialogue agents.
- Interactivity leads to higher worker satisfaction, and better perceived enjoyability by Chinese-speaking users.
- Also, identified language specificity of crowdsourcing platforms. Helps to inform crowdsourcing practices.

# Ultimate Goal? Mind-Connecting!

# Acknowledgement

Ministry of Science and Technology, Taiwan / 科技部
Google / Google
Microsoft Research Asia / 微軟亞洲研究院
Industrial Technology Research Institute (ITRI) / 工業技術研究院




清華大學人機合作與社群運算實驗室
NTHU Collaborative and Social Computing Lab (CSC Lab)