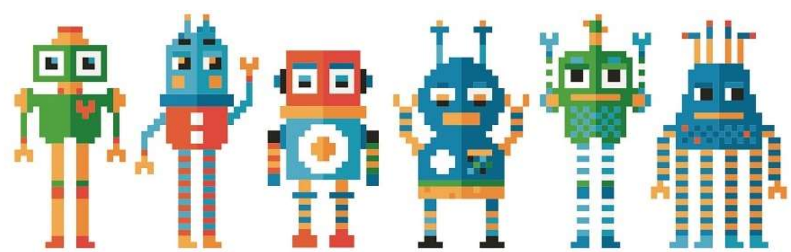
Future Teachers Are Wired Differently: Social Robots in Education

Tony Belpaeme

IDLab – imec, Ghent University, Belgium

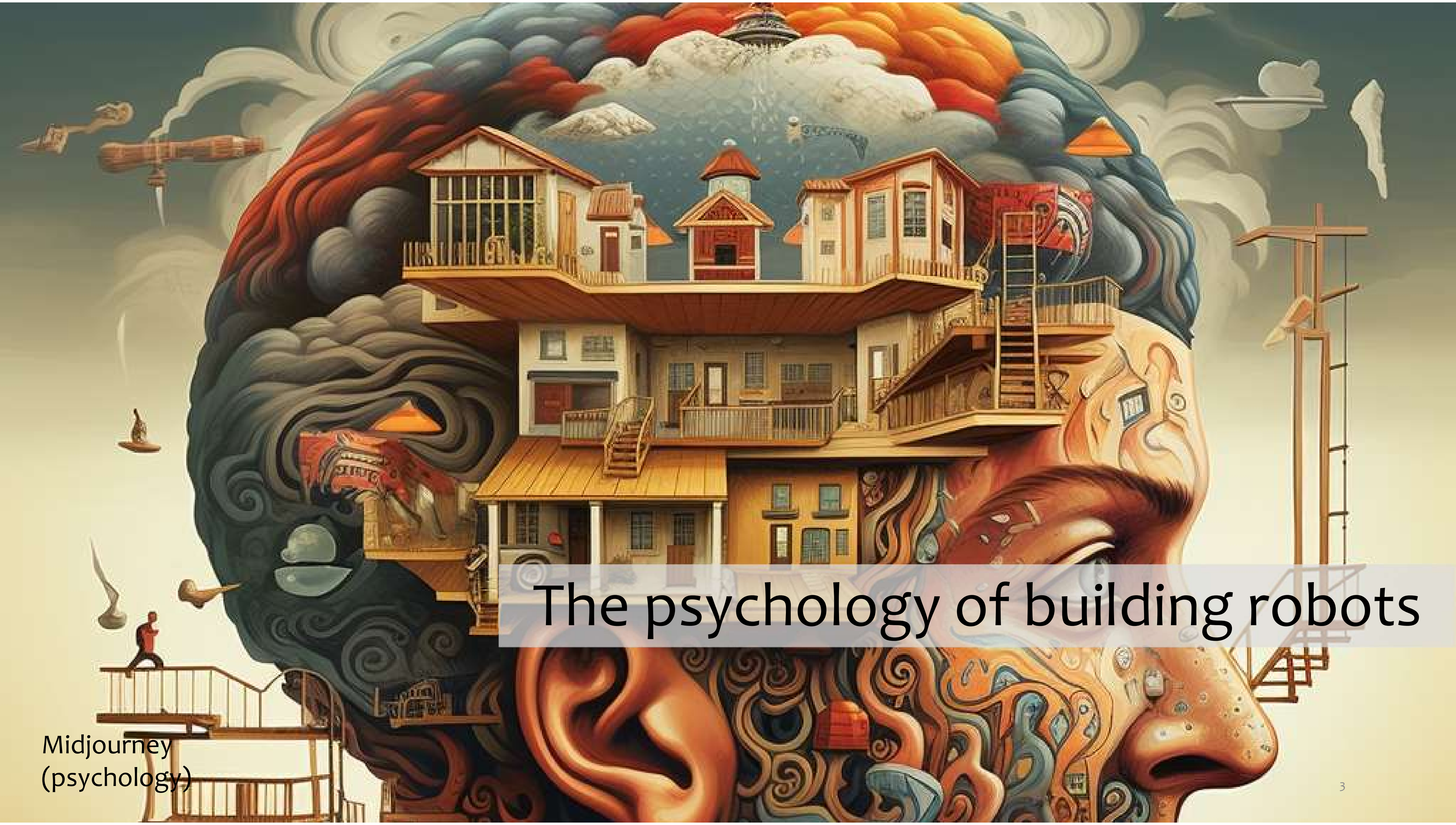


GHENT UNIVERSITY · imec

Today's menu

- The psychology of social robots
- Social robots in education
- Autonomous social robots

The psychology of building robots



Midjourney
(psychology)

Robots tap into our social brain

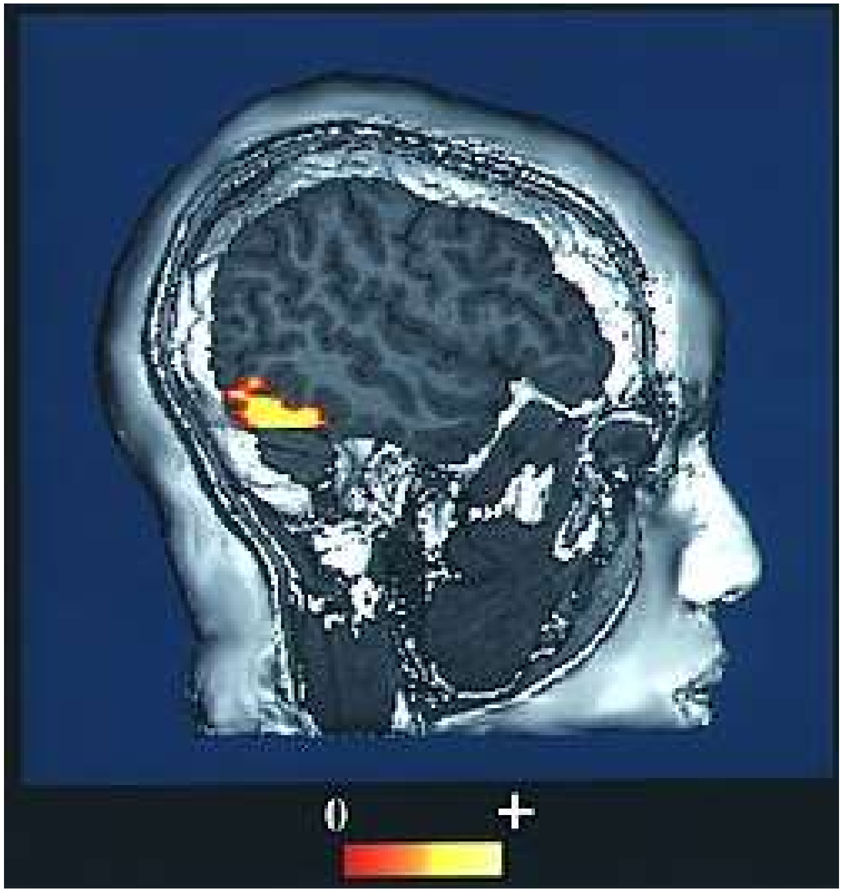
We attribute human-like characteristics to artefacts.

This effect is enhanced when the artefact is animated and responsive.

Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *The American Journal of Psychology*, 243-259.

Pareidolia

- Perceiving human-like features in non-human stimuli.
- Evolutionary psychology explains pareidolia as a hyper response to face-like and human-like features.
- Better to respond to false positives than to not respond.



Fusiform Face Area responds to seeing faces *and* to pareidolia experiences

GHENT UNIVERSITY · imec

(Reddit)

Social robots

Social robots are designed to evoke anthropomorphisation.

But also dynamics, responsiveness, audible responses, linguistic and non-linguistic responses…

This can induce **attention, compliance, conformity** …

University of Washington

Disney Imagineering

GHENT UNIVERSITY · imec

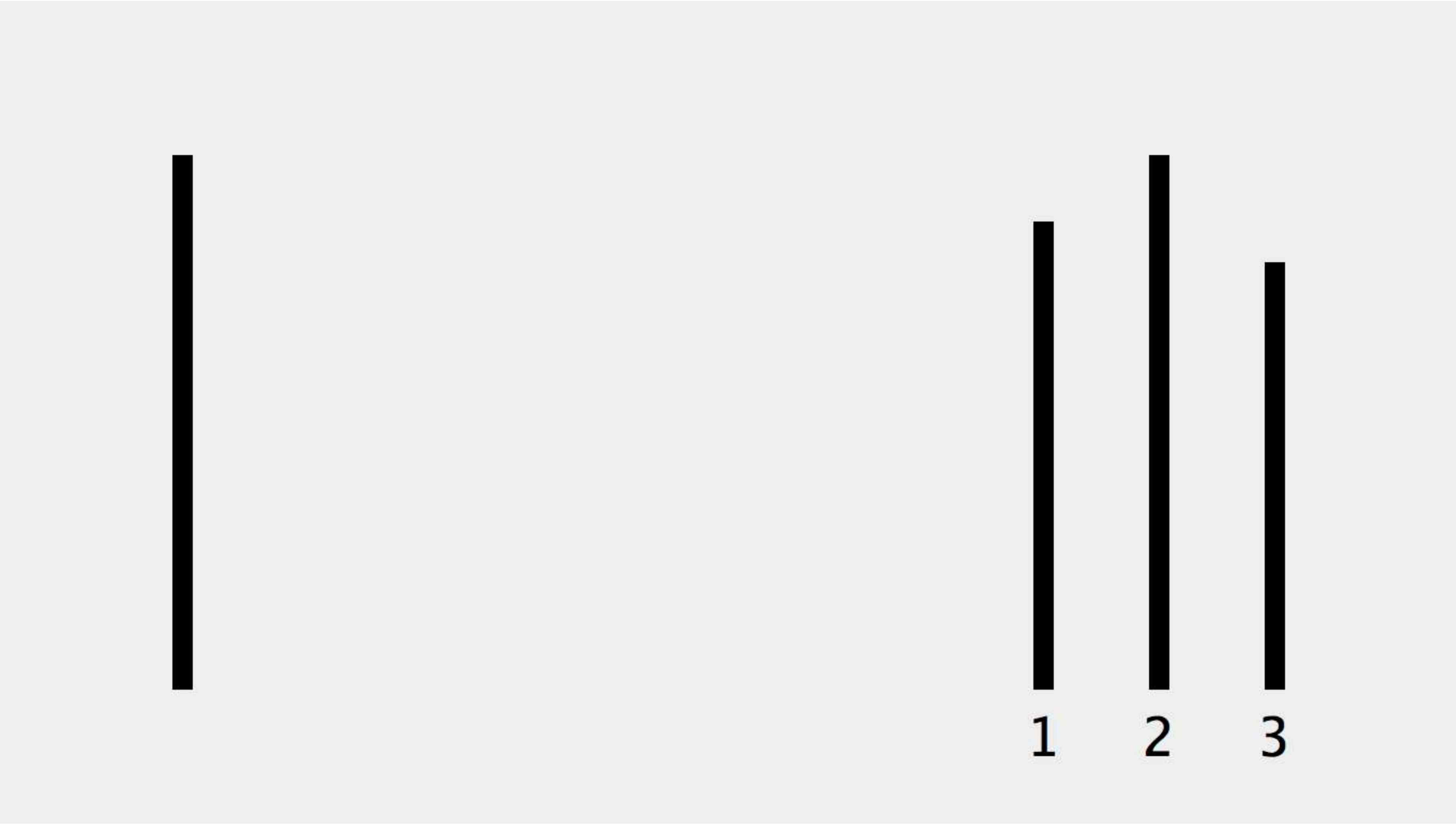
Social conformity

- How much to people conform to social pressure?
- Two kinds of social pressure
  - Informational: uncertainty, so you follow what the others are doing
  - Normative: influence by others
- Solomon Asch's social conformity experiments (1951)

S. E. Asch, Effects of group pressure upon the modification and distortion of judgments, in *Groups, Leadership, and Men*, H. Guetzkow, Ed. (Carnegie Press, 1951), pp. 222–236 .

https://www.simplypsychology.org/asch-conformity.html

GHENT UNIVERSITY · imec

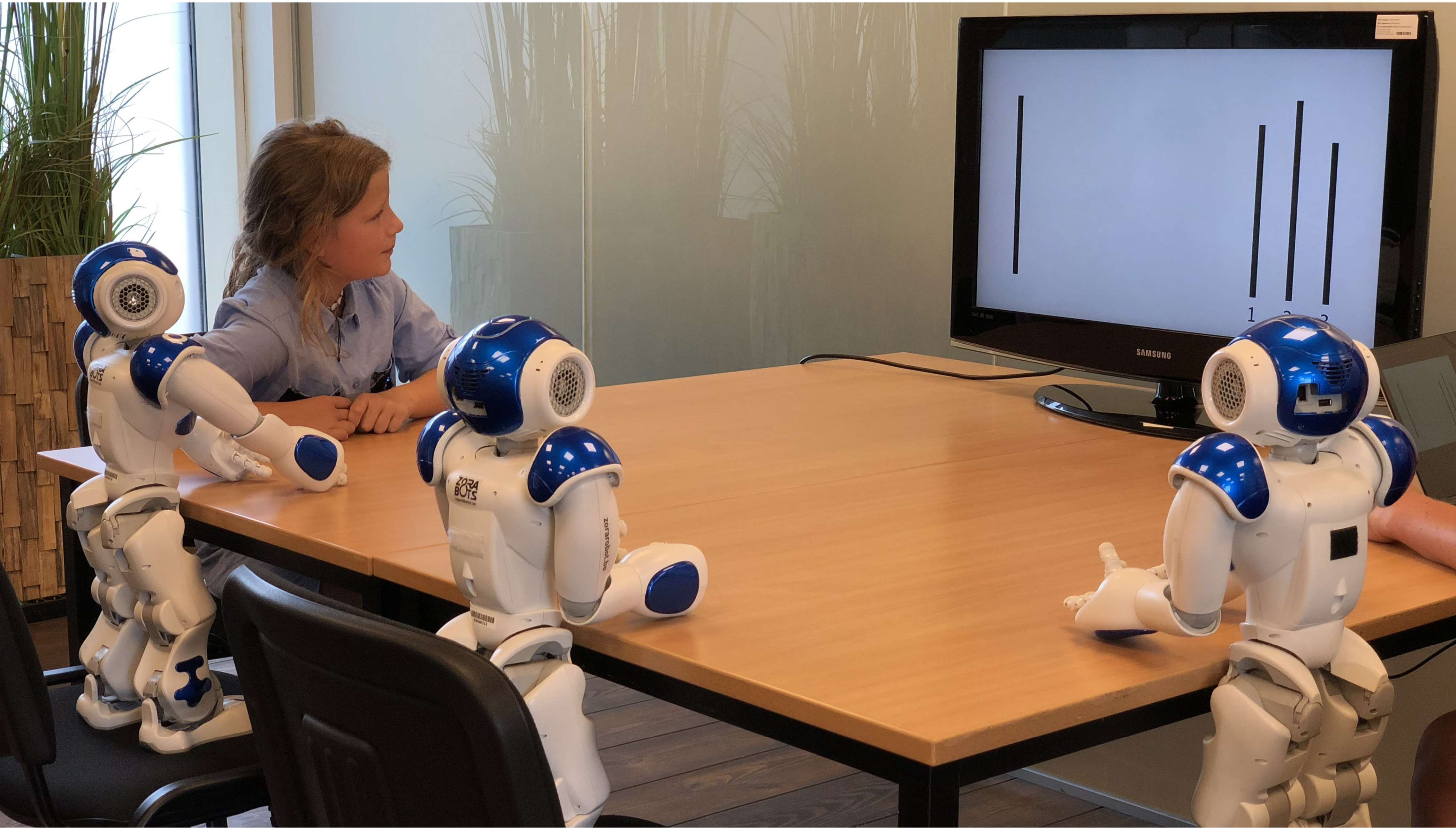

1
2
3

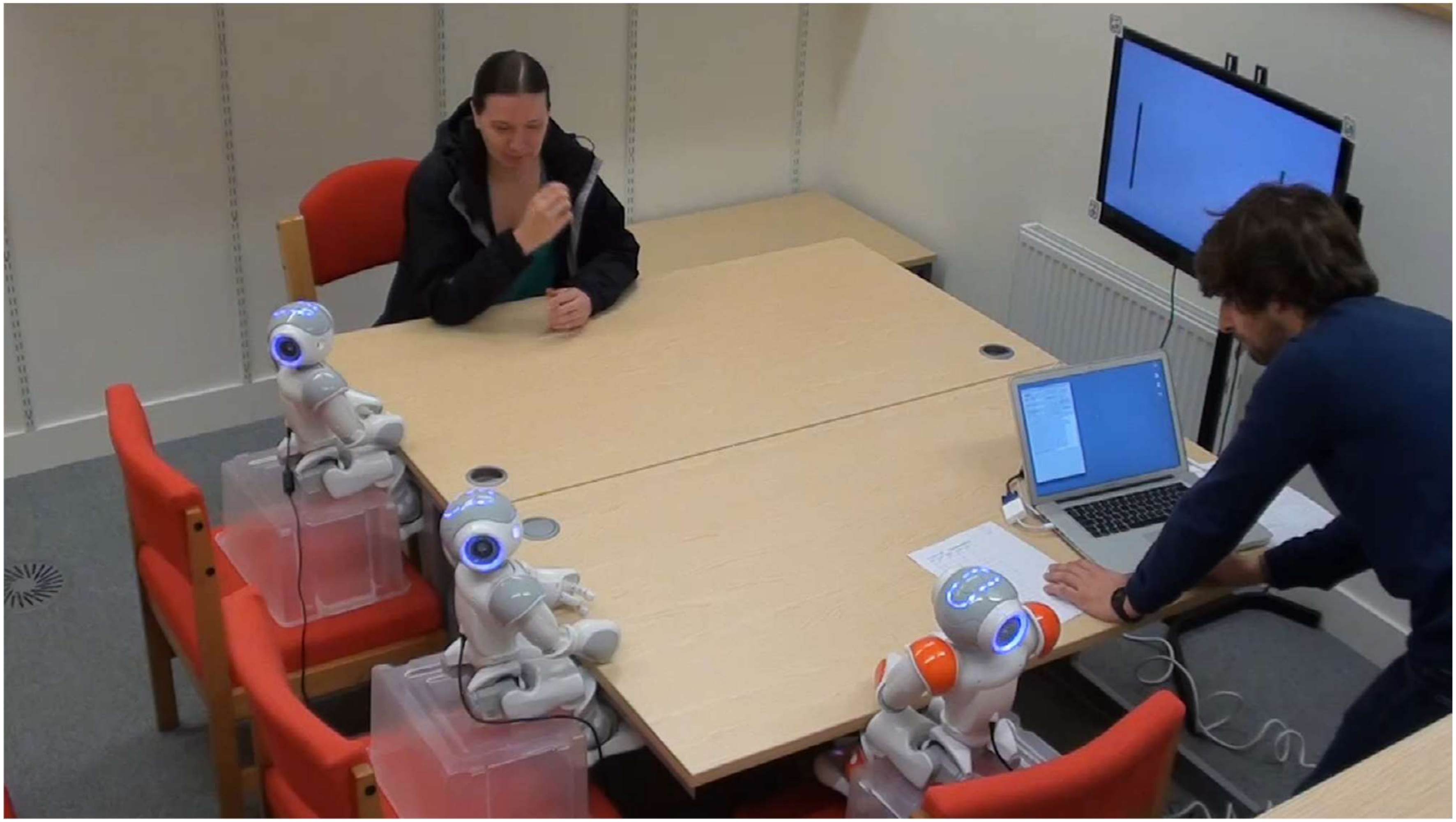
Social conformity

Asch found that people succumb to peer pressure

In **32% of the trials people copied the wrong response.**

Would robots exert social peer pressure as well?


1
SAMSUNG
ZORA BOTS
zorarobot.be



Method

Five experimental conditions

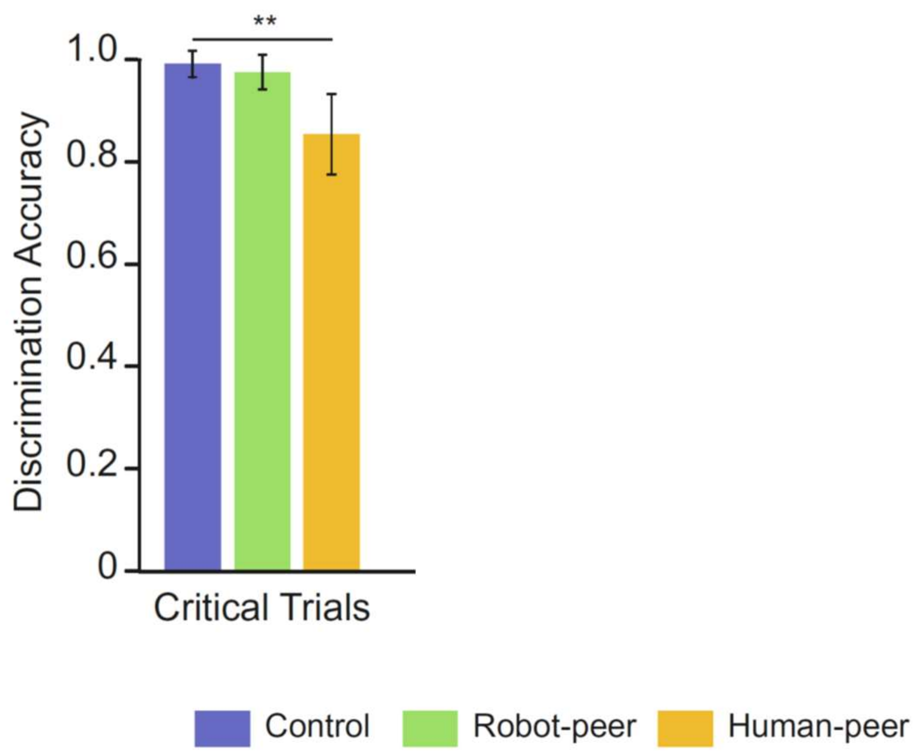
1. Adult alone
2. Adult with robots
3. Adult with other adults
4. Child alone
5. Child with robots
6. ~~Child with other children~~

Participants

- Adults: $M_{age} = 31.0$; $SD = 14.2$
- Children: $M_{age} = 8.5$; $SD = 0.54$

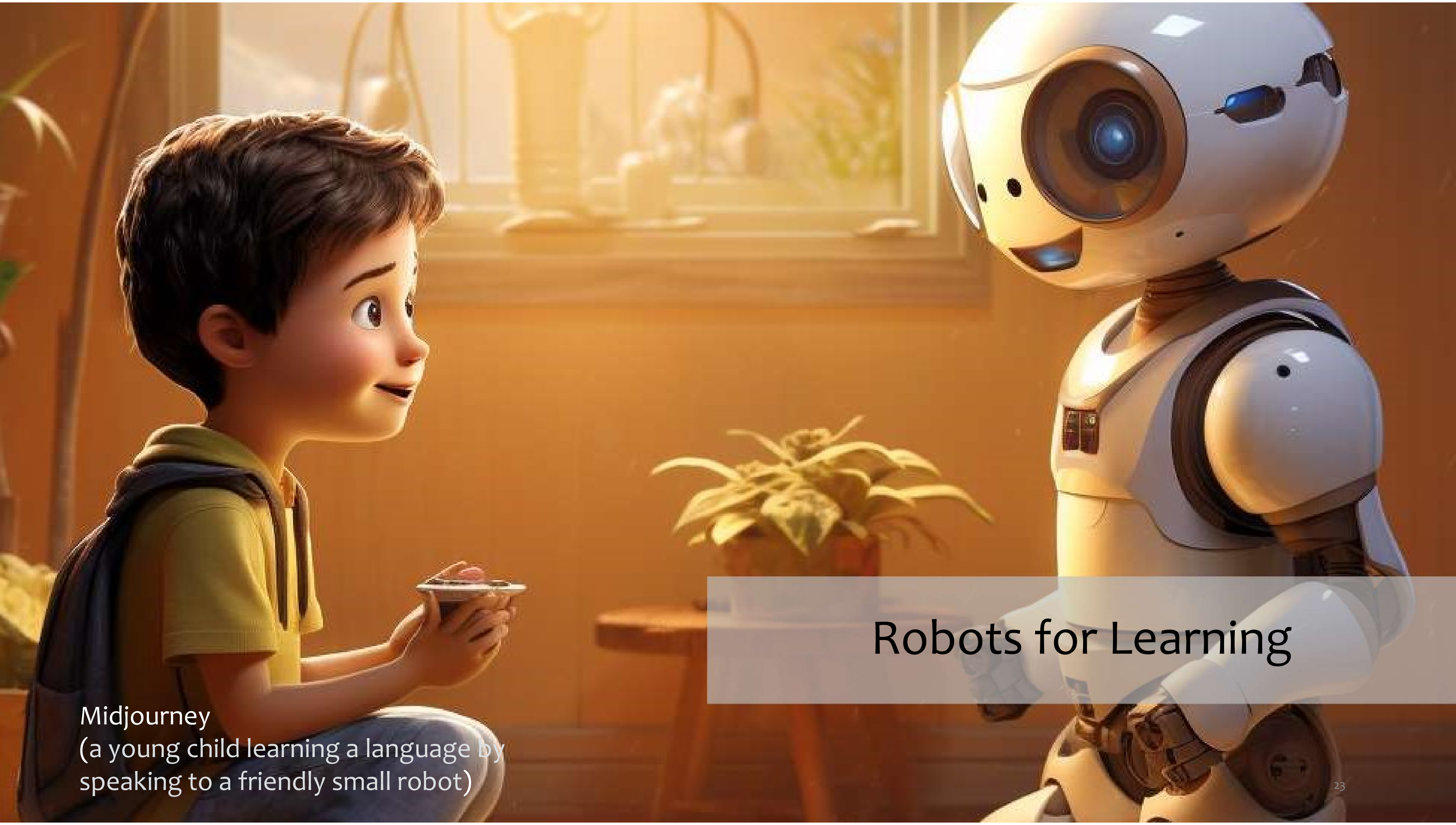
GHENT UNIVERSITY · imec

adults



** p < 0.00001, * p = 0.0009

Anna-Lisa Vollmer, Robin Read, Dries Trippas, and Tony Belpaeme (2018) Children conform, adults resist: A robot group induced peer pressure on normative social conformity, *Science Robotics* , 3(21), eaat7111.



Robots for Learning

Midjourney
(a young child learning a language by
speaking to a friendly small robot)

Why robots?

Why indeed?

- Robots are expensive
- Robots are difficult to use at scale
- We already have enough technology (PC, tablets, smartphones, …)

Good reasons

- Learning is **socially gated** (see Patricia Kuhl).
- Their physical and social presence makes robot tutors effective.
- Social robots can offer one-to-one tuition.
- They can achieve both cognitive and affective outcomes.
- Potential benefits for neurodiverse learners.

Belpaeme, T., Kennedy, J., Ramachandran, A., Scassellati, B., & Tanaka, F. (2018). Social robots for education: A review. *Science Robotics*, *3*(21), eaat5954.

The Journal of Genetic Psychology, 153(3), 269–273

Using a Personal Robot to Teach Young Children

THOMAS W. DRAPER
WANDA W. CLAYTON
Early Childhood Education Laboratory
Brigham Young University

ABSTRACT. Seventy-five preschool children were instructed about birds by a human teacher, a moving personal robot, a stationary personal robot, and a tape recorder. How much the children learned and how much attention the children paid were compared for each type of instruction. The children learned when they were taught by the human teacher and when they were taught by the animated and the stationary robots. The children paid more attention to the live teacher and to the moving robot than they did to the stationary robot or to the tape recorder. The difference between the amount of attention the children paid to the animated robot and the amount of attention they paid to the human teacher was not statistically significant.

CHILDREN YOUNGER THAN 5 YEARS OF AGE learn differently from children older than 7 years of age. Young children's learning is more dependent on concrete, three-dimensional illustrations and social factors, such as personal liking of the teacher and animate teaching style, than is the learning of older children (Bredekamp, 1987; Taylor, 1985). Because of the special needs of young learners, many experienced teachers and early childhood education experts have expressed doubts about the value of computer-controlled presentations for teaching young children. Much of the high technology used with young children has been labeled developmentally inappropriate because it consists of messages requiring relatively sophisticated symbolic reasoning, and hypothetical problem solving (Barnes & Hill, 1983; Brady & Hill, 1984; Cuffaro, 1984; Haugland & Shade, 1988; Tan, 1985; Zajonc, 1984).

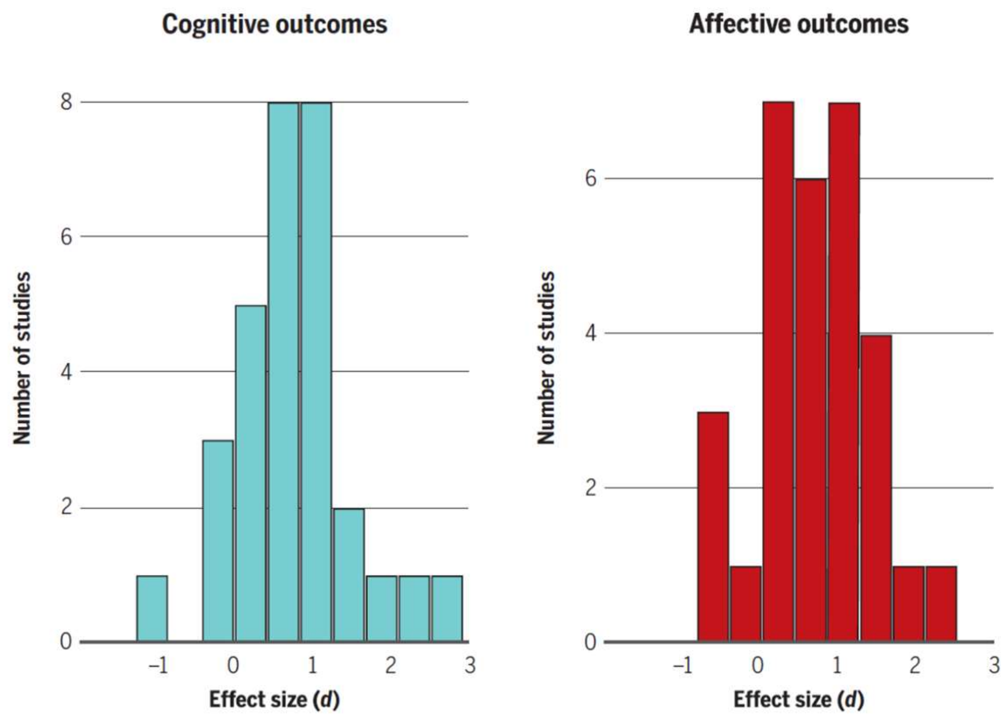
In the present study, we attempted to use technology to teach young children in a manner that is more consistent with the tenets of developmental

Address correspondence to Thomas W. Draper, Early Childhood Education Laboratory, Brigham Young University, Provo, UT 84602.

GHENT UNIVERSITY · imec

Effect sizes of outcomes

- 37 results compared a robot to alternative tech tutoring.
- **Cognitive *d* = 0.70**
- **Affective *d* = 0.59**
- Human tutor achieve cognitive outcomes of d = 0.79
- Positive affective outcomes do not mean positive cognitive outcomes, or vice versa.



Belpaeme, T., Kennedy, J., Ramachandran, A., Scassellati, B., & Tanaka, F. (2018). Social robots for education: A review. *Science Robotics*, 3(21), eaat5954.

Second language learning

- Current classroom setup is ill-suited for language tuition
  - “Broadcast mode” of education doesn’t fit how people acquire first and second languages.
  - Social interaction is important for language acquisition.

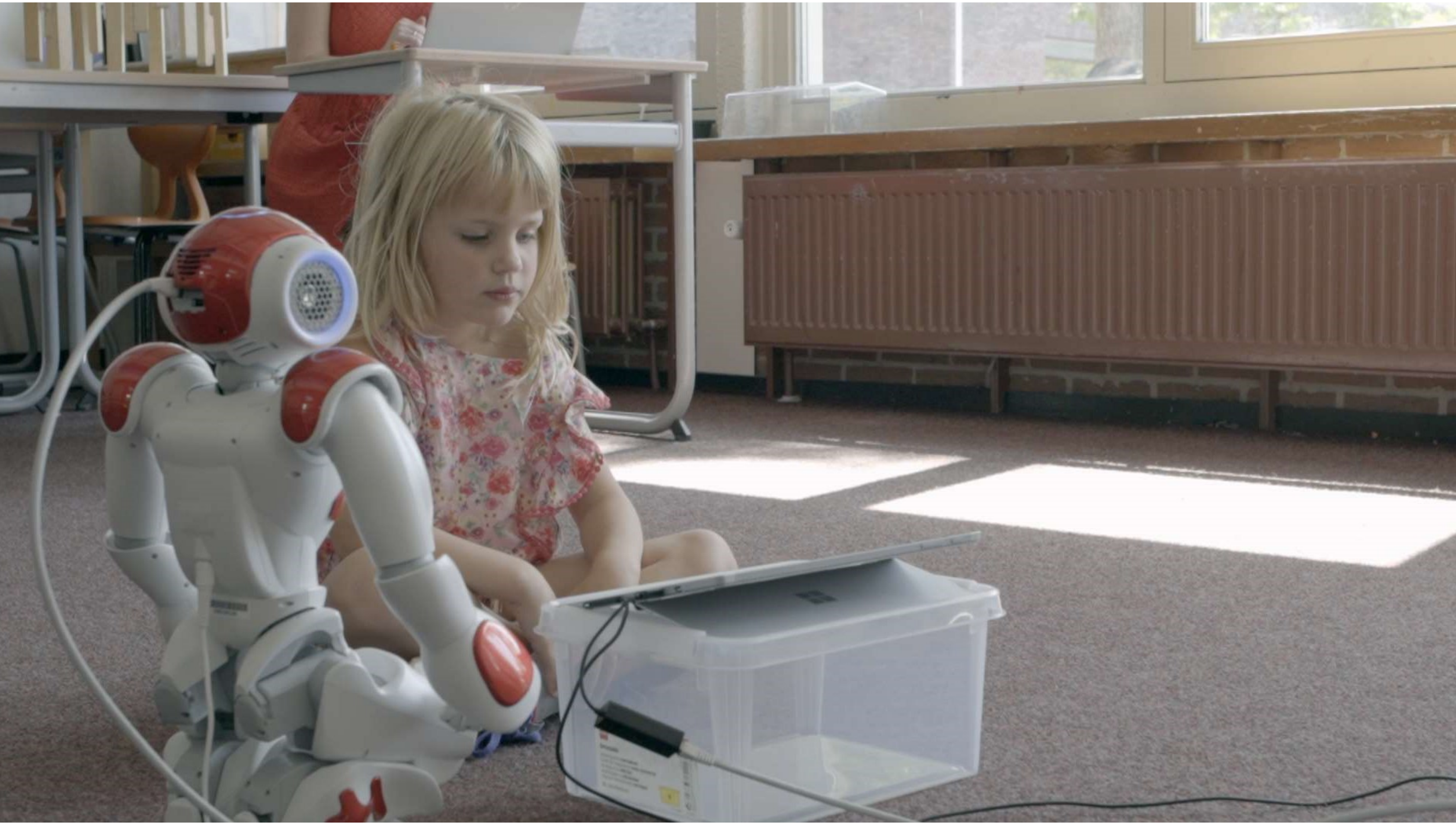
- Benefits of social robots
  - Interaction possible
  - One-to-one tuition
  - Pronunciation
  - Foreign language anxiety

Kuhl, P. K. (2010). Brain mechanisms in early language acquisition. Neuron, 67(5), 713-727.
Snow, C. E., & Hoefnagel-Höhle, M. (1978). The critical period for language acquisition: Evidence from second language learning. *Child development*, 1114-1128.

Possibly the greatest challenge of all

- **Vocabulary learning works**, beyond that things become very complicated.
- **Conflicting age demands.** Critical Period Hypothesis: learn a language before puberty.
- **Technical challenges** prevent a dyadic conversation with the robot.

van den Berghe, R., Verhagen, J., Oudgenoeg-Paz, O., van der Ven, S., & Leseman, P. (2018). Social Robots for Language Learning: A Review. *Review of Educational Research*, 0034654318821286.



Video available at www.l2tor.eu

L2TOR is a European project that investigates how preschool children can learn a second language from a social robot.

Large-scale study

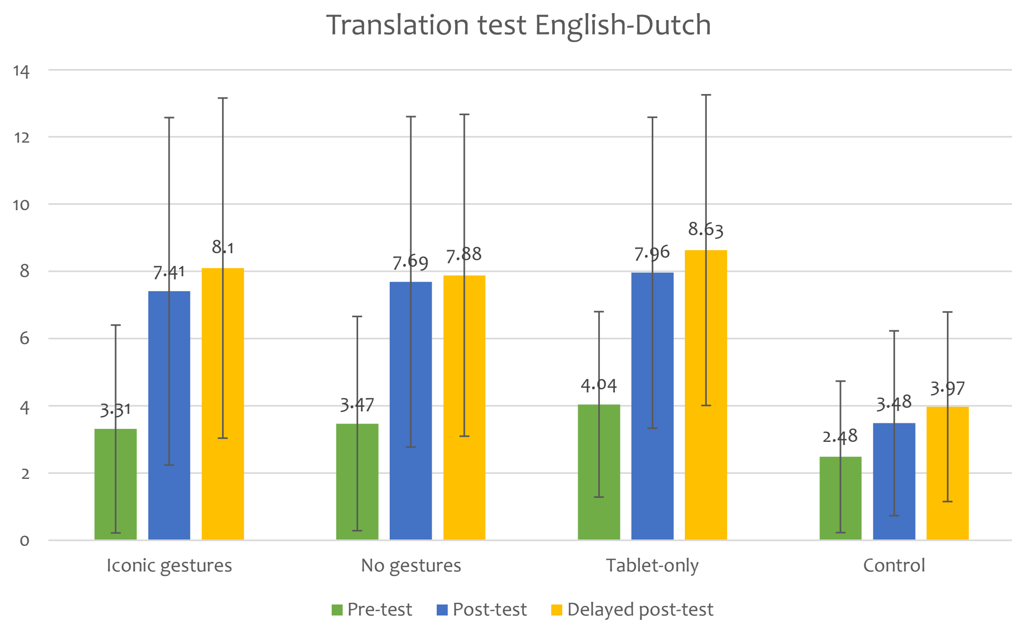
- **208 children** in 9 Dutch primary schools, 194 completed.
- Mean age = 5;8 (SD = 0;5), 34 English target words, embedded in grammatical context.
- 1 introduction, 7 learning sessions. **Two months per child.**
- Challenging words: space, number, actions
- Four conditions
  1. Robot supports teaching with iconic gestures.
  2. Robot does not use gestures.
  3. No robot, only tablet.
  4. Control conditions.

GHENT UNIVERSITY · imec

Vogt, P., van den Berghe, R., de Haas, M., Hoffman, L., Kanero, J., Mamus, E., ... & Papadopoulos, F. (2019, March). Second Language Tutoring Using Social Robots: A Large-Scale Study. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (pp. 497-505). IEEE.

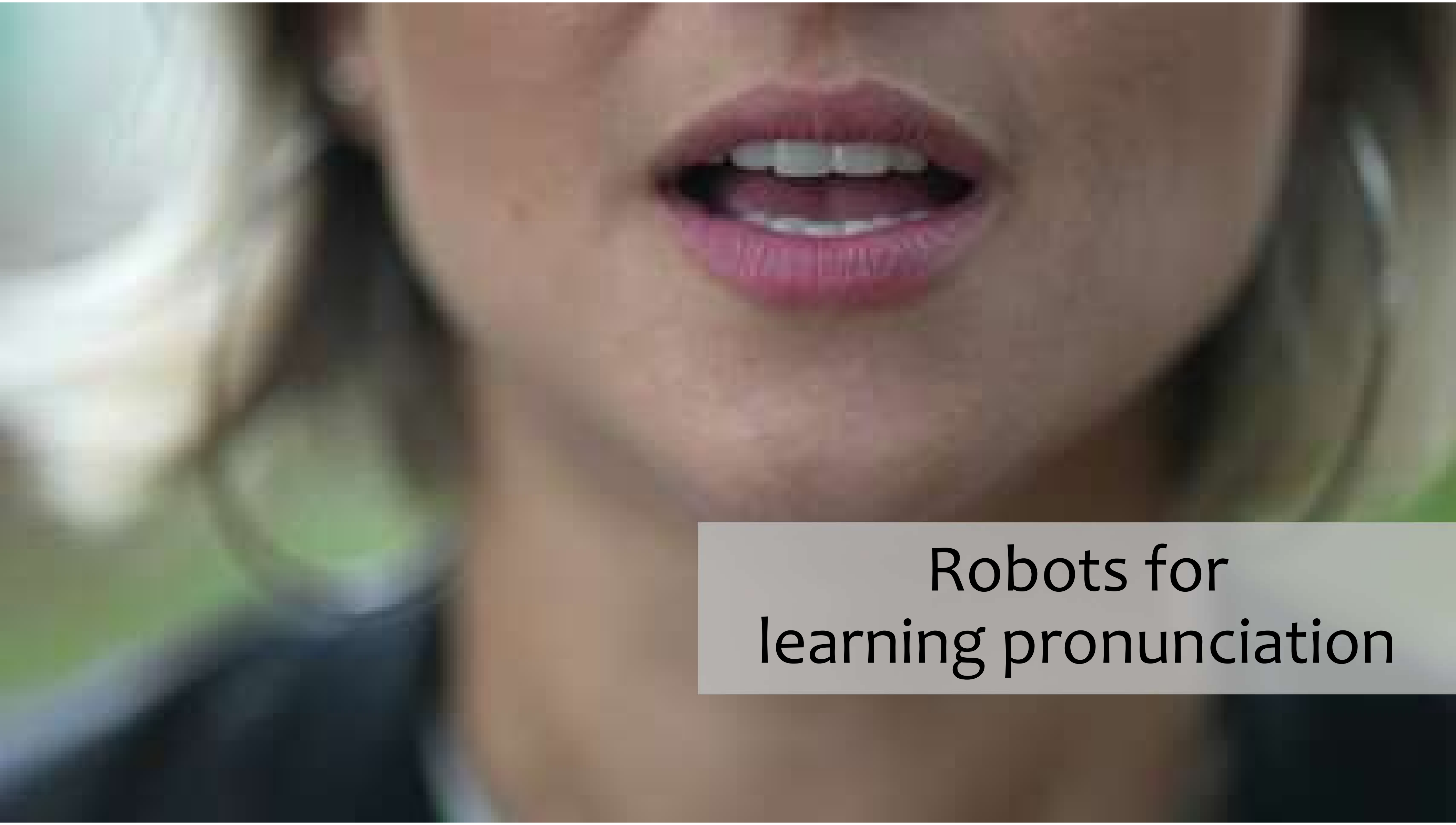
Results

- Learning is surprisingly slow.
- Children learn in all conditions, except in control condition: **no advantage for robot teaching**



GHENT UNIVERSITY · imec

Vogt, P., van den Berghe, R., de Haas, M., Hoffman, L., Kanero, J., Mamus, E., ... & Papadopoulos, F. (2019, March). Second Language Tutoring Using Social Robots: A Large-Scale Study. In *2019 14th ACM/IEEE Int. C. on Human-Robot Interaction (HRI)* (pp. 497-505). IEEE.


Robots for
learning pronunciation

Audiovisual speech

- Audiovisual speech perception augments L1 language perception.
- Some evidence that L2 learners also benefit from seeing lips.
- **Hypothesis: pronunciation will improve when learning with a robot that uses audiovisual speech.**

GHENT UNIVERSITY · imec

Amioka, S., Janssens, R., Wolfert, P., Ren, Q., Pinto Bernal, M.J., Belpaeme, T. (2023) Limitations of Audiovisual Speech on Robots for Second Language Pronunciation Learning. *HRI2023*.

Methodology

- Furhat robot, with motion captured audiovisual animation
- Robot teaches pronunciation of 30 Japanese words.
- 3 conditions
  - Matched lip synching
  - Mismatched lip synching
  - Computer generated

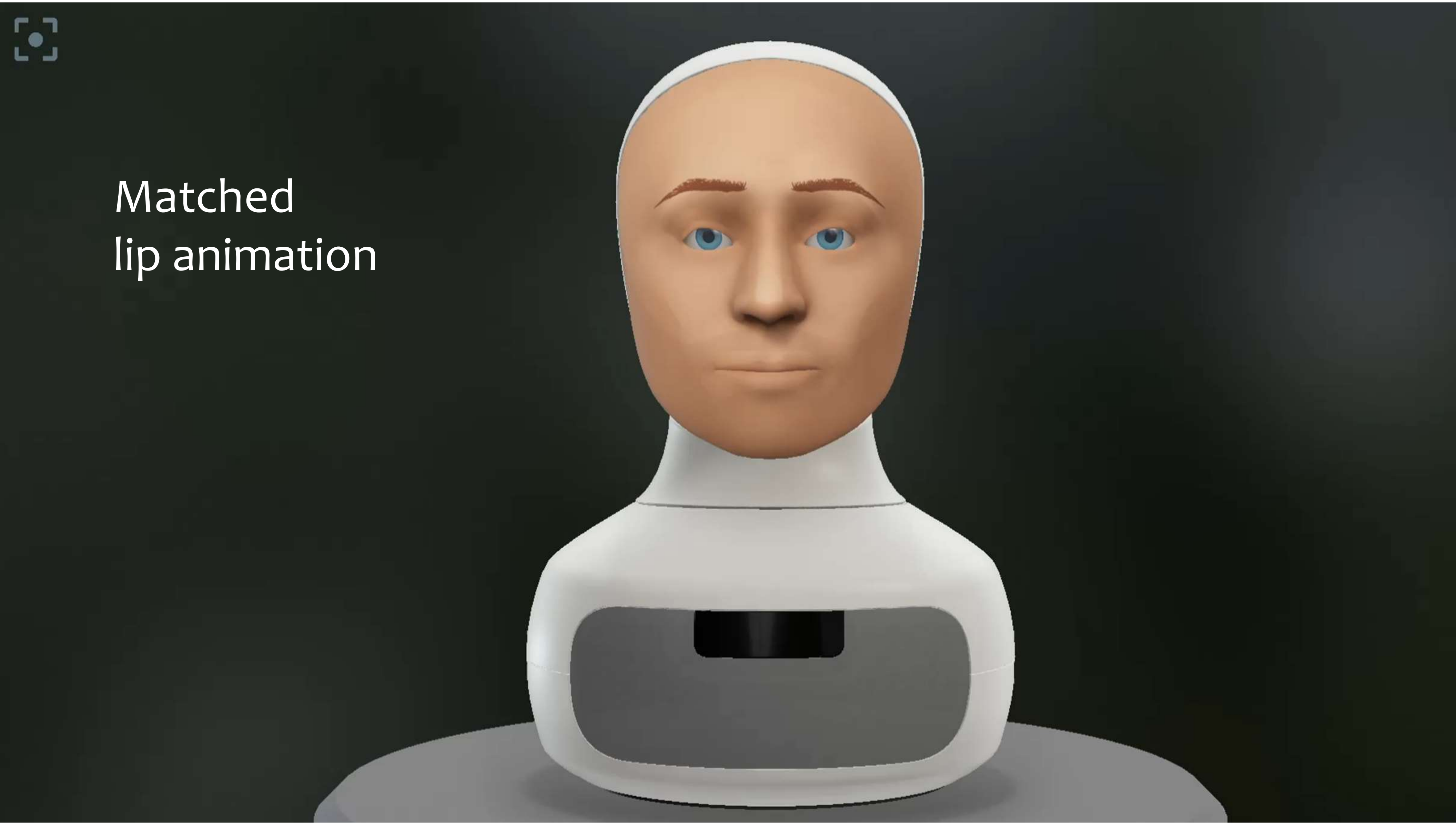
Methodology

- 27 participants recruited in Flanders (native Dutch speakers)
- Within subject design: all participants see all 3 conditions
- Their pronunciation is recorded and sent to Japan: scored by 20 native Japanese speakers from Tokyo region

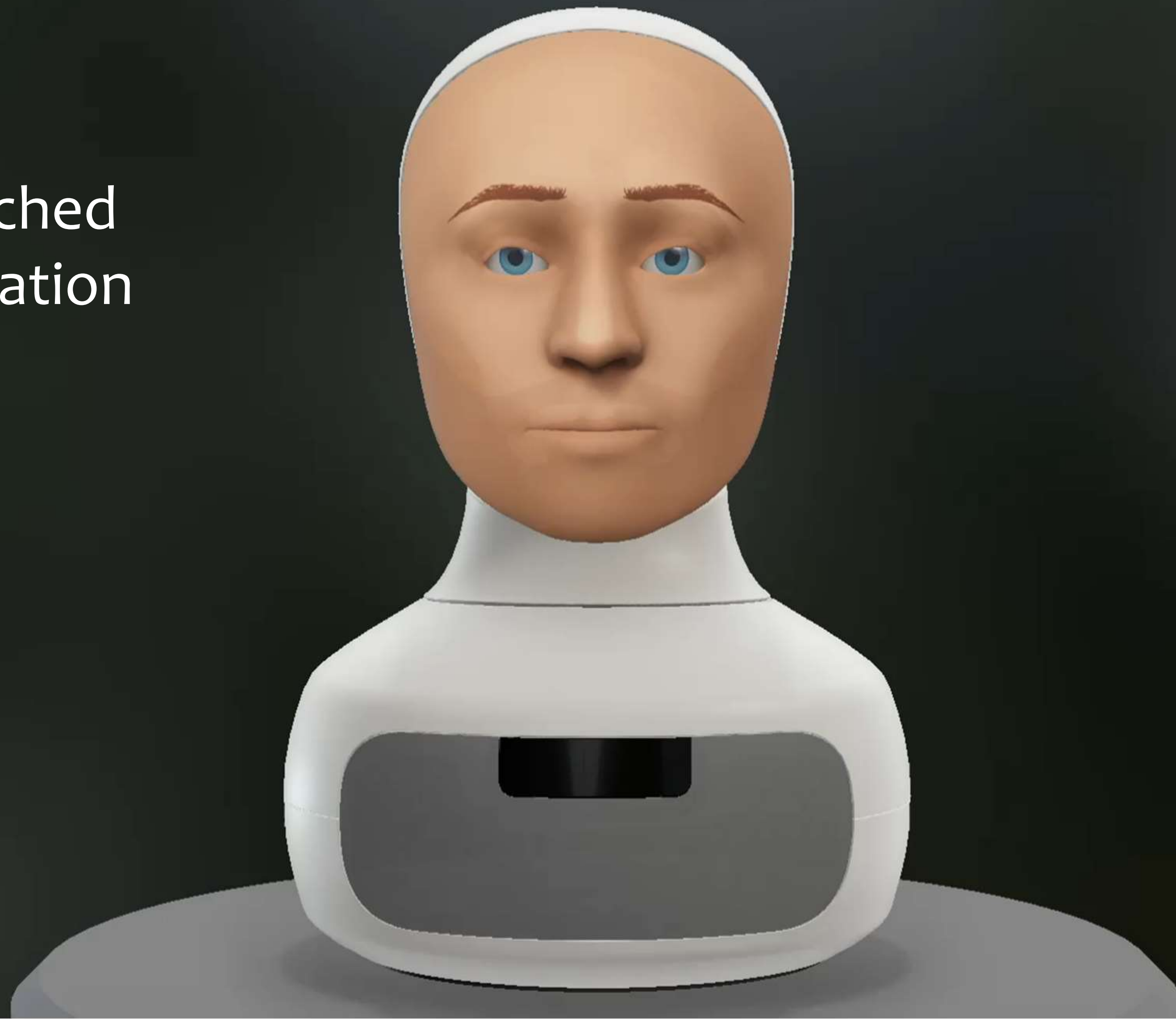
Japanese	English
いち	one
ご	five
しち	seven
おはようございます	good morning
おやすみなさい	good night
ありがとうございます	thank you very much
どういたしまして	you're welcome
つたえられなかった	I couldn't tell you
かぜぐすり	cold medicine
りょこう	travel

GHENT UNIVERSITY · imec


MAAK
DAT MEE

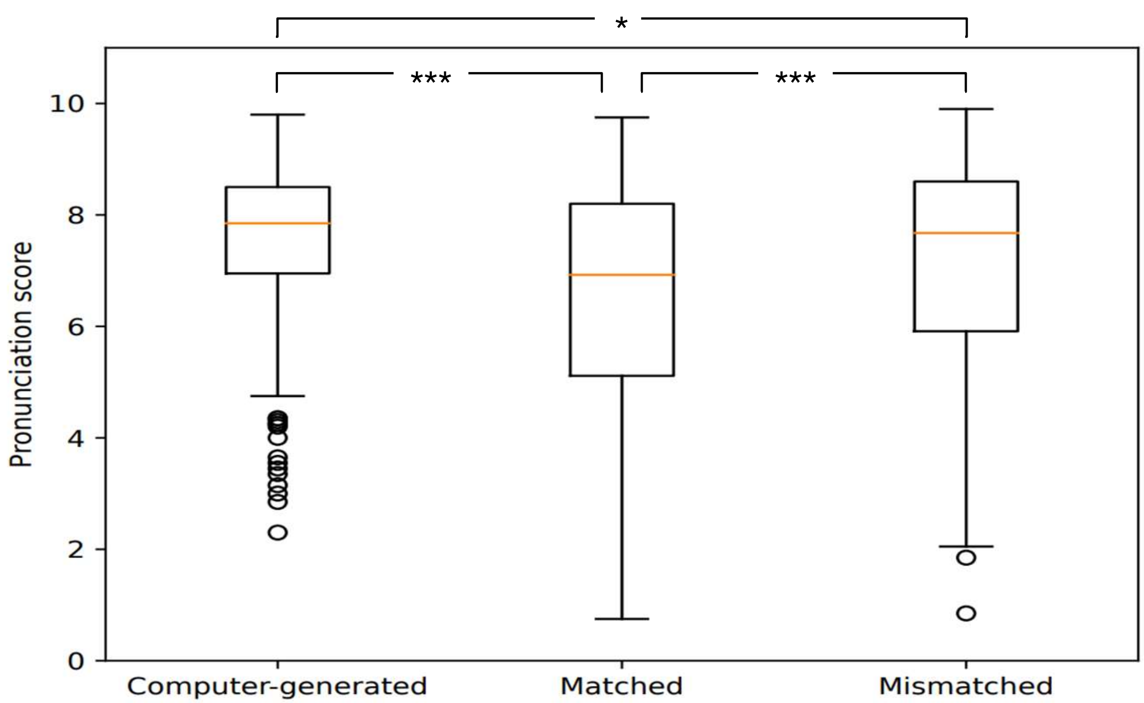

Matched
lip animation

Mismatched
lip animation



Results

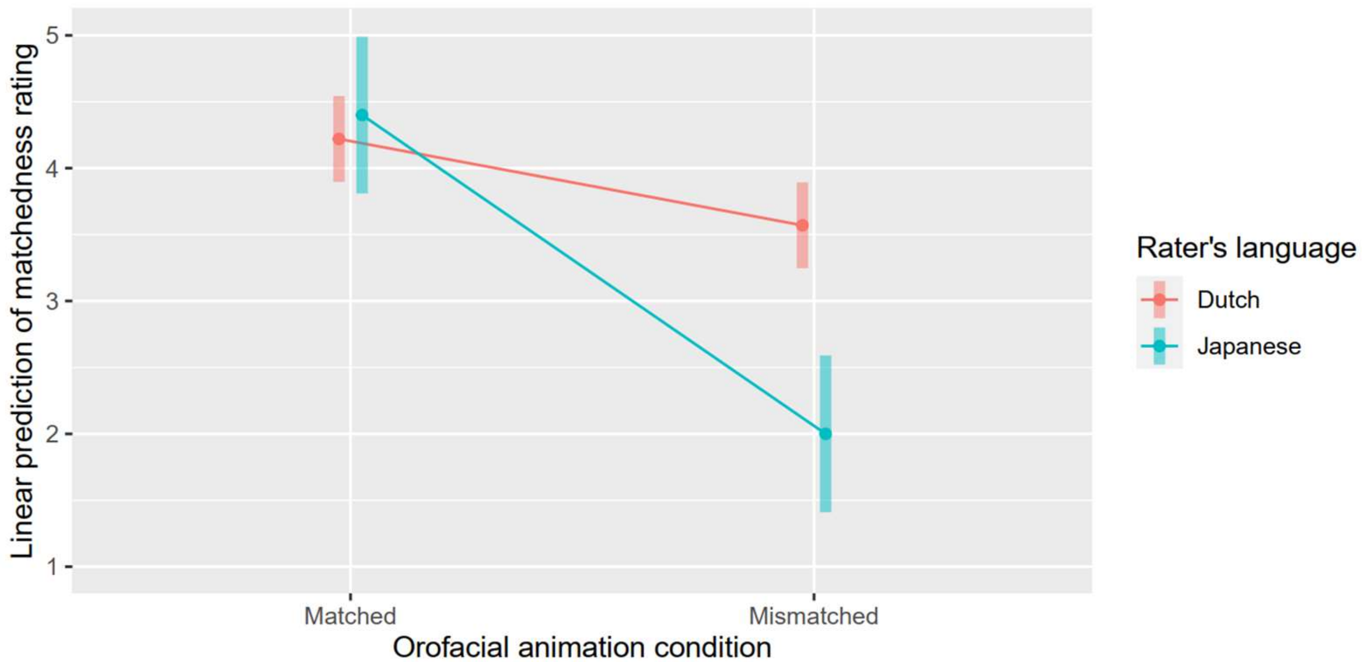
Audiovisual speech does not aid pronunciation learning 😳



*	p < 0,05
**	p < 0,01
***	p < 0,001

GHENT UNIVERSITY · imec

What is going on?



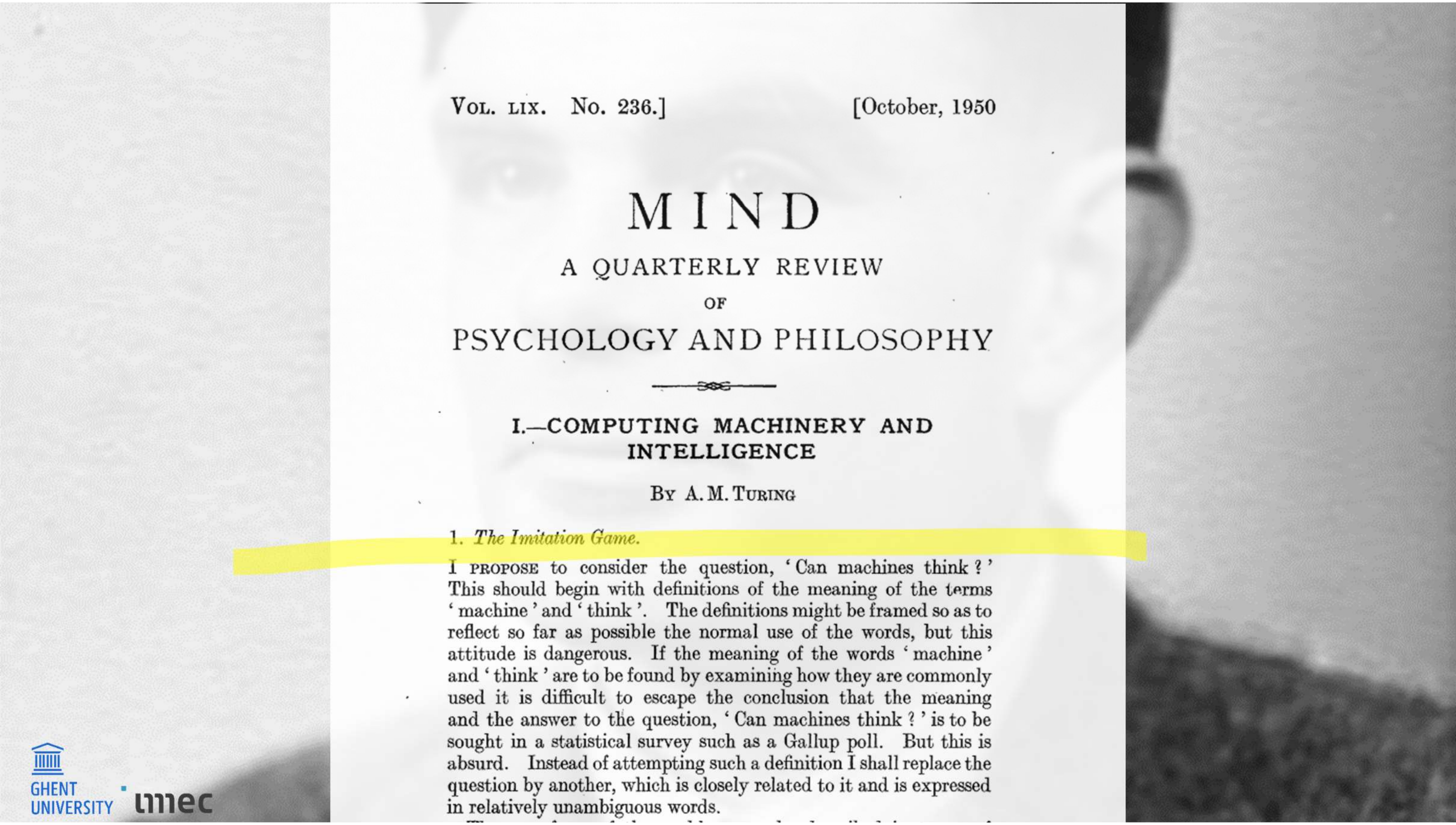
Our participants did not notice the mismatched audiovisual speech.

Orofacial cues do not help learn a language that you do not master yet!

GHENT UNIVERSITY · imec

Building social robots

Midjourney
(a group of social robots)



VOL. LIX. No. 236.] [October, 1950

MIND

A QUARTERLY REVIEW OF PSYCHOLOGY AND PHILOSOPHY

I.—COMPUTING MACHINERY AND INTELLIGENCE

BY A. M. TURING

1. *The Imitation Game.*

I PROPOSE to consider the question, 'Can machines think?' This should begin with definitions of the meaning of the terms 'machine' and 'think'. The definitions might be framed so as to reflect so far as possible the normal use of the words, but this attitude is dangerous. If the meaning of the words 'machine' and 'think' are to be found by examining how they are commonly used it is difficult to escape the conclusion that the meaning and the answer to the question, 'Can machines think?' is to be sought in a statistical survey such as a Gallup poll. But this is absurd. Instead of attempting such a definition I shall replace the question by another, which is closely related to it and is expressed in relatively unambiguous words.

GHENT UNIVERSITY · imec

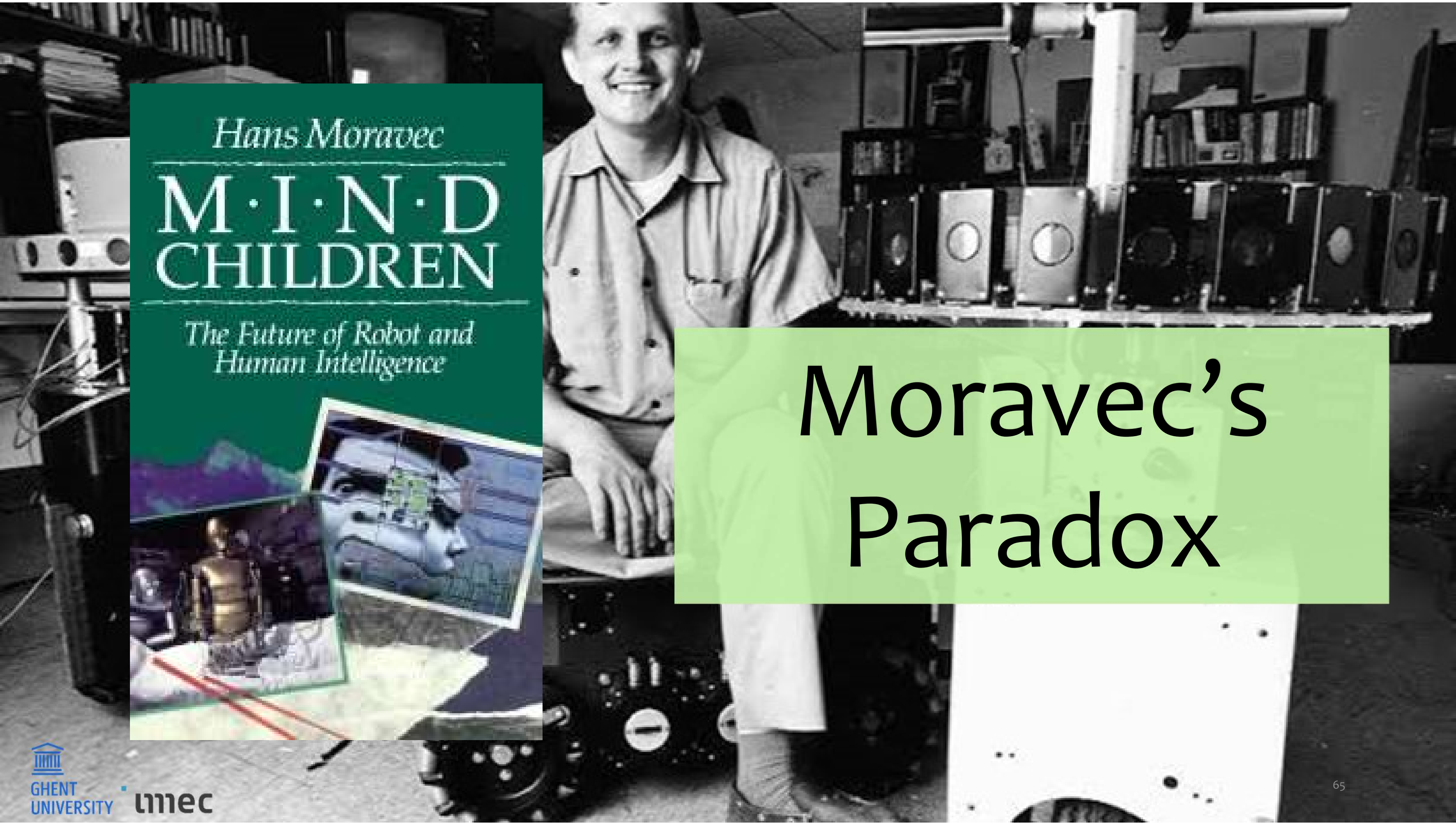
Out with the Wizard

Wizard of Oz approaches still amount for **50% of all research.**

Useful for a quick and cheap study or as a stub for underperforming technology.

But the goal is **autonomous human-robot interaction**

GHENT UNIVERSITY · imec



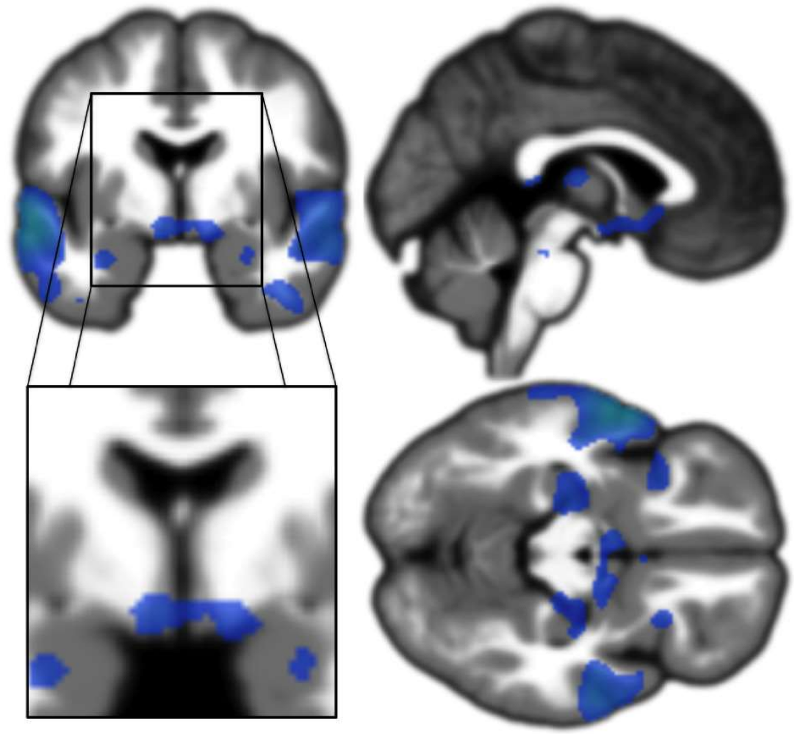
Moravec's Paradox

Building autonomous social robots is hard

Social interaction is **brain complete**.

All cognitive faculties are active and coordinated during social interaction.

Artificial social cognition needs artificial equivalents of these cognitive faculties and their interaction.



Rauchbauer, B., Nazarian, B., Bourhis, M., Ochs, M., Prévot, L., & Chaminade, T. (2019). Brain activity during reciprocal social interaction investigated using conversational robots as control condition. *Philosophical Transactions of the Royal Society B*, *374*(1771), 20180033.

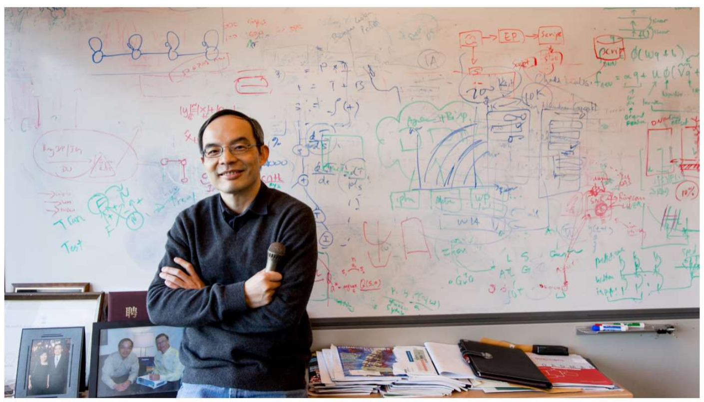
GHENT UNIVERSITY · imec

Speech recognition: super human?

Microsoft claims new speech recognition record, achieving a super-human 5.1% error rate

BY TODD BISHOP on August 20, 2017 at 7:44 pm

5 Comments | Share 820 | Tweet | Share | Reddit | Email



Xuedong Huang, a Microsoft Technical Fellow in AI and Research, leads Microsoft's Speech and Language Group. (Microsoft Photo)

IBM speech recognition is on the verge of super-human accuracy

Chris Weller
Mar. 9, 2017, 4:58 PM 1,494

FACEBOOK | LINKEDIN | TWITTER | EMAIL | PRINT

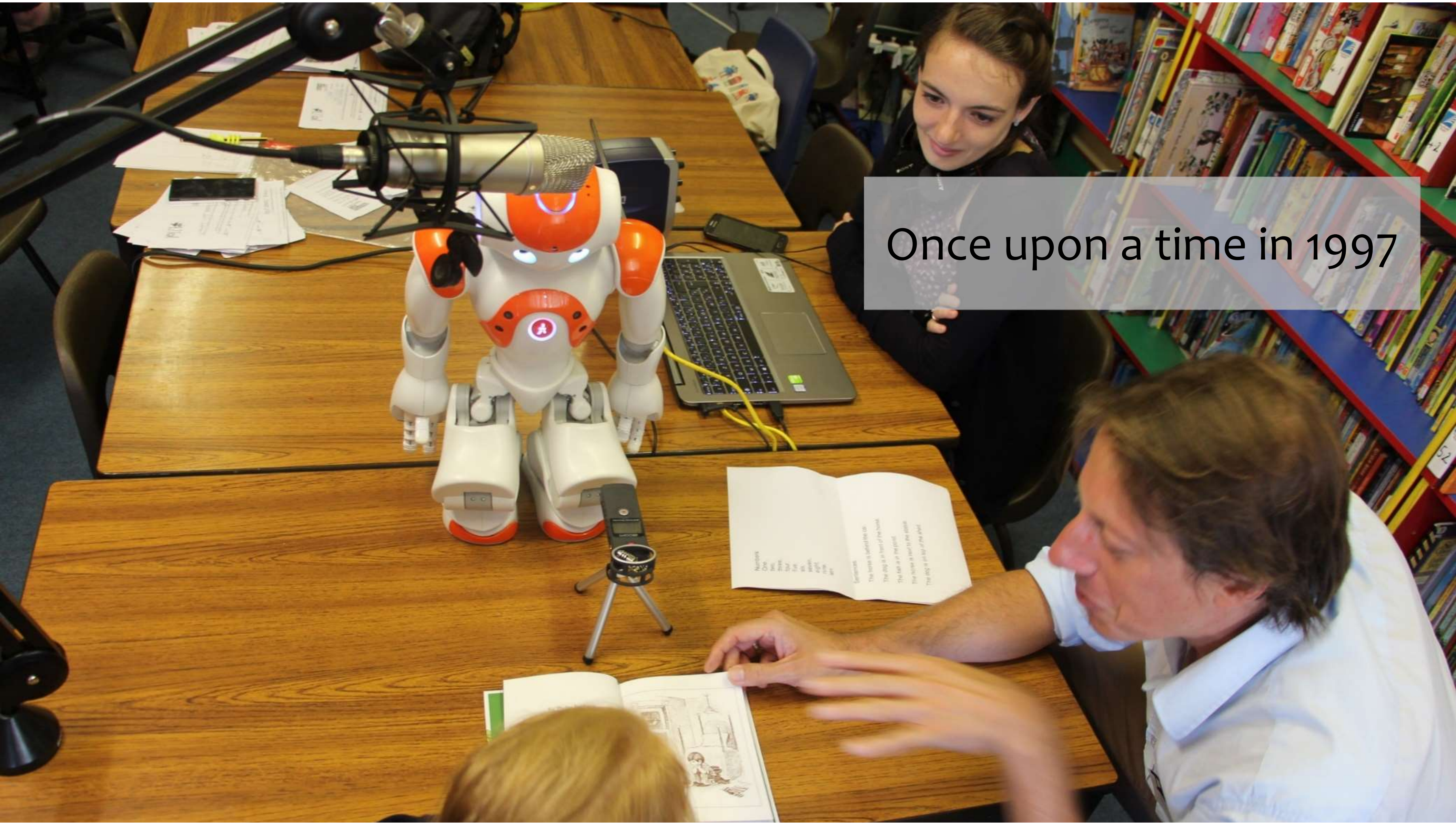
In the world of speech recognition software, 5.1% is kind of a magic number.

Companies that can create software with error rates falling in that ballpark are essentially matching the capabilities of humans, who miss roughly 5% of the words in a given conversation.

ASR for atypical populations

Youtube: [Amazon Alexa Gone Wild!!!]

Still problematic for minorities, especially for atypical populations

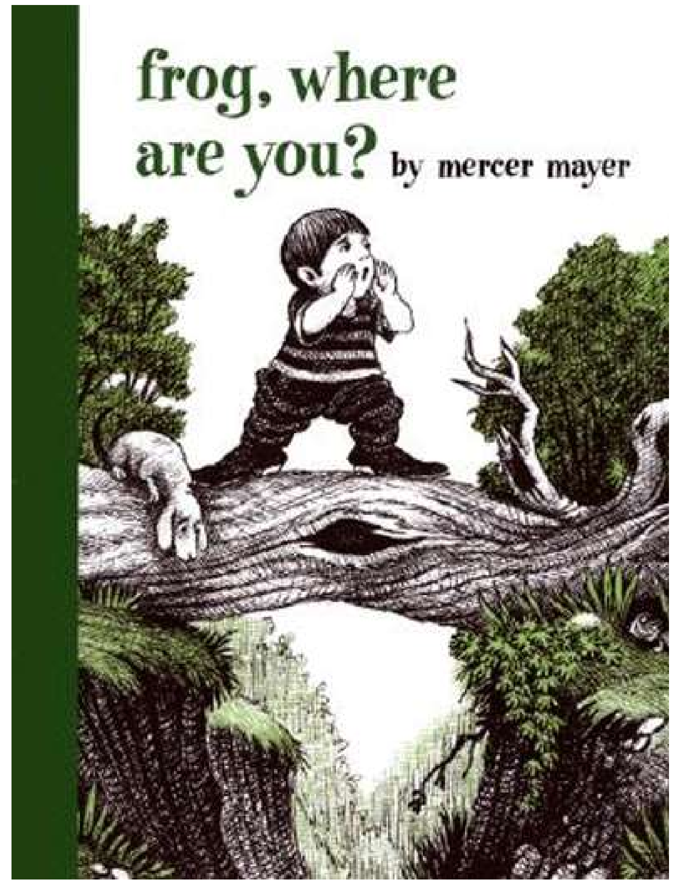

Once upon a time in 1997

Methodology

Counting

Spontaneous speech

- Children’s speech in a school setting in England.
- 11 children, average age M=4.9, SD=0.3; 5F/6M
- Three kinds of utterances
  - Words (“one”, “two”, “three”, …)
  - Simple sentences (“The horse is in the stable”, …)
  - Spontaneous speech
- Three recording devices
  - NAO (V5.0, running NaoQi V2.1.4).
  - Studio grade microphone (Rode NT1-A)
  - Portable audio recorded (Zoom H1)



ASR performance for child speech

	Google M LD [95%CI]	*% rec.*	Bing M LD [95%CI]	*% rec.*	Sphinx M LD [95%CI]	*% rec.*	Nuance M LD [95%CI]	*% rec.*
fixed (n=34)	**0.34** [0.24,0.44]	*11.8 [38]*	0.64 [0.56,0.71]	*0 [0]*	0.68 [0.64,0.73]	*0 [0]*	0.76 [0.73,0.80]	*0 [0]*
spontaneous (n=222)	**0.39** [0.36,0.43]	*6.8 [17.6]*	0.64 [0.61,0.67]	*0.5 [2.4]*	0.80 [0.77,0.84]	*0 [0]*	0.80 [0.78,0.82]	*0 [0]*
spontaneous clean only (n=83)	**0.40** [0.35,0.45]	*6.0 [16.9]*	0.63 [0.58,0.68]	*1.2 [1.2]*	0.78 [0.72,0.85]	*0 [0]*	0.78 [0.75,0.81]	*0 [0]*

Only 6% to 12% of words recognised correctly

Kennedy, J., Lemaignan, S., Montassier, C., Lavalade, P., Irfan, B., Papadopoulos, F., ... & Belpaeme, T. (2017, March). Child speech recognition in human-robot interaction: evaluations and recommendations. In *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction* (pp. 82-90).

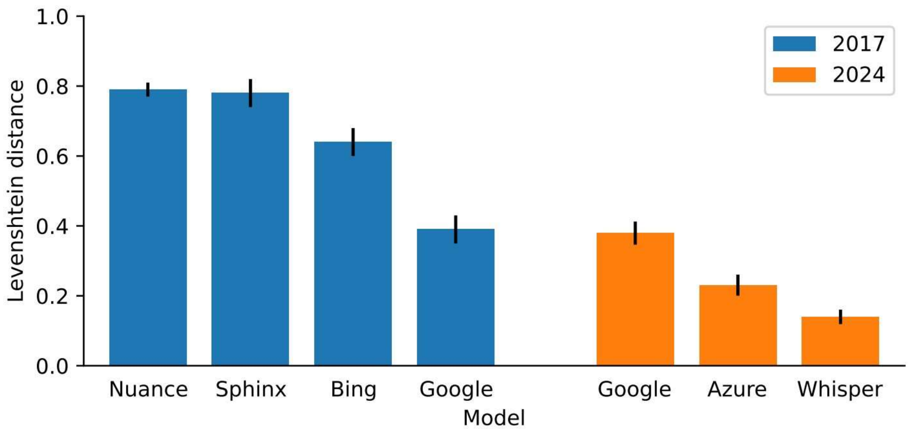
Fast forward to October 2023: OpenAI's Whisper 1550M model

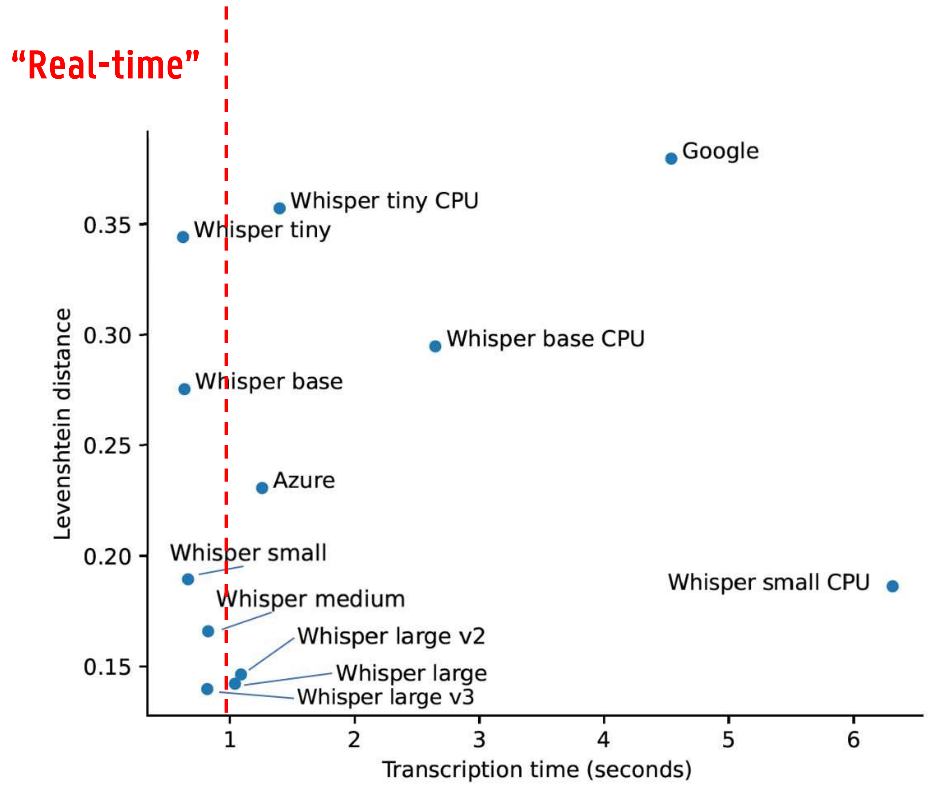
A boy looking at the frog. He's looking at it too. He's on the whole pack in the pond. The frog is gone. The boy found a shoe. The dog fell out.

The dog is looking and the kid is watching. The dog is looking and the kid is watching. The frog is getting out of the jar when the kid and the dog is asleep. The kid woke up, so did the dog and he saw that the frog was gone. He was searching for him everywhere but they couldn't find him.

Recordings at https://zenodo.org/records/200495

Performance of Whisper

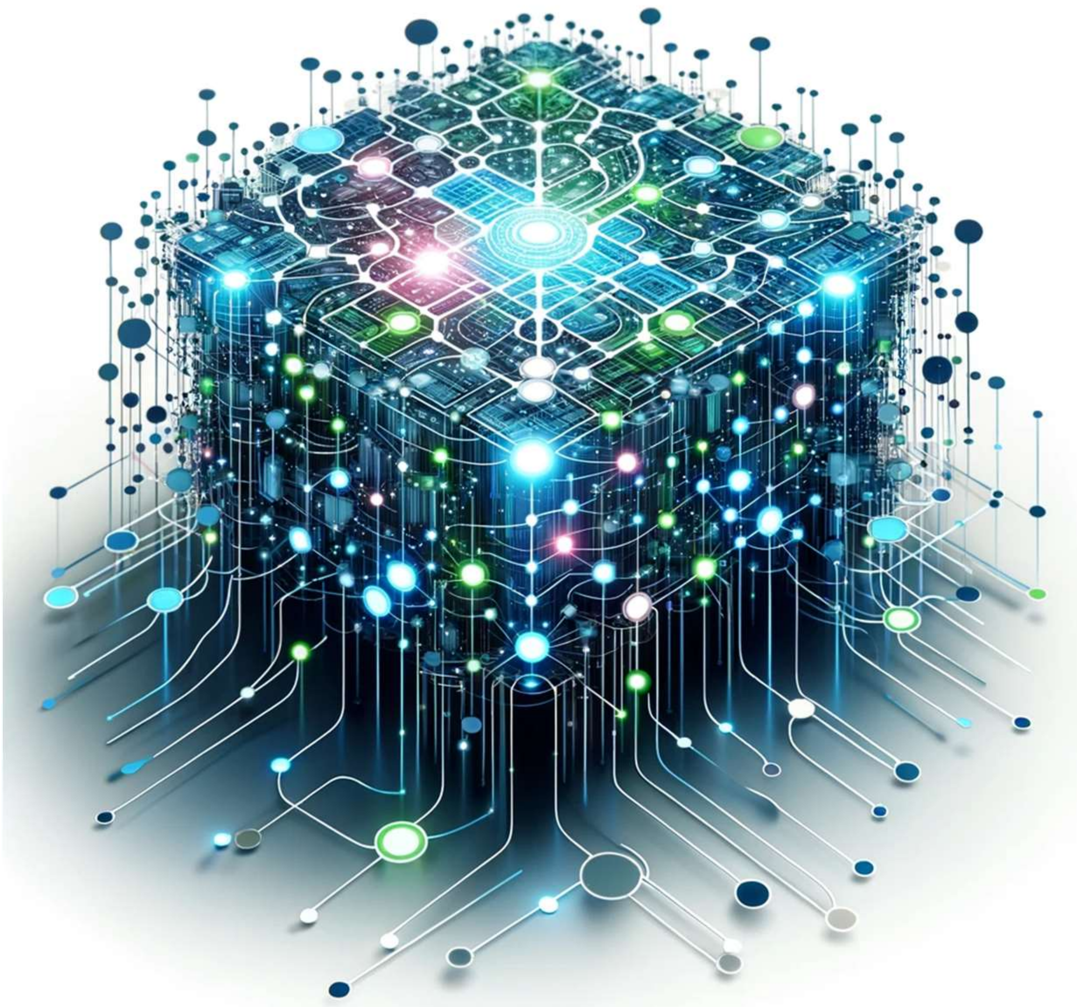




GHENT UNIVERSITY · imec

Large Language Models

- Unprecented leap in quality of generative language performance
- Interactive and robust NLI.
- Adaptation through one shot/few shot prompting.
- Democratised access to LLMs.



GHENT UNIVERSITY · imec

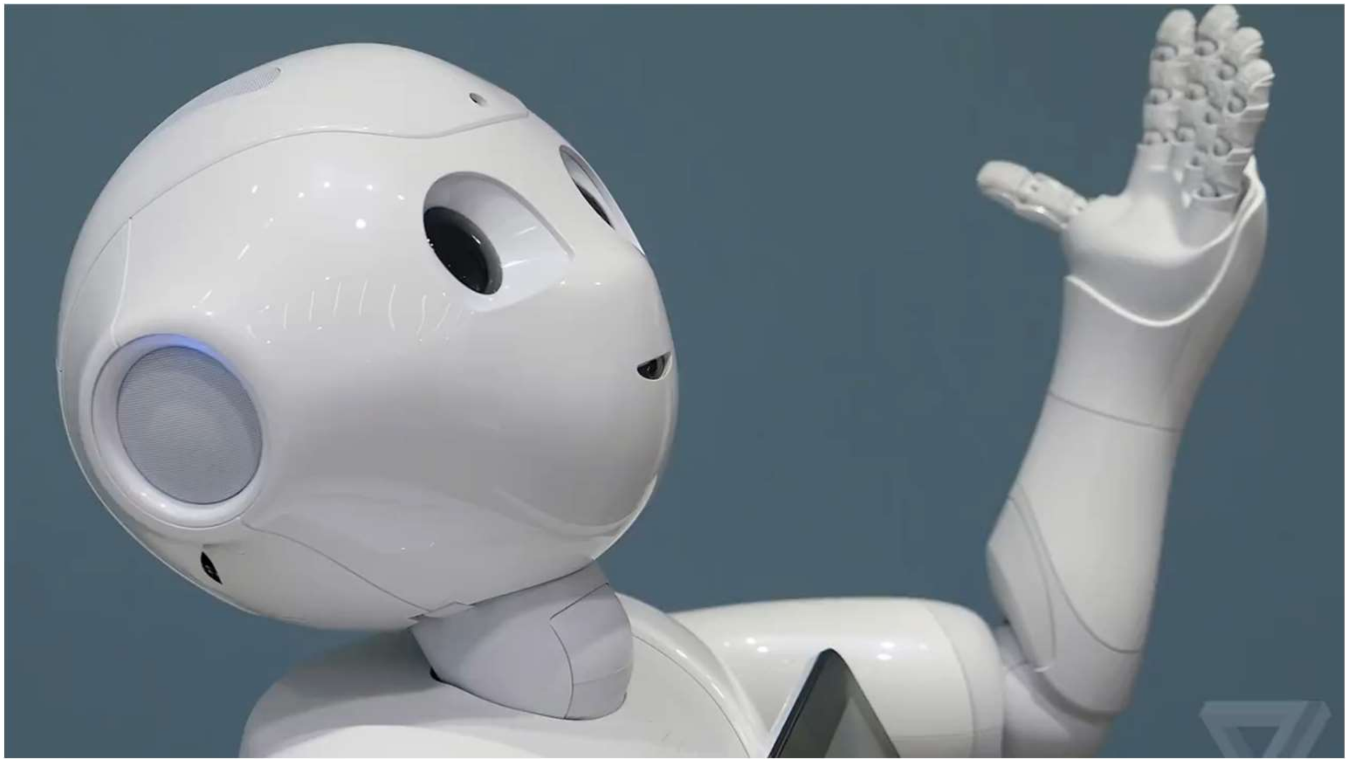
LLMs and robots

- Abilities needed for social robots
  - Open-domain dialogue capabilities
  - Storing state within a conversation
  - Can cover for poor performance by ASR
- Speed “good enough” for conversation
- Prompting to create robot persona
- Storing salient elements for next conversation

```
You are Pepper. Pepper is a helpful,
creative, clever, empathic, friendly,
and emotional robot. Now, you will have
a natural and friendly conversation
with your friend called {user_name}.
The next information is about the last
conversation and the information that
you already know about your friend.
Information: ###{last_conversation}###.
The current date and time are
{now.strftime('%m/%d/%Y %H:%M')} and
the location where you are speaking is
{location}. The conversation will be in
{language}. Keep your responses in that
language, brief, short and chatty.
```

Work by Maria Jose Pinto Bernal

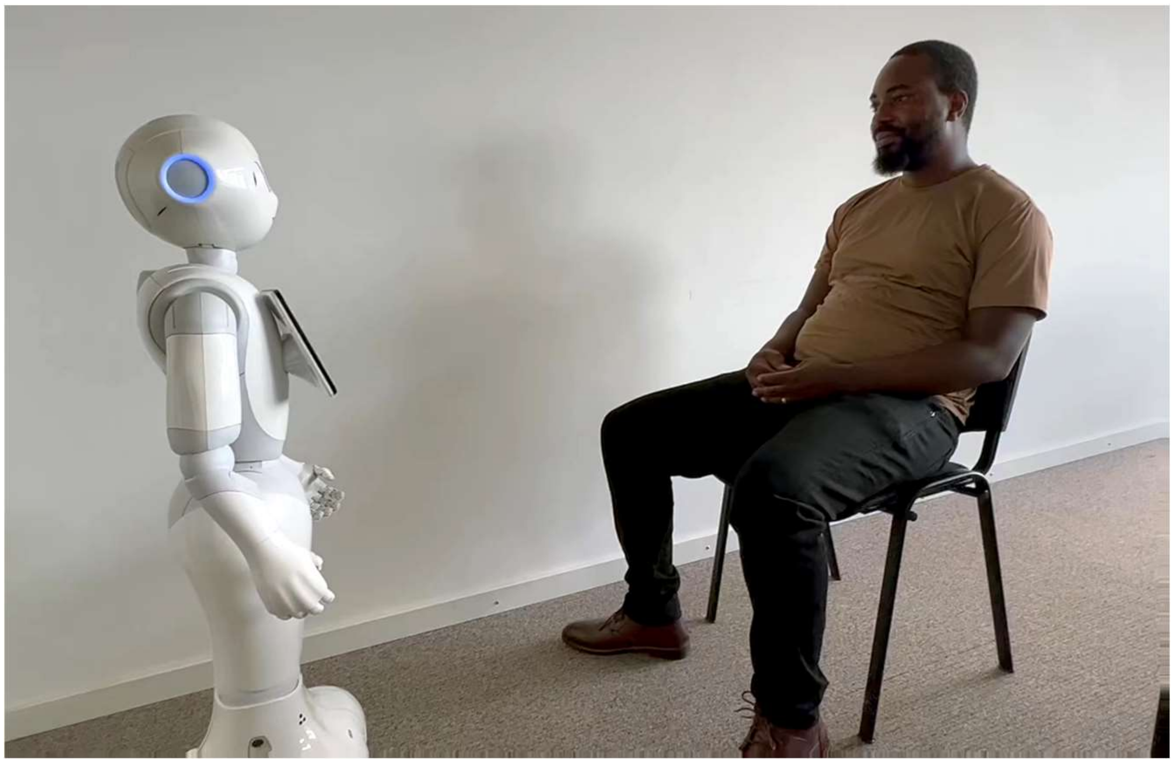
Large Language Models for embodied interaction



GHENT UNIVERSITY · imec

Work by Maria Jose Pinto Bernal

Tanzanian Swahili: 0.0080% of ChatGPT training data



GHENT UNIVERSITY · imec

Source https://commoncrawl.github.io/cc-crawl-statistics/plots/languages

Multimodal interaction

LLMs incorporate **knowledge** about the physical and social world.

Even when LLMs are **disembodied.**

This is exactly the **world knowledge** we have been waiting for in AI.

Ahn, M., Brohan, A., Brown, N., Chebotar, Y., Cortes, O., David, B., ... & Yan, M. (2022). Do as I can, not as I say: Grounding language in robotic affordances. *arXiv preprint arXiv:2204.01691*.

GHENT UNIVERSITY · imec

Multimodal interaction with GPT4

Challenges

- Prompt engineering turns into an art form
- Integrating across modalities
  - The sequential nature of processing (TTS, LLM, ASR) is not the right way.
  - Dynamics and timing
- Integrating with services
- Computational effort needed
  - But the expense of running multimodal models is coming down – Llama 3
- Different embodiments and physical capabilities
  - “That is really great. I love swimming in the sea too.”
- …

“Cognitive skills” of LLMs

- Understanding of physical and social world
- (Good enough) Understanding of causality
- Reasoning by analogy
- Creativity
- Theory of Mind
- Aligned to human values
- …

Note: all on a surface level, but LLMs pass the “duck test” for many cognitive skills.

GHENT UNIVERSITY · imec

Generative AI for language education

LLM and Stable Diffusion to generate adapted learning material on the fly.

GHENT UNIVERSITY · imec

Verhelst, E., Janssens, R., Demeester, T., & Belpaeme, T. (2024, March). Adaptive Second Language Tutoring Using Generative AI and a Social Robot. In Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction (pp. 1080-1084).

Are the social robots coming?

For the first time AI has broken through **Moravec's paradox.**

Social interaction while being a hard technical problem is no longer an impossible technical problem

Thank you!

Giulio Antonio Abbo

Ruben Janssens

James Kennedy

Séverin Lemaignan

Maria Jose Pinto Bernal

Qiaoqiao Ren

Edger Rutatola

Eva Verhelst

Pieter Wolfert

GHENT UNIVERSITY · imec

Wrap up

- Learning is socially gated and "social robots" are likely to play a role in our future learning.
- Artificial Intelligence is for the first time breaking through Marovec's paradox.
- Social AI will lead to a learning technology revolution.

Thank you!

Giulio Antonio Abbo
Madeleine Bartlett
Paul Baxter
Ruben Janssens
James Kennedy
Séverin Lemaignan
Qiaoqiao Ren
Edger Rutatola
Emmanuel Senft
Pieter Wolfert
Serge Thill
Eva Verhelst

L2TOR
VALAWAI
fwo Research Foundation Flanders
Opening new horizons

GHENT UNIVERSITY
imec


1. narrow interactions

Teleoperation, aka **Wizard of Oz,** in HRI research makes up

50% of all studies

Useful for a quick and cheap study or as a stub for underperforming technology

Baxter, P et al. (2016). From characterising three years of HRI to methodology and reporting recommendations. In *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (pp. 391-398). IEEE.

Programming social behaviour by hand is **hard and time consuming**

Limited interaction scope

Not long-term: most interactions are well under 20 minutes

Commercial social robots **underwhelm**

GHENT UNIVERSITY · imec


Input
Action selection
Output

Input

GHENT UNIVERSITY · imec

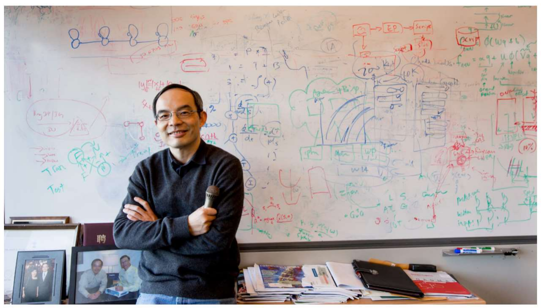
Speech recognition: super human?

Microsoft claims new speech recognition record, achieving a super-human 5.1% error rate

BY TODD BISHOP on August 20, 2017 at 7:44 pm

5 Comments | Share 820 | Tweet | Share | Reddit | Email

Cloud Tech Summit: Tickets on sale here!



Xuedong Huang, a Microsoft Technical Fellow in AI and Research, leads Microsoft's Speech and Language Group. (Microsoft Photo)

GeekWire Newsletters

Subscribe to GeekWire's free newsletters to catch every headline

Enter your email address | Subscribe

Send Us a Tip

Have a scoop that you'd like GeekWire to cover? Let us know.

Send Us a Tip

IBM speech recognition is on the verge of super-human accuracy

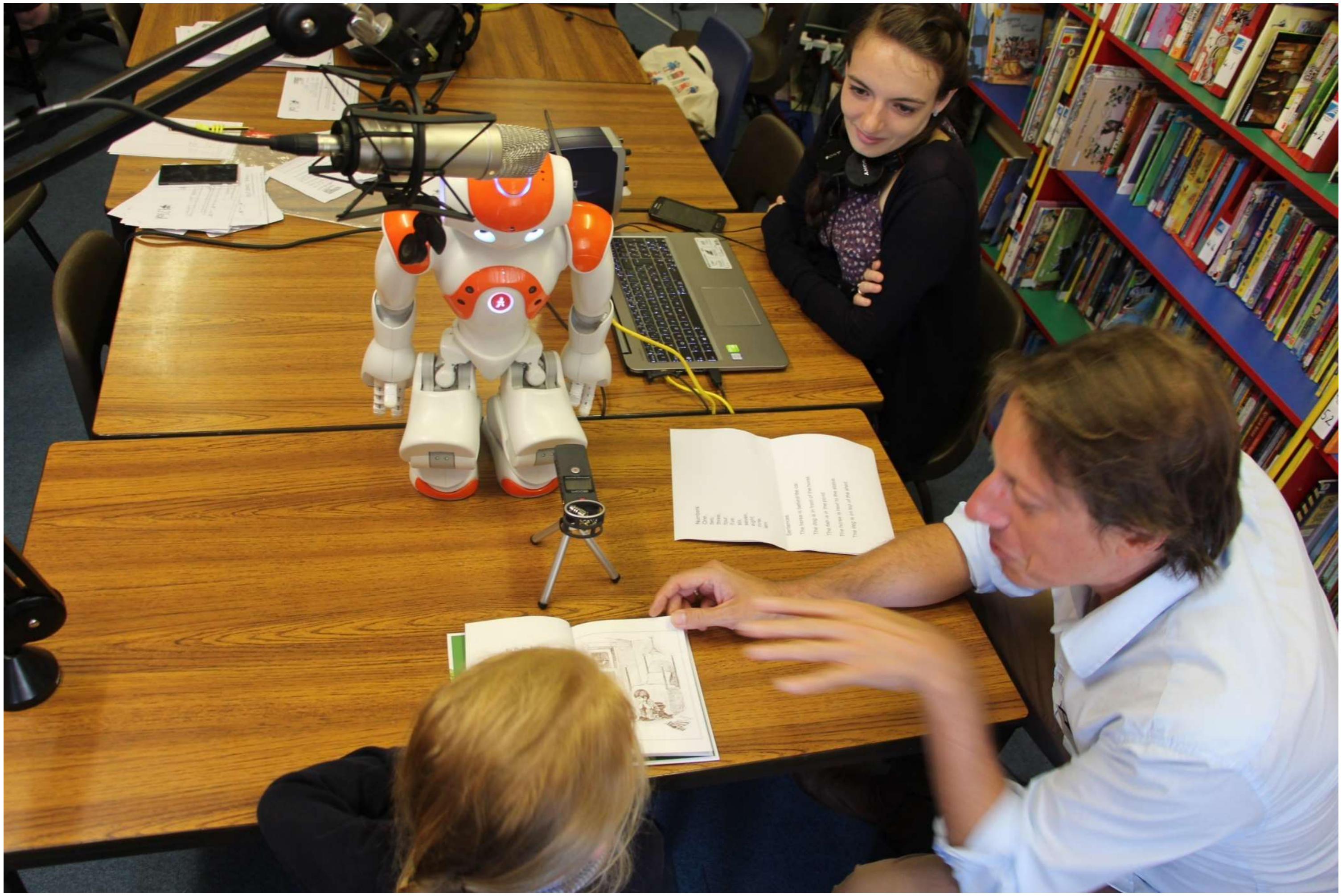
Chris Weller
Mar. 9, 2017, 4:58 PM 1,494

FACEBOOK | LINKEDIN | TWITTER | EMAIL | PRINT

In the world of speech recognition software, 5.1% is kind of a magic number.

Companies that can create software with error rates falling in that ballpark are essentially matching the capabilities of humans, who miss roughly 5% of the words in a given conversation.

GHENT UNIVERSITY · imec



GHENT UNIVERSITY · imec

ASR performance for child speech

	Google		Bing		Sphinx		Nuance	
	M LD [95%CI]	*% rec.*	*M* LD [95%CI]	*% rec.*	*M* LD [95%CI]	*% rec.*	*M* LD [95%CI]	*% rec.*
fixed (n=34)	0.34 [0.24,0.44]	*11.8 [38]*	0.64 [0.56,0.71]	*0 [0]*	0.68 [0.64,0.73]	*0 [0]*	0.76 [0.73,0.80]	*0 [0]*
spontaneous (n=222)	0.39 [0.36,0.43]	*6.8 []*	0.64 [0.61,0.67]	*0.5 []*	0.80 [0.77,0.84]	*0 [0]*	0.80 [0.78,0.82]	*0 [0]*
spontaneous clean only (n=83)	0.40 [0.35,0.45]	*6.0 []*	0.63 [0.58,0.68]	*1.2 []*	0.78 [0.72,0.85]	*0 [0]*	0.78 [0.75,0.81]	*0 [0]*

6% to 38% words correctly recognised

Range of recognition success of Google ASR engine with no grammatical constraints. **Google is the best performing engine (compared to Sphinx, Bing, and Nuance).**

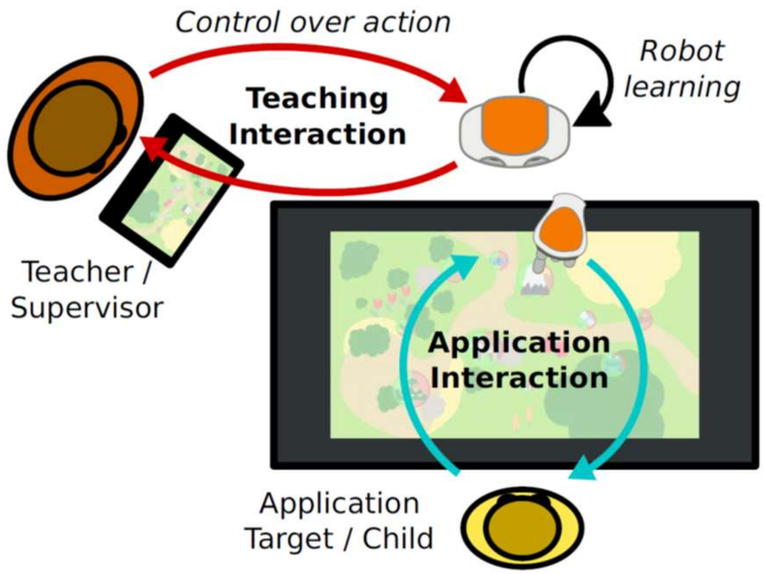
GHENT UNIVERSITY · imec

Kennedy, J., Lemaignan, S., Montassier, C., Lavalade, P., Irfan, B., Papadopoulos, F., ... & Belpaeme, T. (2017, March). Child speech recognition in human-robot interaction: evaluations and recommendations. In *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction* (pp. 82-90).

Action selection

Data-driven action selection

- Learning for robots
  - **Reinforcement Learning** requires exploration and mistakes to be cheap.
  - **Learning from Demonstration** better suited for kinestethic teaching.
- Needed
  - Online learning
  - Active involvement of the expert during learning.
  - Understanding of what the AI learns.
  - Retaining control over the AI.
- Solution is **SPARC: supervised progressively autonomous robot competencies.**

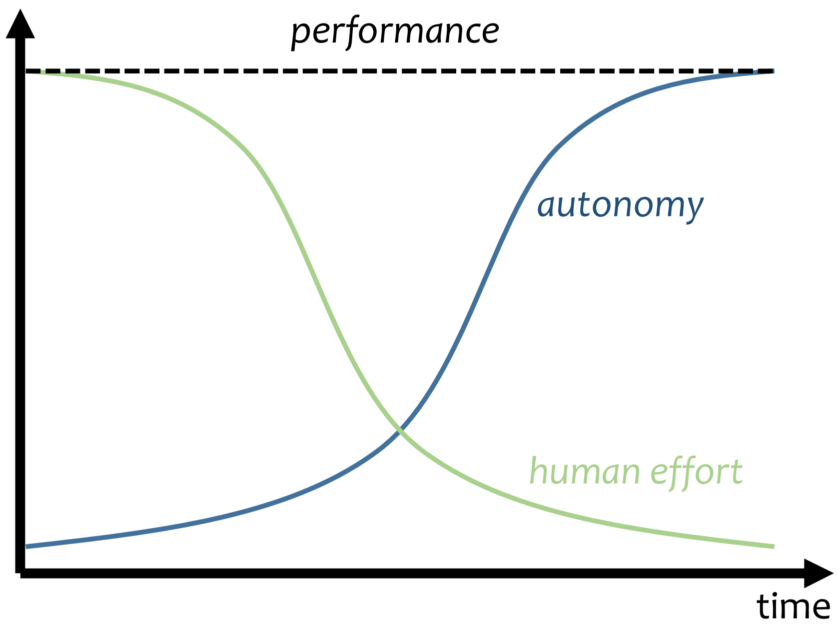


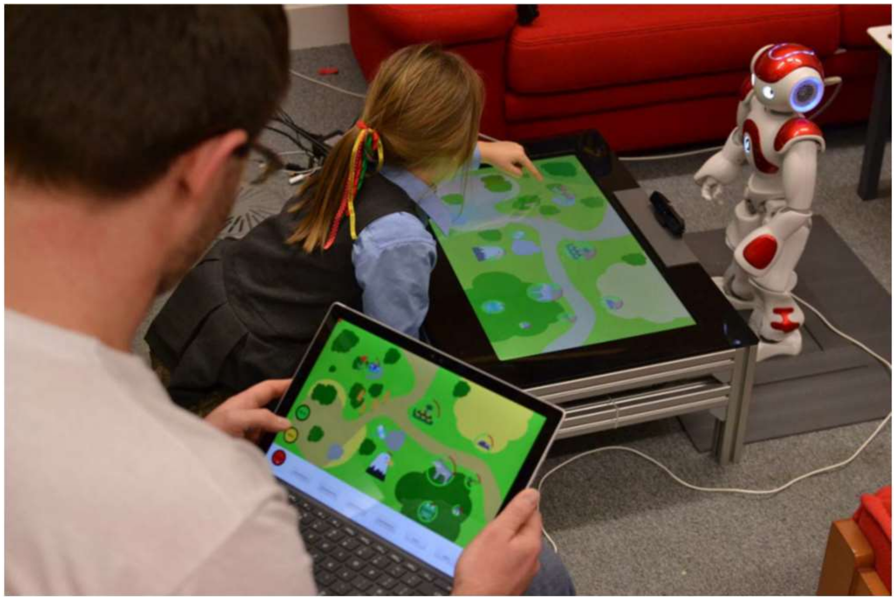
Senft, E., Lemaignan, S., Baxter, P. E., Bartlett, M., & Belpaeme, T. (2019). Teaching robots social autonomy from in situ human guidance. *Science Robotics*, 4(35).


9:41
New Message
Cancel
To: Jane Appleseed
Would you like to grab some coffee with me in a
"a"
about
an
q w e r t y u i o p
a s d f g h j k l
z x c v b n m
123
space
return

SPARC: Core concepts

Robot progressively learns appropriate autonomous behavior from in situ human demonstrations and guidance.

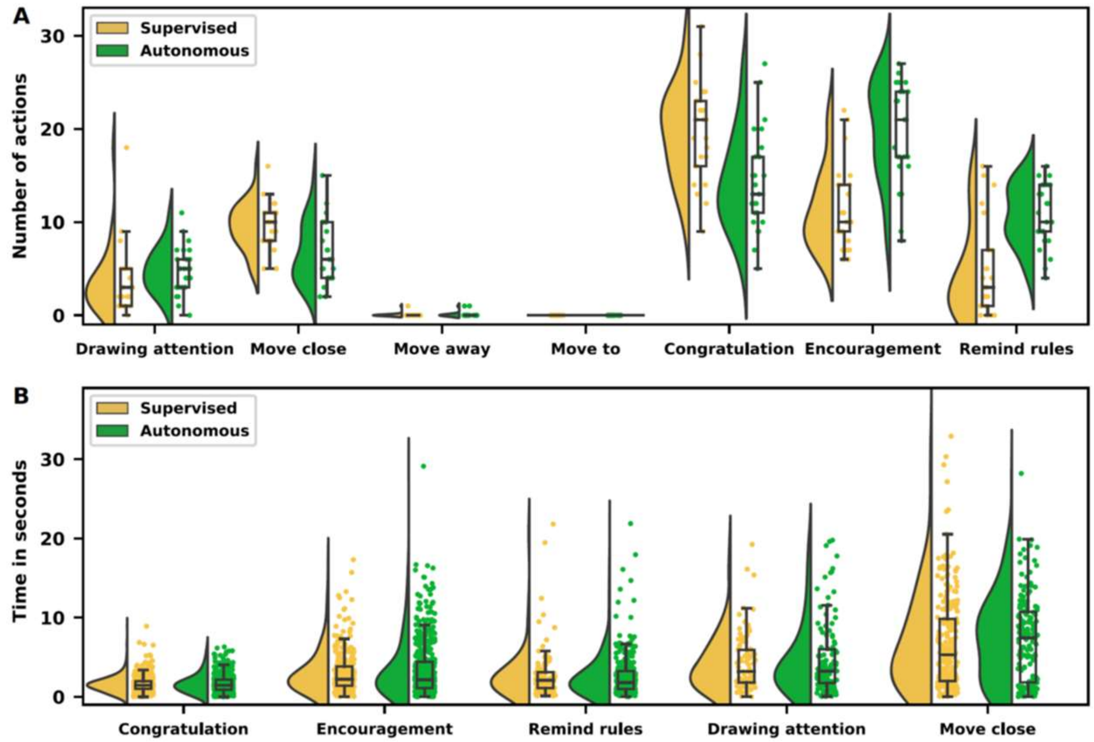




Senft, E., Lemaignan, S., Baxter, P. E., Bartlett, M., & Belpaeme, T. (2019). Teaching robots social autonomy from in situ human guidance. *Science Robotics*, 4(35).

Results

- Three conditions
  - SPARC robot
  - Autonomous robot
  - Passive robot (control condition)
- 75 children (n = 75; age: M = 9.4, SD = 0.71; 37 female), each of the three conditions was allocated 25 children.
- Challenging learning domain
  - 210 input dimensions and output action space of 655 actions.



Senft, E., Lemaignan, S., Baxter, P. E., Bartlett, M., & Belpaeme, T. (2019). Teaching robots social autonomy from in situ human guidance. *Science Robotics*, 4(35).

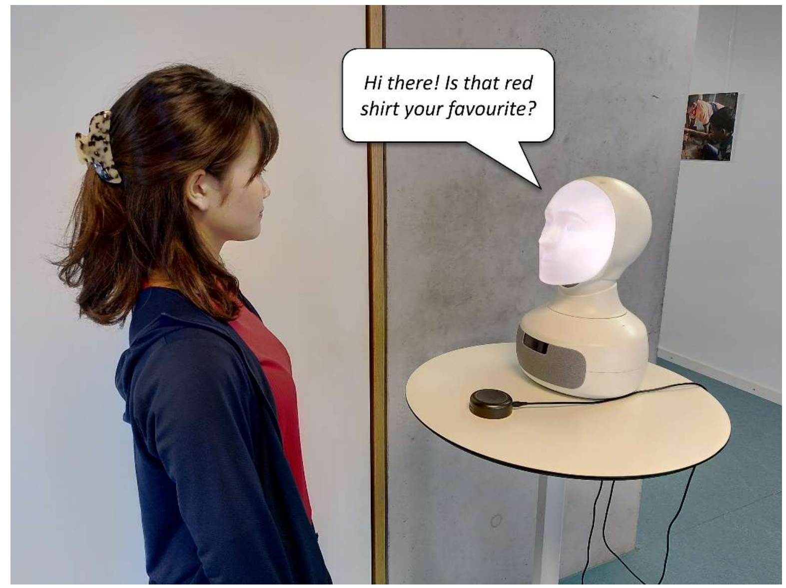
Output

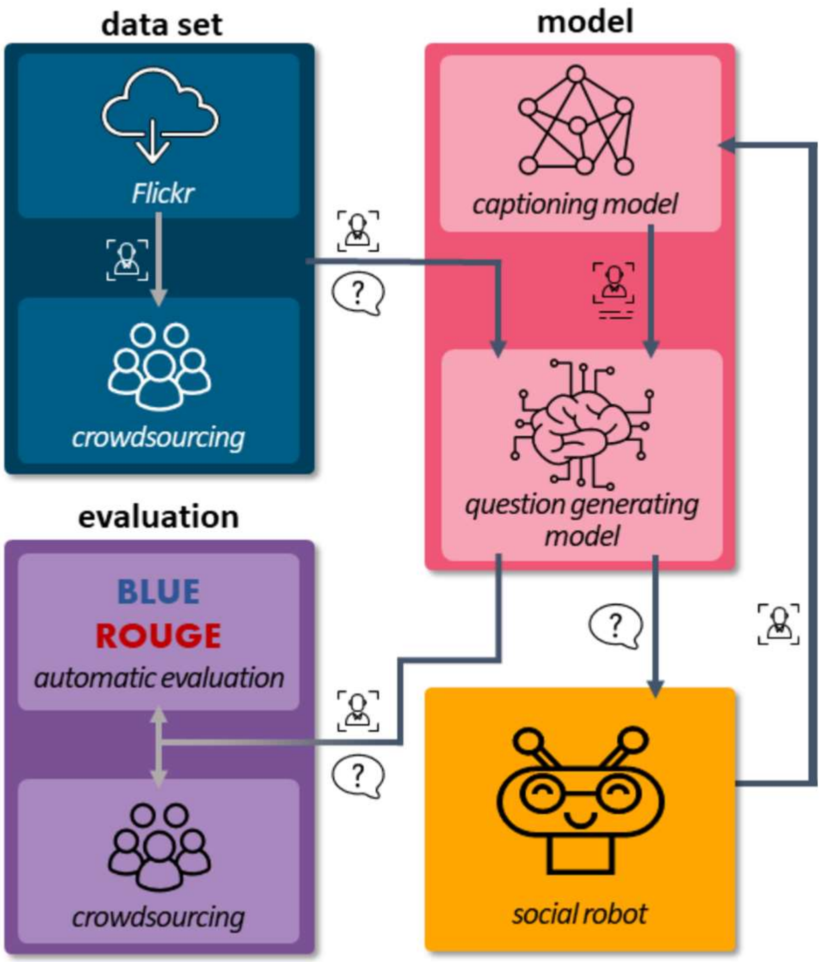
Multi-modal interaction

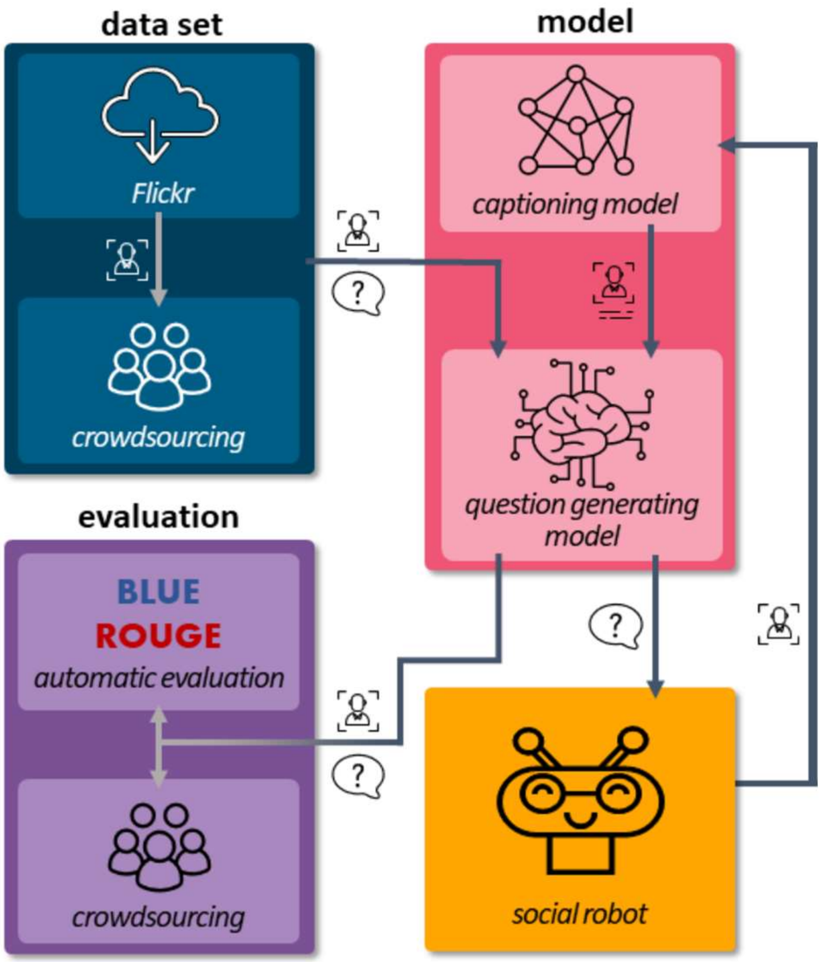
Current **chatbots are uni-modal**

Phatic communication or small talk

All good conversation starts with a
conversation starter.
Can these be generated?

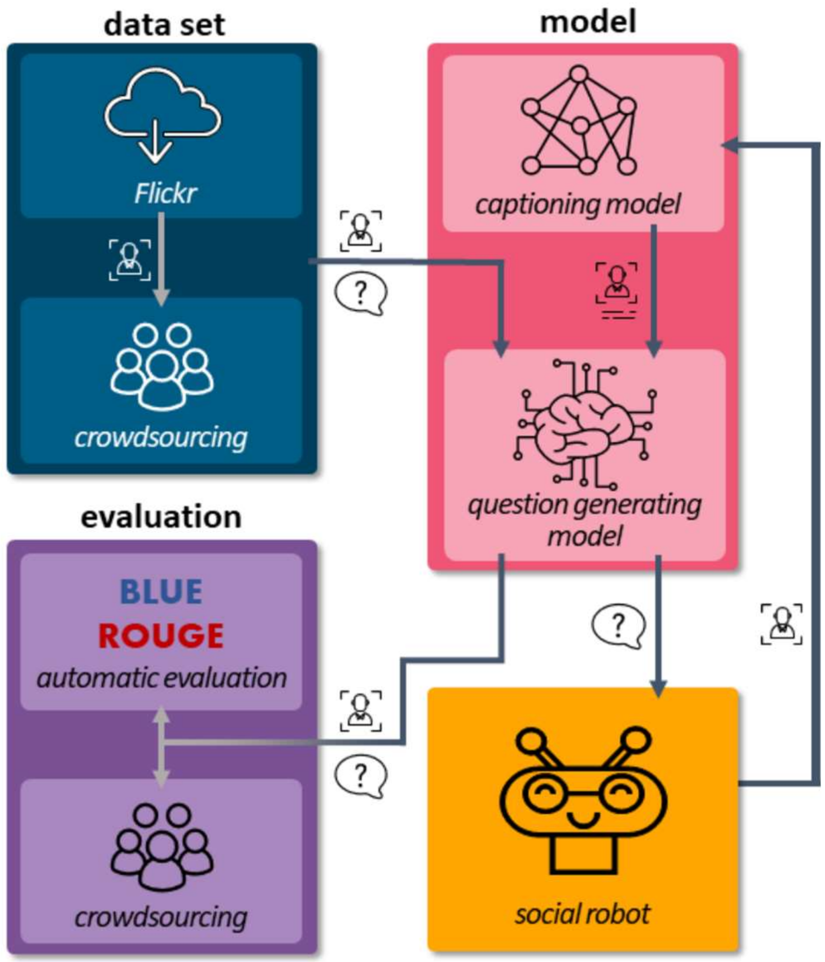


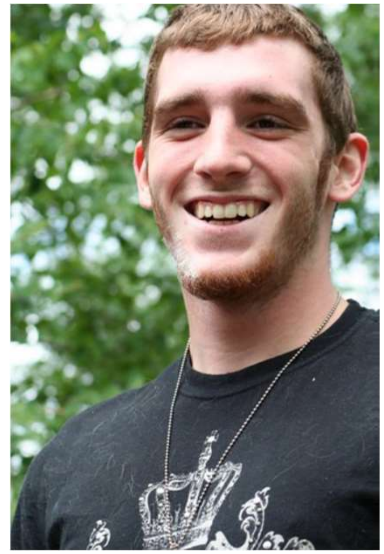

data set
Flickr
crowdsourcing
model
captioning model
question generating model
evaluation
BLUE
ROUGE
automatic evaluation
crowdsourcing
social robot



Have you always had freckles? Were you born with red hair? Is white your favourite colour? Is that cool hat typical of your country? Do your parents also have blue eyes? Where did you get that scarf? Do you go to bars often? What kind of drinks do you like? Where did you get that necklace? Do you put the braids in your hair?

GHENT UNIVERSITY · imec





man with brown hair. the eyes are brown. the man has a beard. the man has short hair. the man is smiling. the man is wearing a black shirt. the man is wearing a necklace. the shirt is white. white letters on shirt.

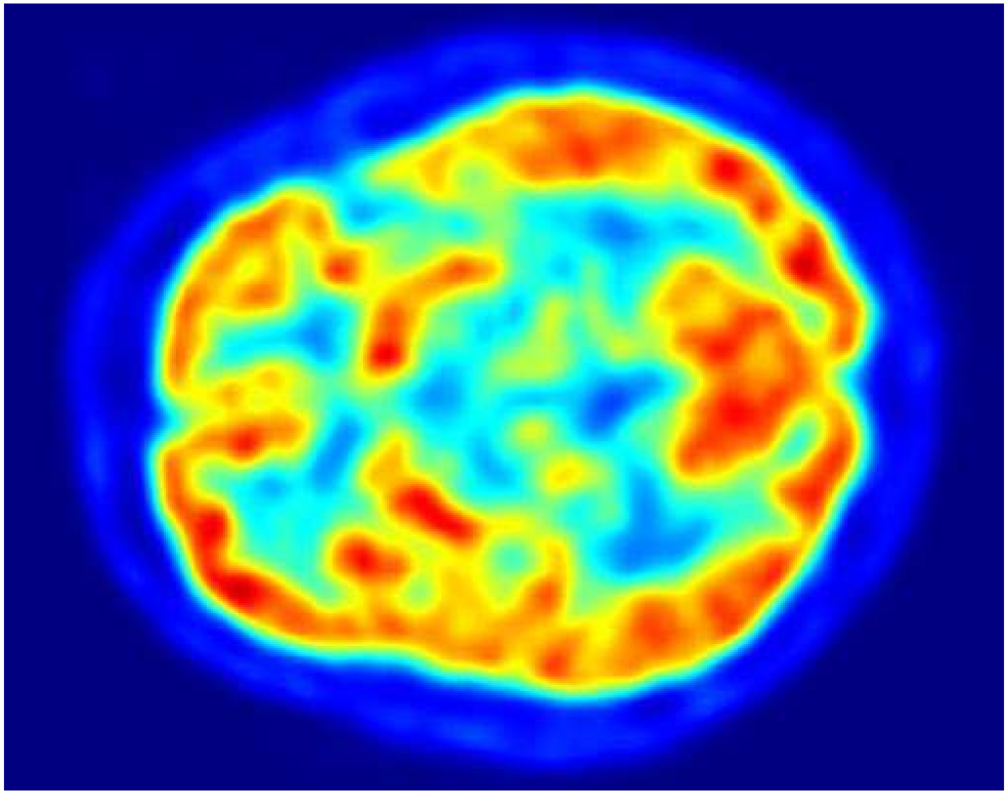
Models

- Dense captioning network
- Neural generative model: BART
- Standard retrieval model: TF-IDF
- Neural retrieval model: Transformer-based Bi-Encoder

Demonstration

A call to arms for AI

- Interaction recruits all our cognitive faculties
- An effort by the entire AI department
- People and robots meet halfway
- Even simple systems can make a large impact



Sci-news.com