ISTITUTO ITALIANO
DI TECNOLOGIA

Motor invariants in action recognition

Giorgio Metta
Cognitive Humanoids Laboratory
http://www.cognitivehumanoids.eu

Dept. of Robotics, Brain and Cognitive Sciences

Cogsys
Cognitive Systems

RobotCub

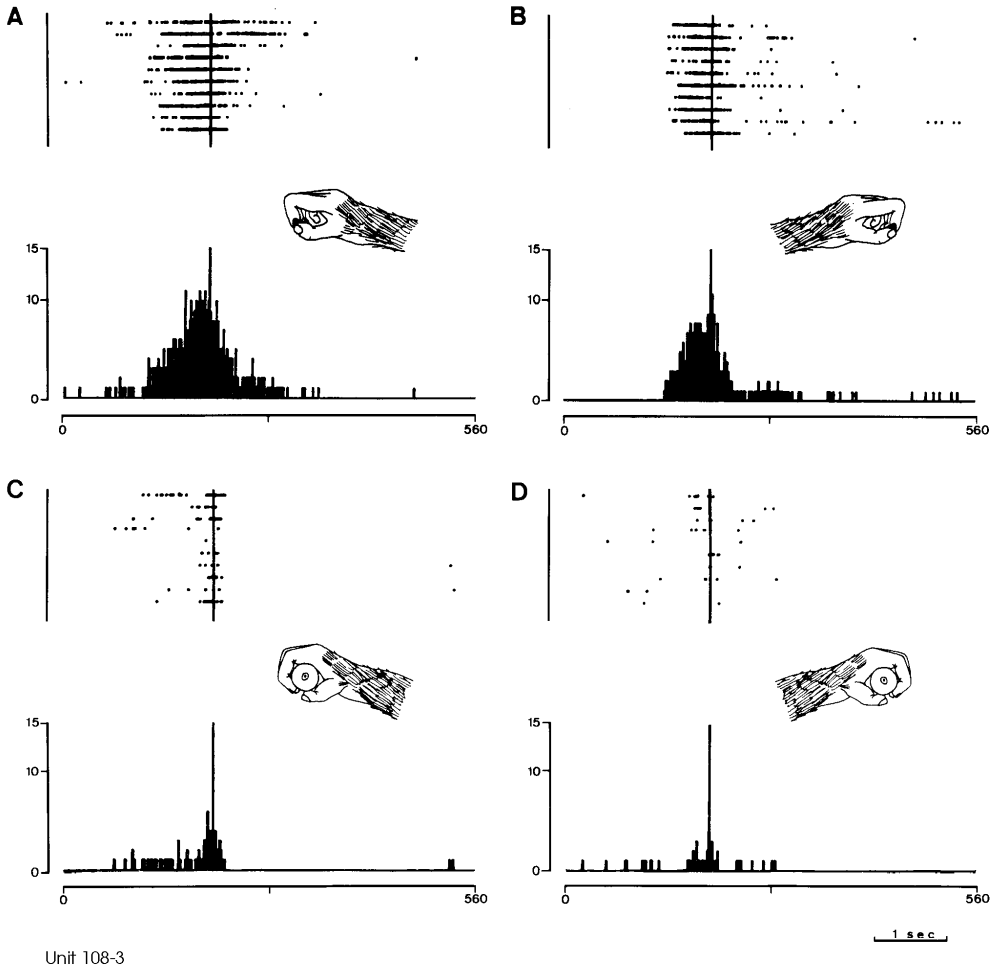
Alvin Liberman's idea

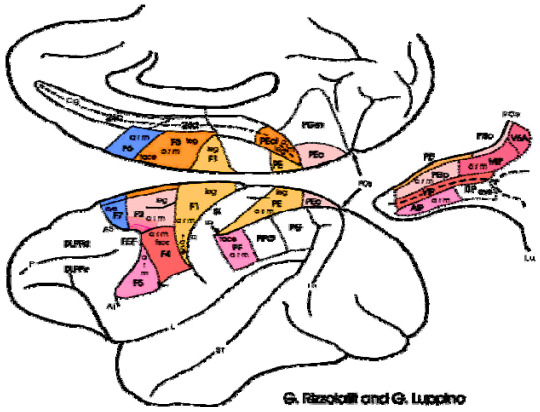
"Thus, it appeared that the objects of speech perception were not to be found at the acoustic surface. They might, however, be sought in the underlying motor processes, if it could be assumed that the acoustic variability required for an invariant percept resulted from the temporal overlap, in different contexts, of correspondingly invariant units of production"

(Liberman & Mattingly, 1985, page 2)

- Perhaps it is the same in other modalities…
  - Rizzolatti et al. 1992: discovery of mirror neurons
  - Fadiga et al. 1999: mirror effects due to motor imagery
  - Rizzolatti & Arbib 1998: Mirror neurons and language
  - Fadiga et al. 2002: TMS experiment on speech listening

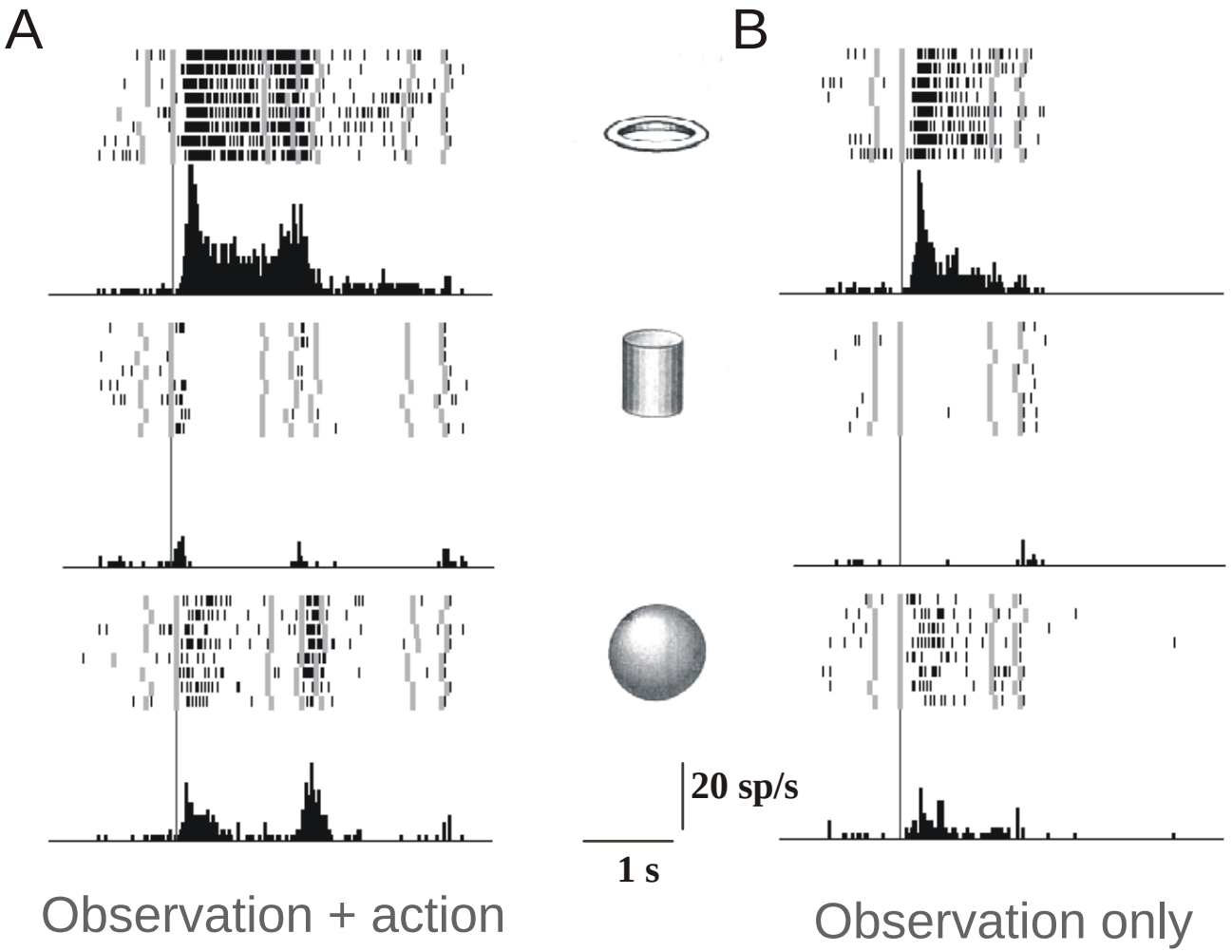
Grasping neurons





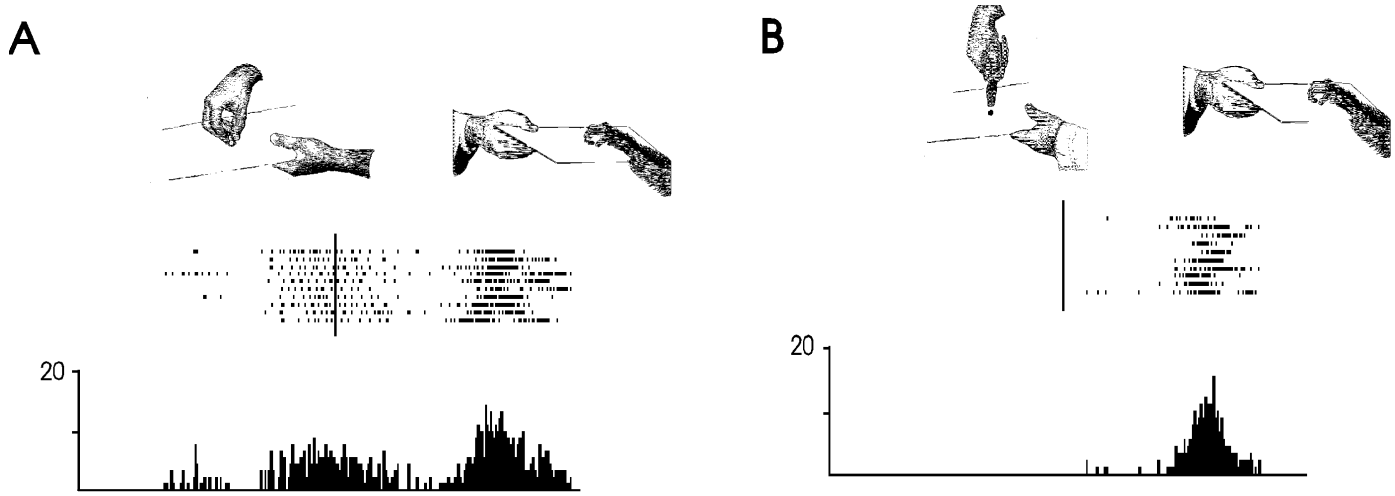
Fadiga et al. (various sources)

F5 canonical neurons



Mirror Neurons

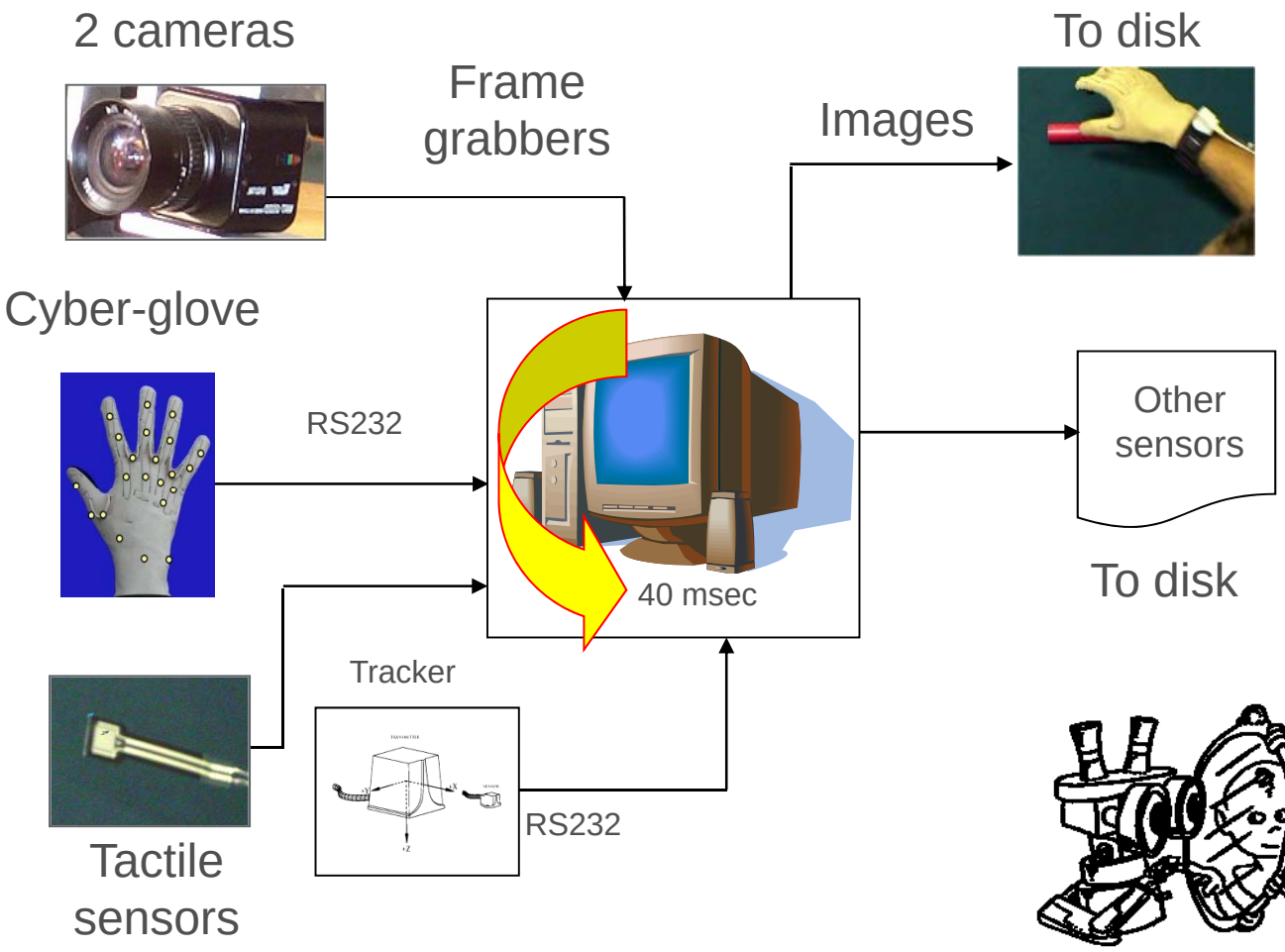
The neuron is activated by “seeing” someone else’s hand performing a manipulative action **and** while the monkey is performing the same action



The type of action seen is relevant

From: Fadiga, L., L. Fogassi, V. Gallese, and G. Rizzolatti, *Visuomotor Neurons: ambiguity of the discharge or "motor“ Perception?* Internation Journal of Psychophysiology, 2000. **35**: p. 165-177.

Data from human grasping

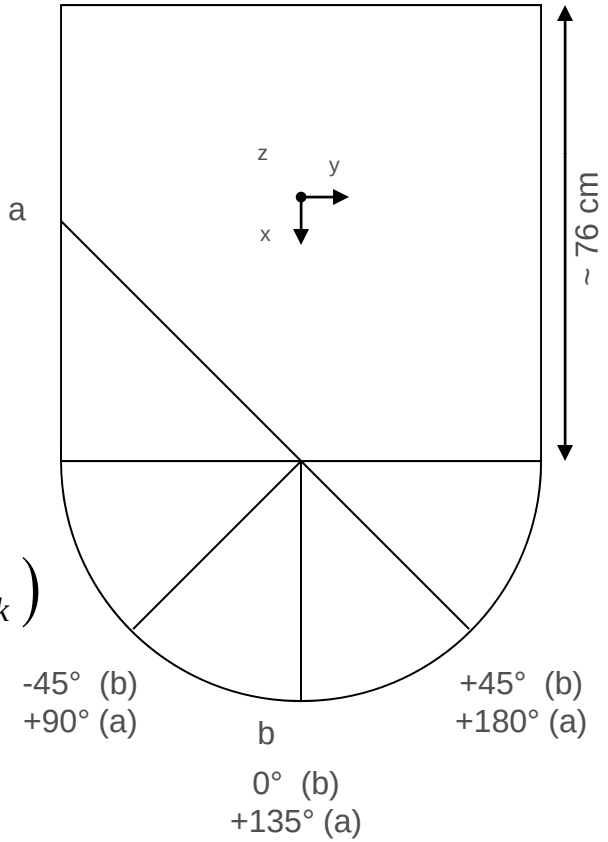


EU project MIRROR (2001-2003)

Bayesian classifier

{G_i}: set of gestures
F: observed features
{O_k}: set of objects

168 sequences per subject
10 subjects
6 complete sets

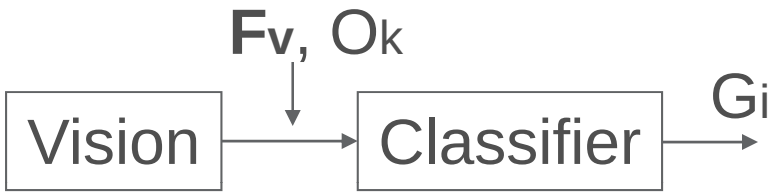


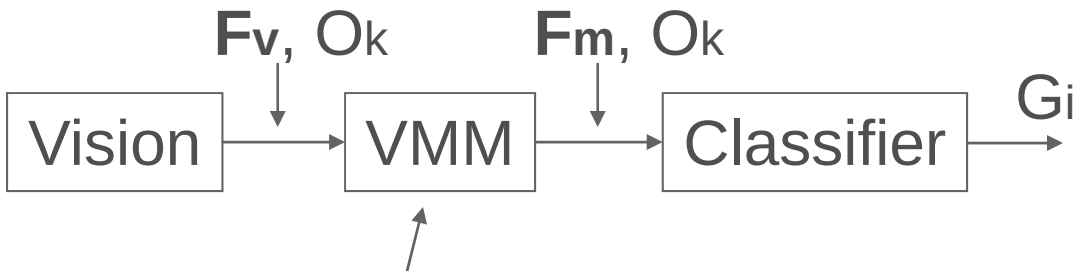
$p(G_i|O_k)$: priors (affordances)
$p(\mathbf{F}|G_i,O_k)$: likelihood to observe **F**

$$p\left(G_i \mid \mathbf{F}, O_k\right) = p\left(\mathbf{F} \mid G_i, O_k\right) p\left(G_i \mid O_k\right) / \, p\left(\mathbf{F} \mid O_k\right)$$

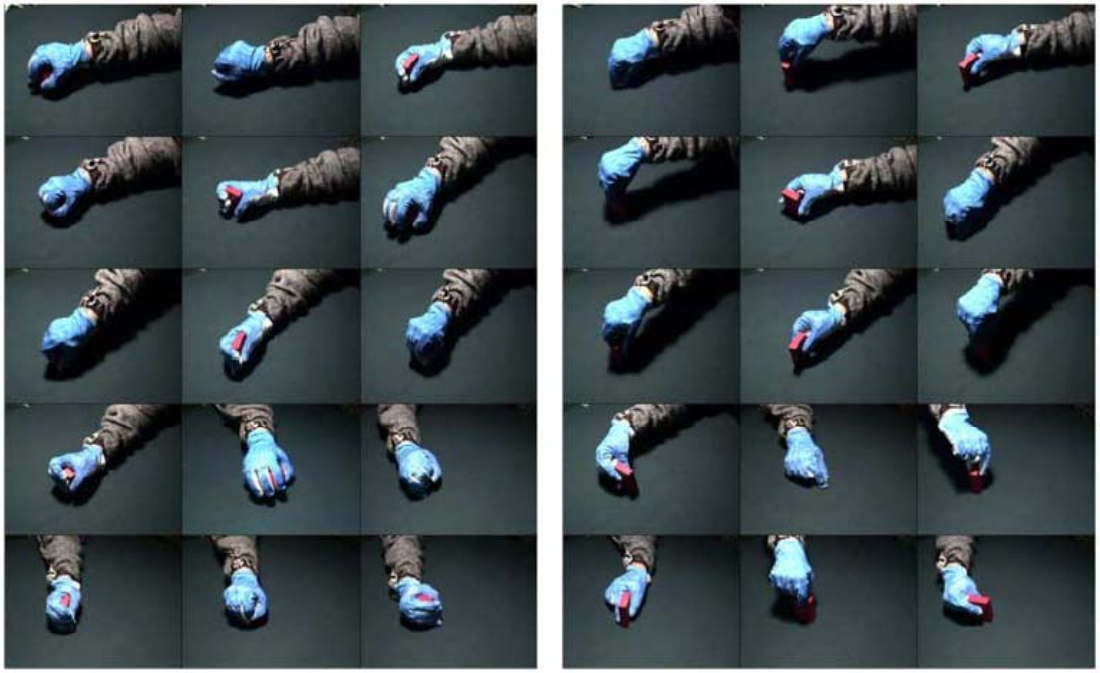
$$\hat{G}_{MAP} = \arg\max_{G_i} \left(G_i \mid \mathbf{F}, O_k\right)$$

Two types of experiments



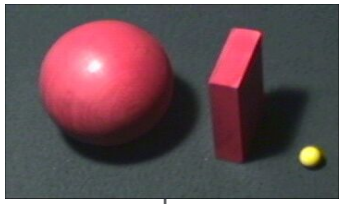


Role of motor information in action understanding

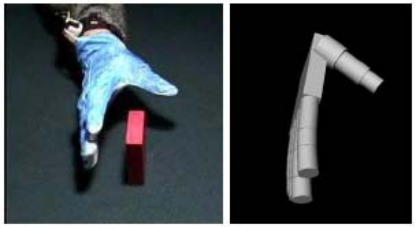


Grasping actions

Object affordances (priors)



Visual space Motor space



Classification
(recognition)

Understanding mirror neurons: a bio-robotic approach. *G. Metta, G. Sandini, L. Natale, L. Craighero, L. Fadiga*. Interaction Studies. Volume 7 Issue 2. 2006

Some results

	Exp. I (visual)	Exp. II (visual)	Exp. III (visual)	Exp. IV (motor)
	Training			
# Sequences	16	24	64	24
# of view points	1	1	4	1
Classification rate	100%	100%	97%	98%
# Features	5	5	5	15
# Modes	5-7	5-7	5-7	1-2
	Test			
# Sequences	8	96	32	96
# of view points	1	4	4	4
Classification rate	100%	30%	80%	97%

Additional neurophysiology

Current Biology 19, 1–5, March 10, 2009 ©2009 Elsevier Ltd All rights reserved DOI 10.1016/j.cub.2009.01.017

Report

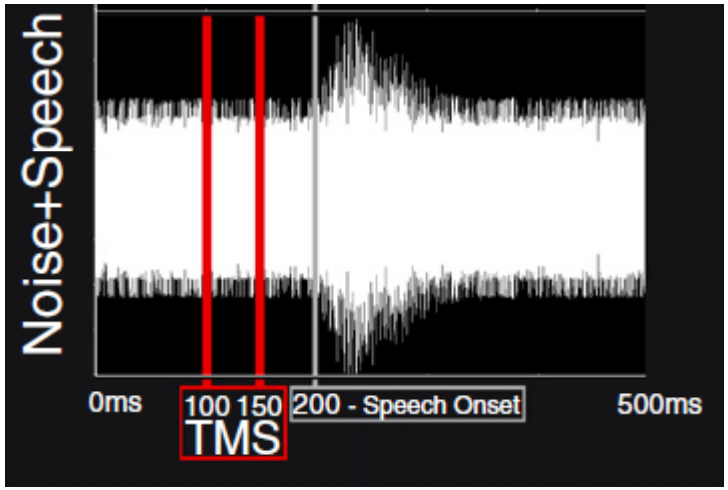
The Motor Somatotopy of Speech Perception

Alessandro D'Ausilio,[1] Friedemann Pulvermüller,[2] Paola Salmas,[3] Ilaria Bufalari,[1] Chiara Begliomini,[1] and Luciano Fadiga[1,3,*]
[1]DSBTA
Section of Human Physiology
University of Ferrara
Ferrara 44100
Italy
[2]Cognition and Brain Sciences Unit
Medical Research Council
Cambridge CB2 7EF
UK
[3]IIT, The Italian Institute of Technology
Genova 16163
Italy

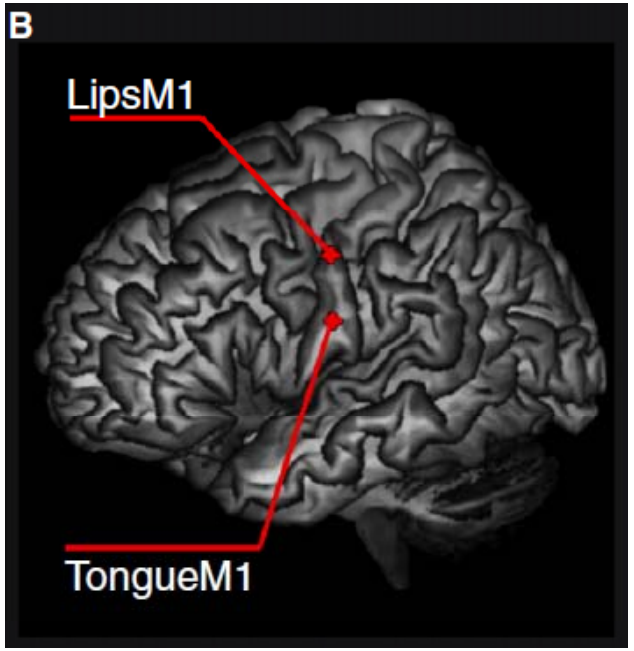
(MTSP) [3], an early precursor of a new zeitgeist, most radically postulated that the articulatory gestures, rather than sounds, are critical for both production and perception of speech (see [4]). On neurobiological grounds, fronto-temporal circuits are thought to play a functional role in production as well as comprehension of speech. The coactivation of motor circuits and the concurrent perception of self-produced speech sounds during articulations might lead to correlated neuronal activity in motor and auditory systems, triggering long-term plastic processes based on Hebbian learning principles [15–17]. The postulate of a critical role of actions in the formation of speech circuits is paralleled in more general action-perception theories emphasizing a critical role of action representations in action-related perceptual processes [18]. However, a majority of researchers are still skeptical toward a general role of motor systems in speech perception, admit-

TMS experiment

- Listening to [b] and [p], labial phonemes
- Listening to [t] and [d], dental phonemes

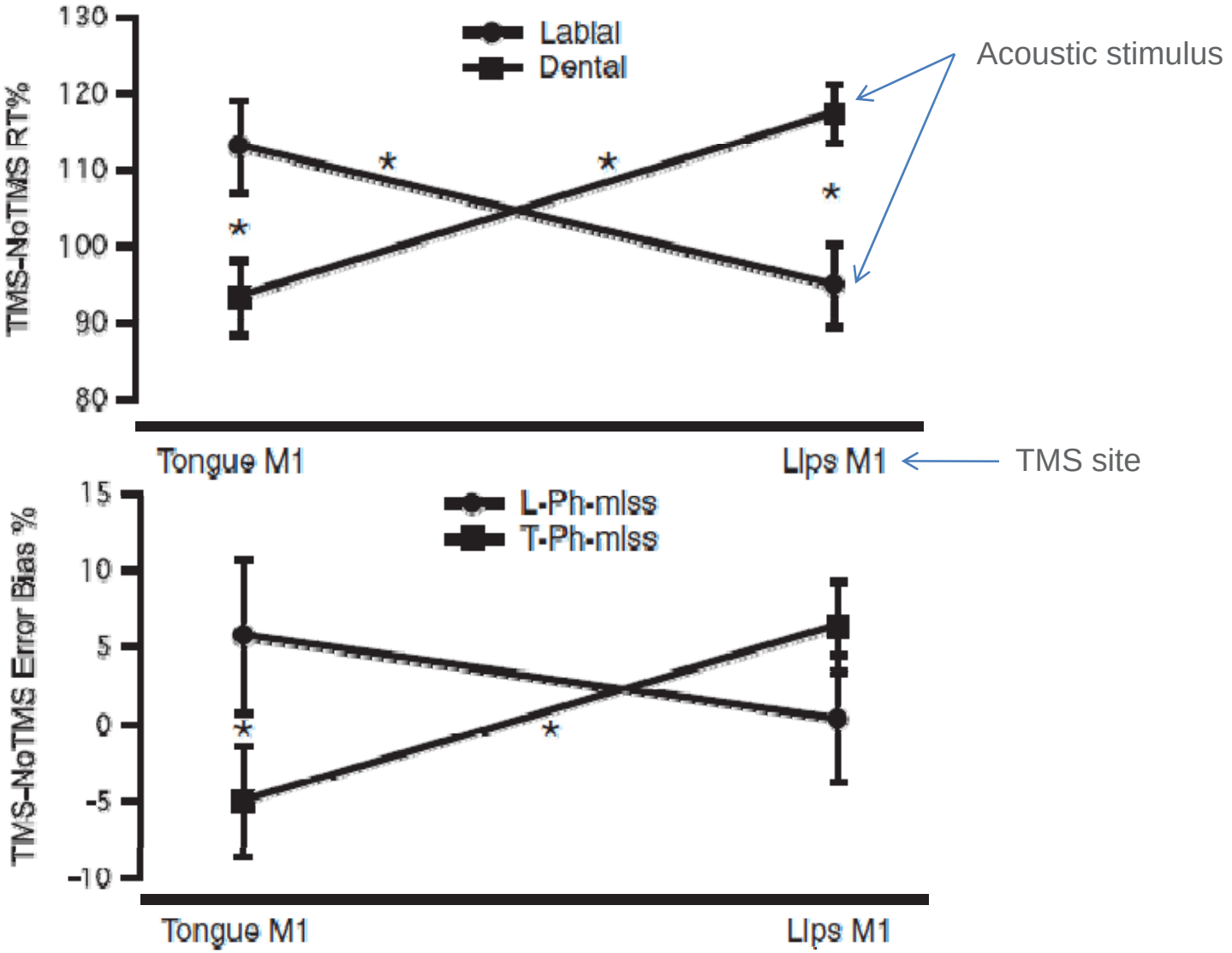


Stimulus

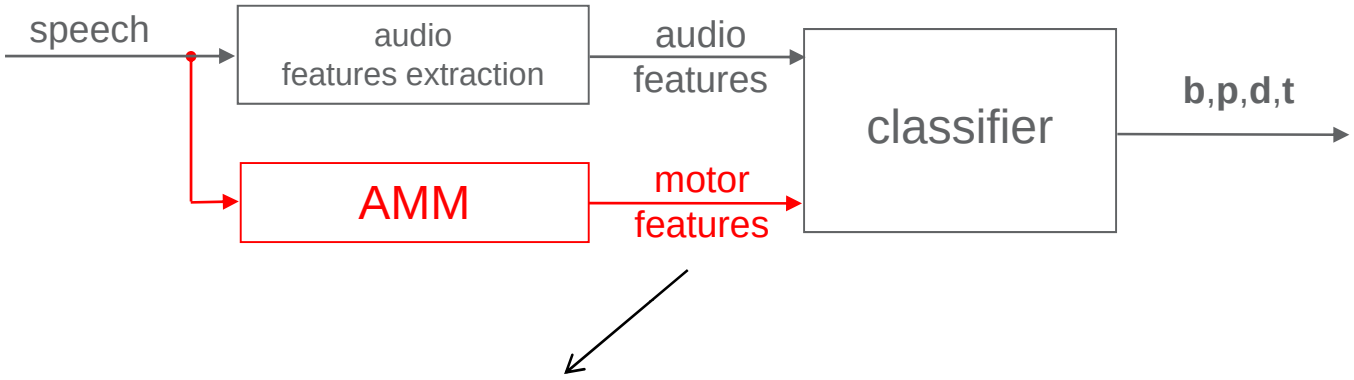


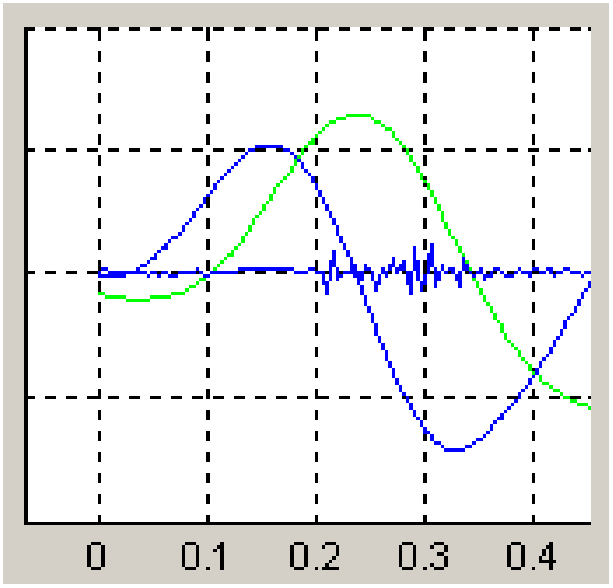
Stimulation

Results


130
120
110
100
90
80
TMS-NoTMS RT%
Labial
Dental
*
*
*
*
Tongue M1
Lips M1
Acoustic stimulus
TMS site
15
10
5
0
-5
-10
TMS-NoTMS Error Bias %
L-Ph-mlss
T-Ph-mlss
*
*
Tongue M1
Lips M1

Motor feature based recognition



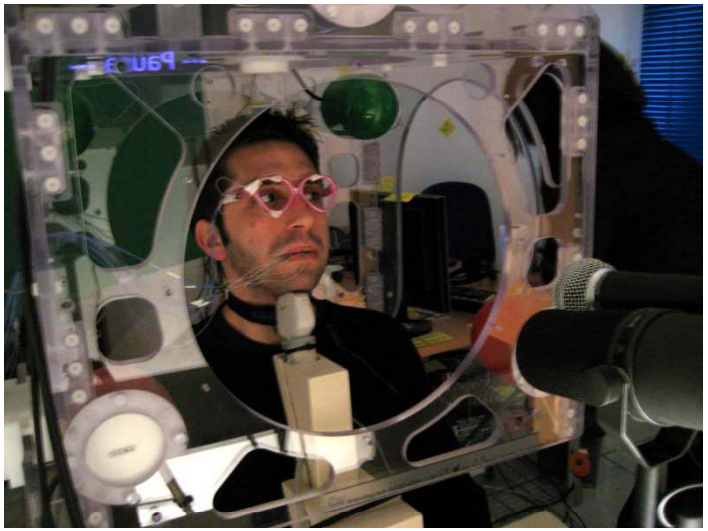


green: lips opening velocity

blue: lips opening acceleration

grey zone: the identified motor invariant for **b**

Data collection

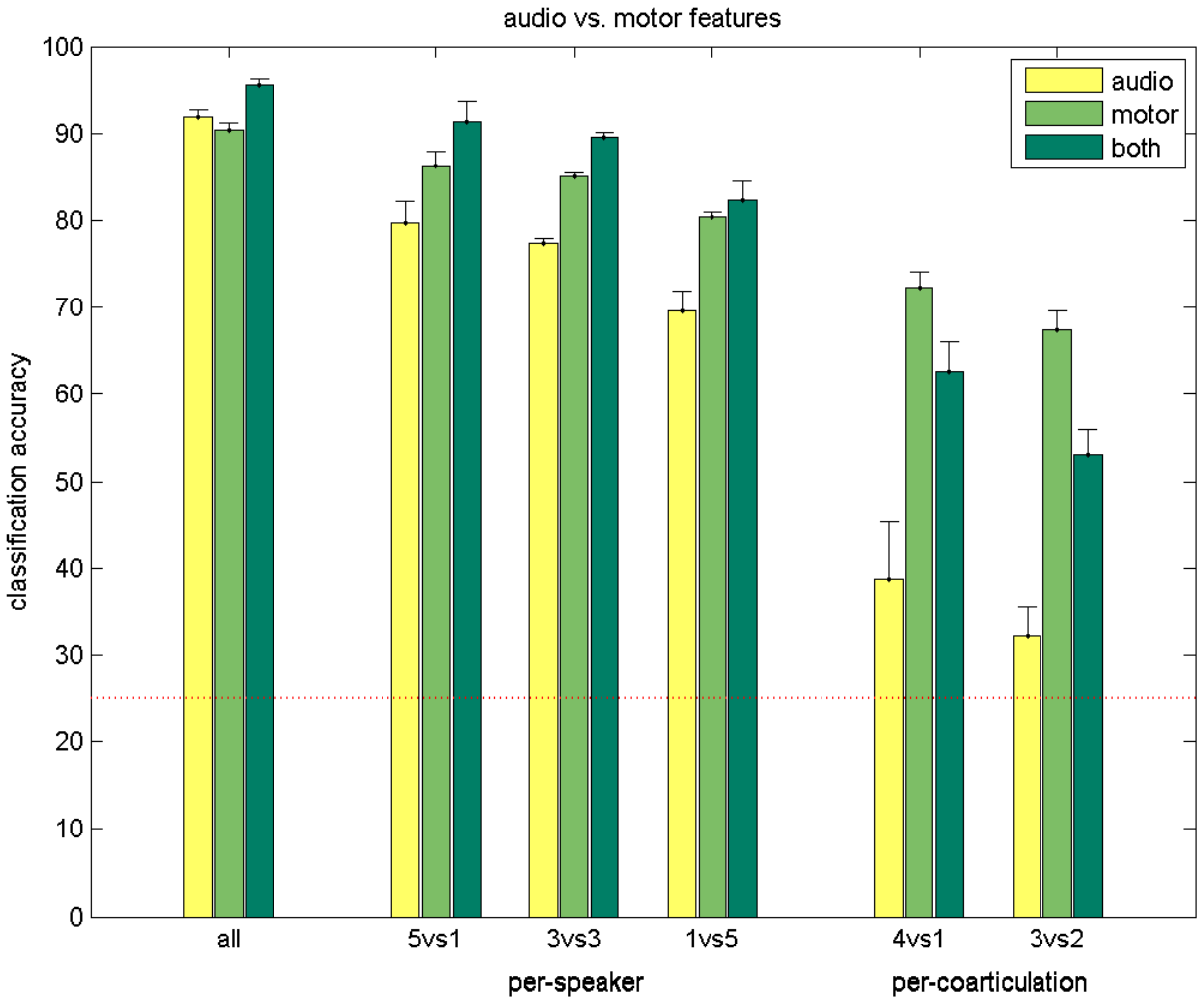


- 9 speakers, 74 (pseudo)words and syllables
- magnetic tracking of tongue, lips and teeth
- ultrasound imaging of tongue
- video of face
- laryngography of vocal folds

THE CONTACT PROJECT

2005-2009

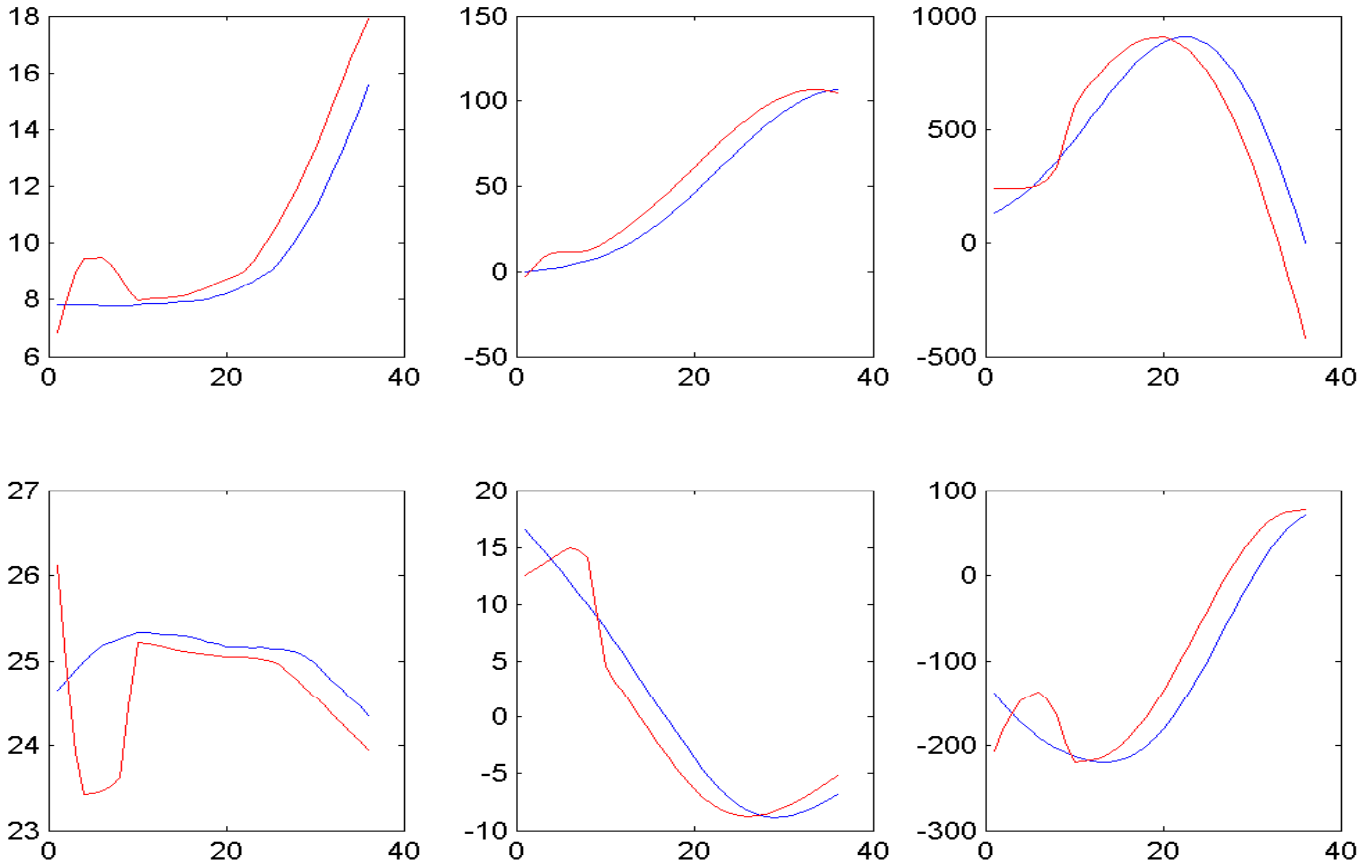
Baseline experiment



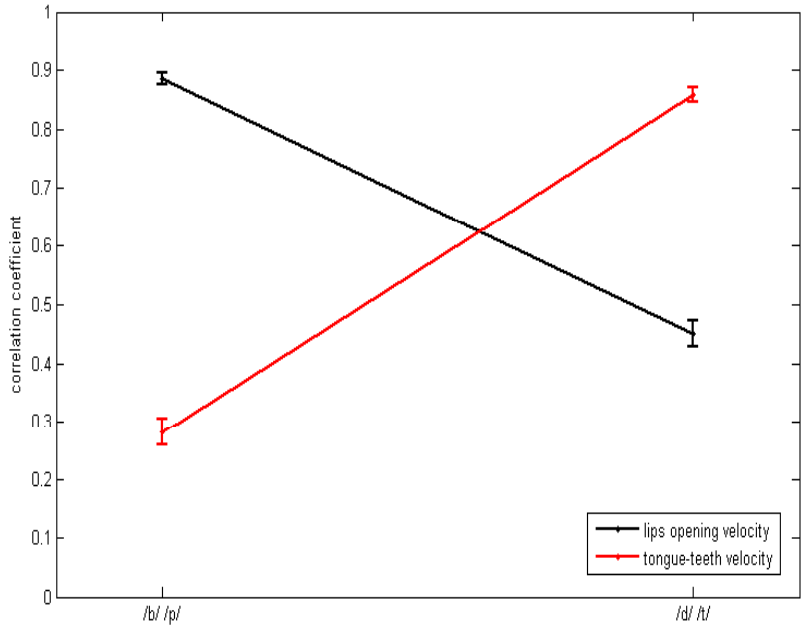
Audio-motor map

- Training the AMM:
  - *input space:* 200ms. Mel-scale spectrogram (20 filters) of speech ($\mathbf{R}^{380}$)
  - *output space:* point-by-point *VliO*, *AliO*, *VttU*, *AttU* over utterance ($\mathbf{R}^{4}$)
  - ANN w/ sigmoidal activation function, cross-validation, regularization, 10 random restart (the best is stored)
- Cross-validation:
  1. over all utterances
  2. per-speaker

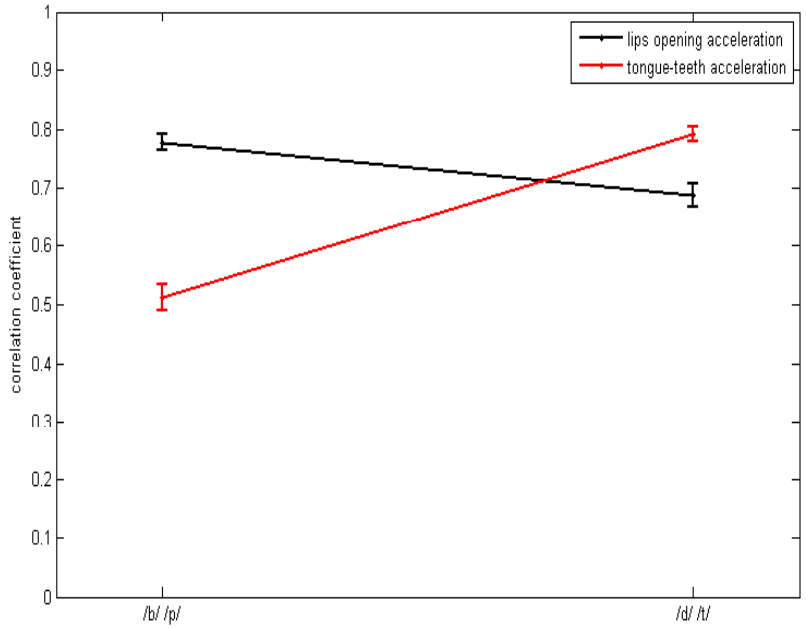
papa


18
16
14
12
10
8
6
0
20
40
150
100
50
0
-50
0
20
40
1000
500
0
-500
0
20
40
27
26
25
24
23
0
20
40
20
15
10
5
0
-5
-10
0
20
40
100
0
-100
-200
-300
0
20
40

Audio-motor map

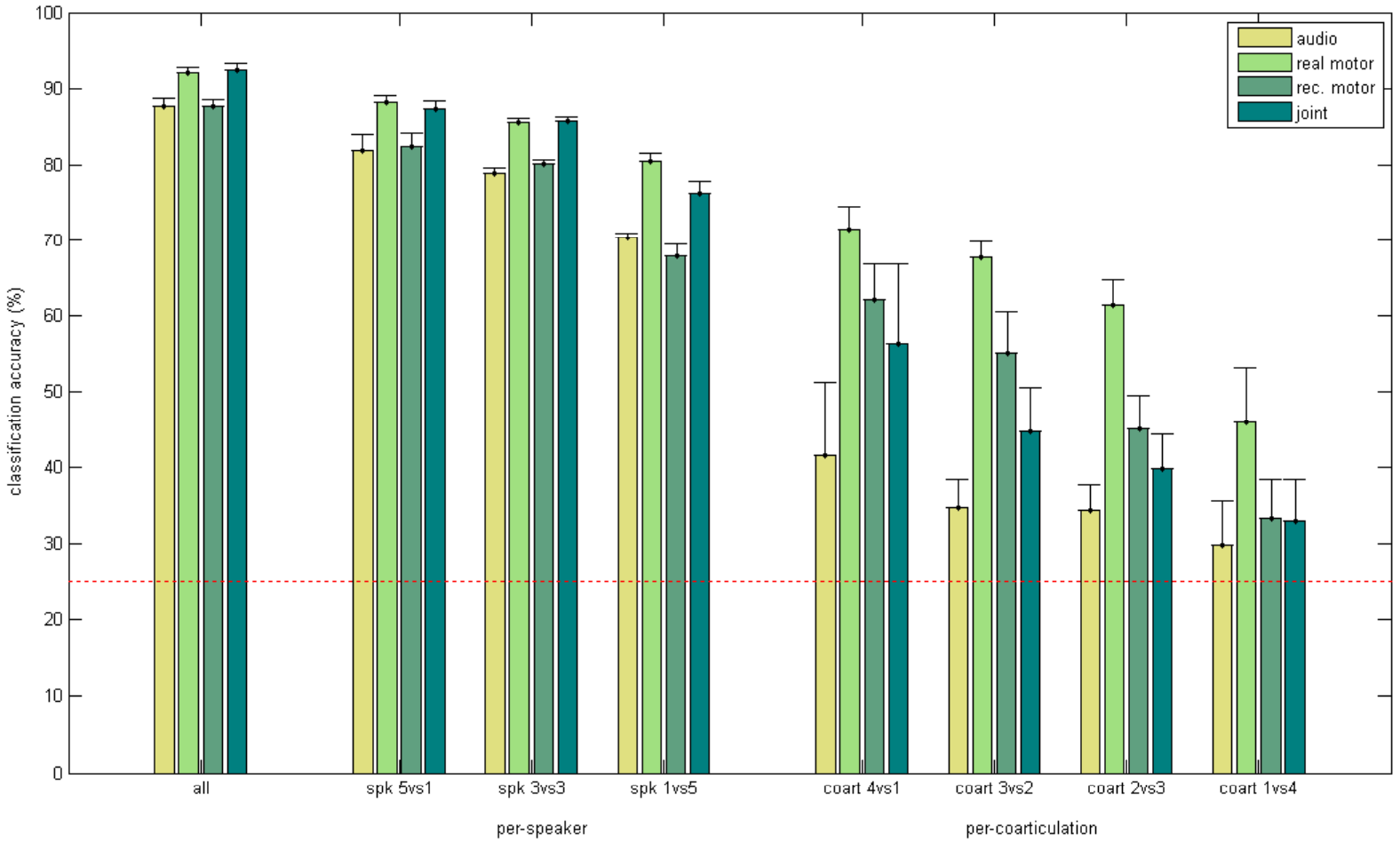


Velocity

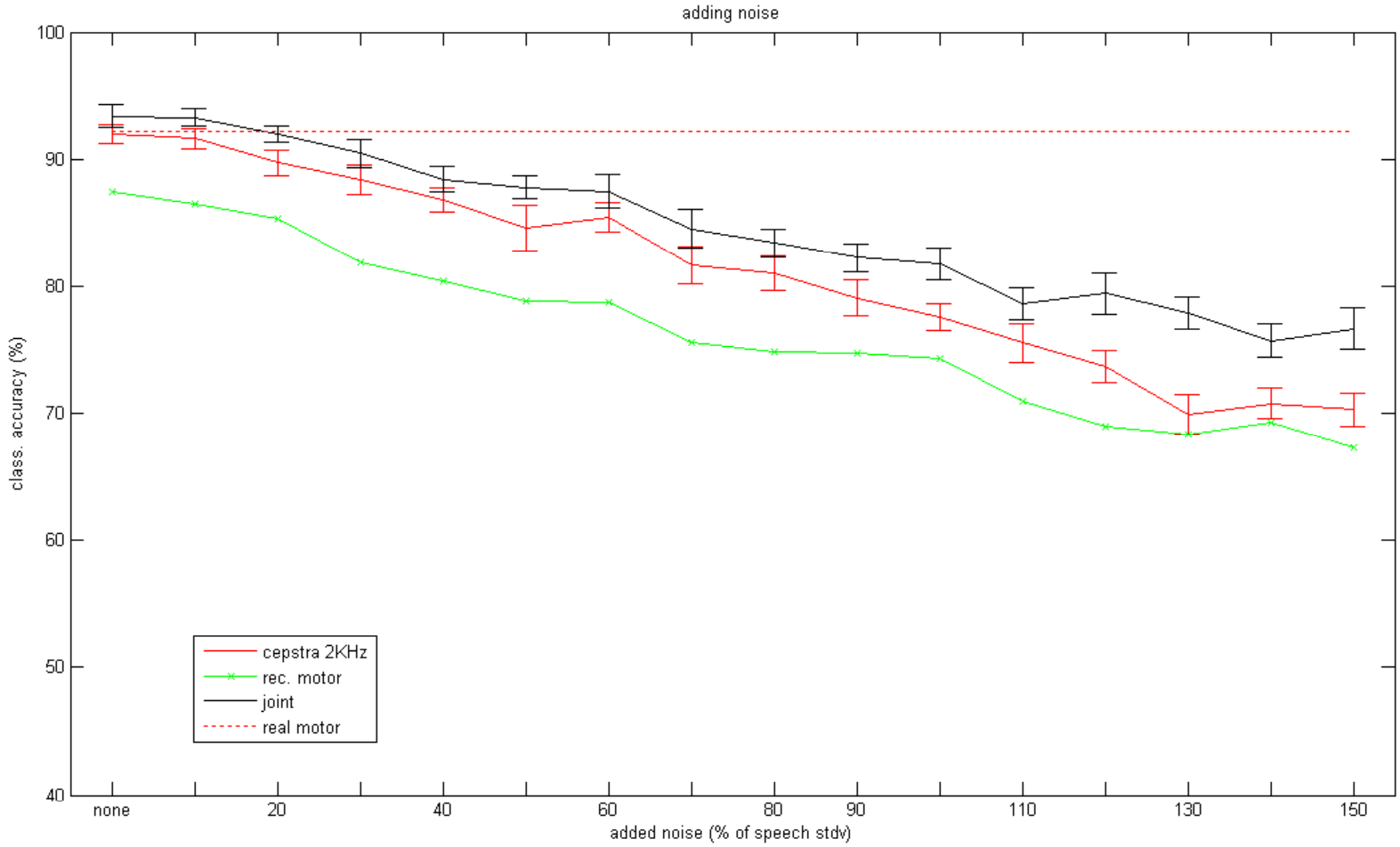


Acceleration

With reconstructed motor signals

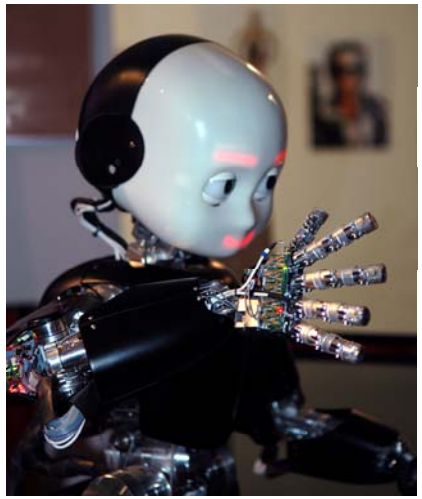


Increasing noise



Conclusions

- The brain uses motor information as “perceptual invariants”
- It might be advantageous to copy this solution in artificial systems
- …which ultimately require a body to generate sensorimotor patterns autonomously (there’s always an excuse to build a humanoid robot)



Sponsors

- EU Commission projects:
  - RobotCub, grant FP6-004370, http://www.robotcub.org
  - CHRIS, grant FP7-215805, http://www.chrisfp7.eu
  - ITALK, grant FP7-214668, http://italkproject.org/
  - Roboskin, grant FP7-231500, http://www.roboskin.eu
  - Poeticon, grant FP7-215843 http://www.poeticon.eu

- More information: http://www.cognitivehumanoids.eu

Cogsys
Cognitive Systems

RobotCub