DComm
Deictic Communication

MANCHESTER 1824
The University of Manchester

Learning to Talk in a Gesture-Rich World: Application in Cognitive Robotics

Gabriella Pizzuto

Cognitive Robotics Lab
Machine Learning and Optimisation Research Group

Developmental Neural Networks Workshop
ALIFE 2019

European Commission
Horizon 2020
European Union funding
for Research & Innovation

MARIE CURIE ACTIONS

Motivation

What are the main questions?

(1) How can an **agent** develop **language** with the help of **deictic gestures**?

(2) How can an **agent** transition from **single word** to **multi-word**?

(Extracted Frames From Video 1 Source) Child using gestures to communicate

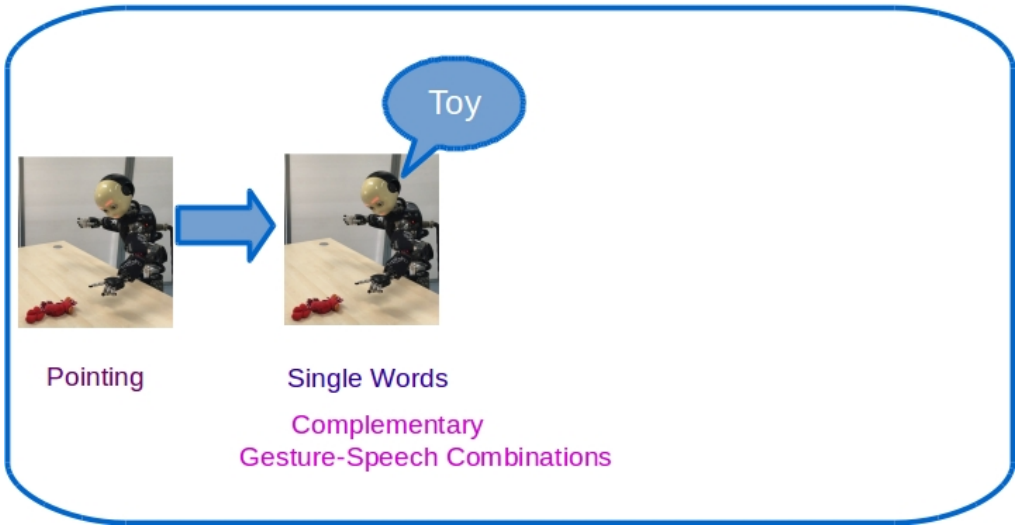
(Extracted Frames From Video 2 Source) Observation: Speech Milestones at 2 years

Early Language Development-An Overview

Pointing

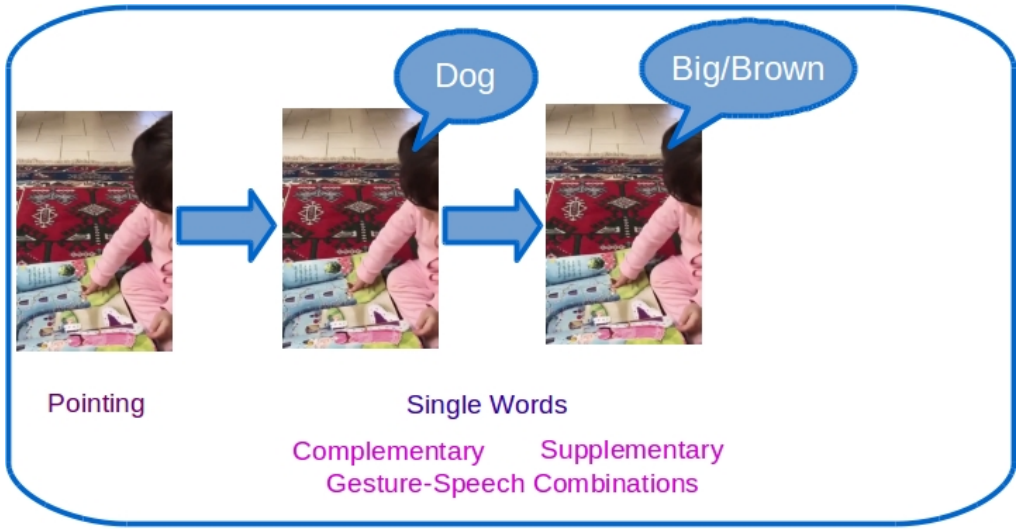
Pointing

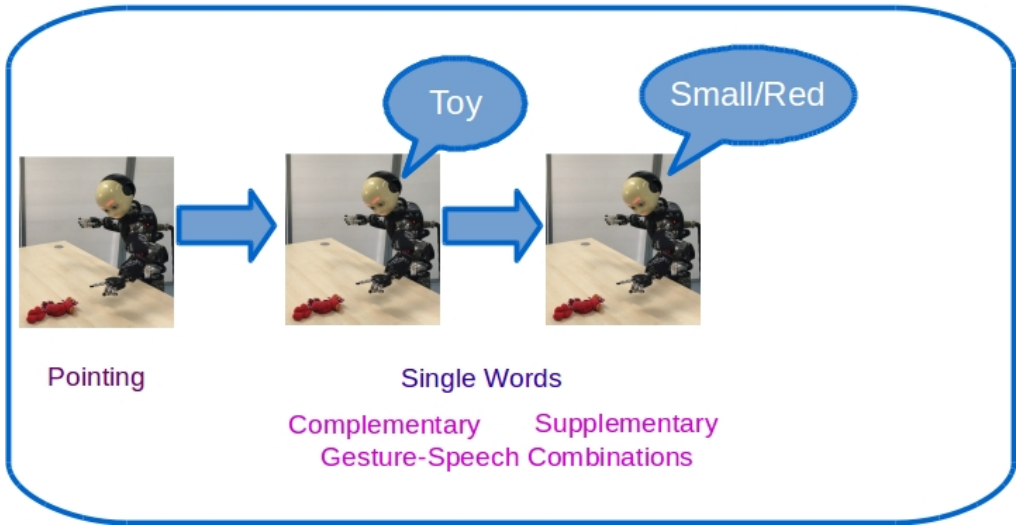
Early Language Development-An Overview



Gestures that convey information **reinforcing** the information conveyed in the accompanying speech.

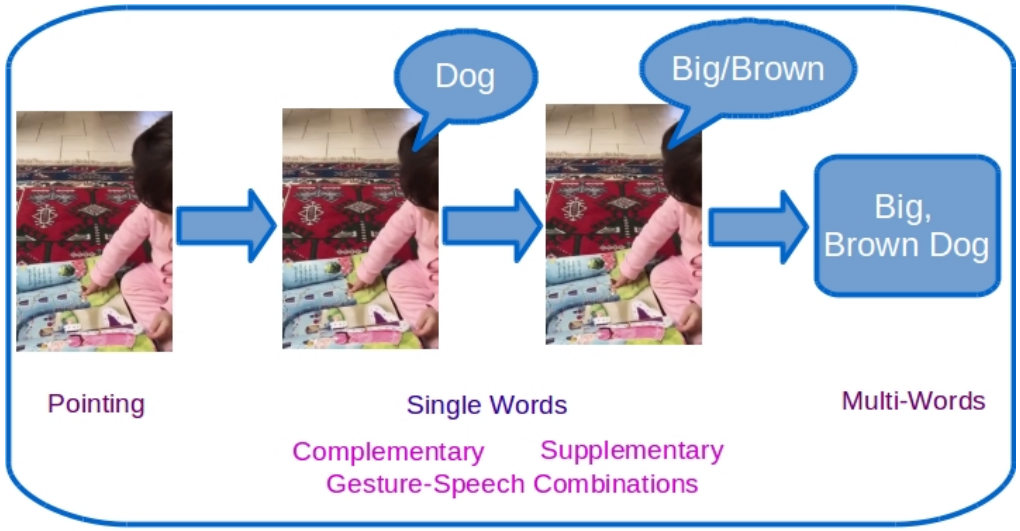
Early Language Development-An Overview

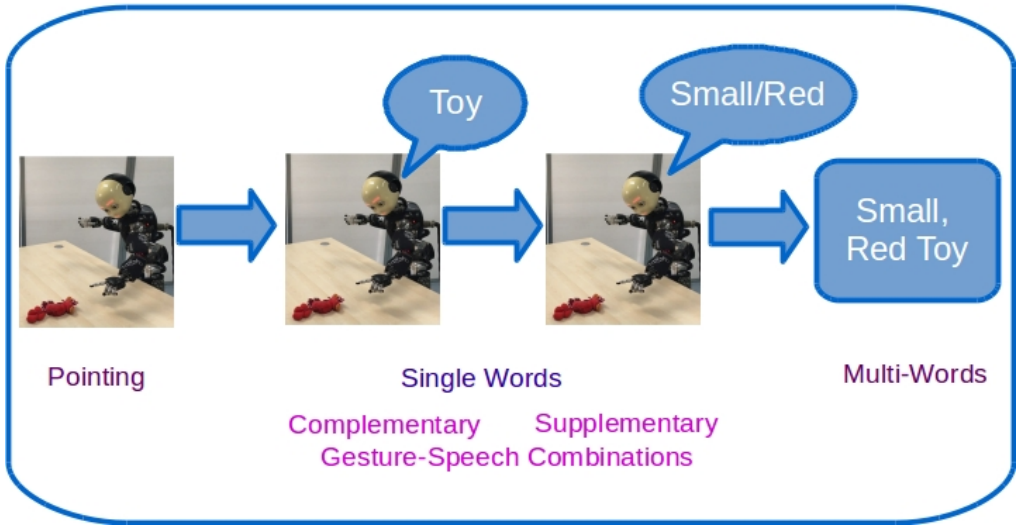




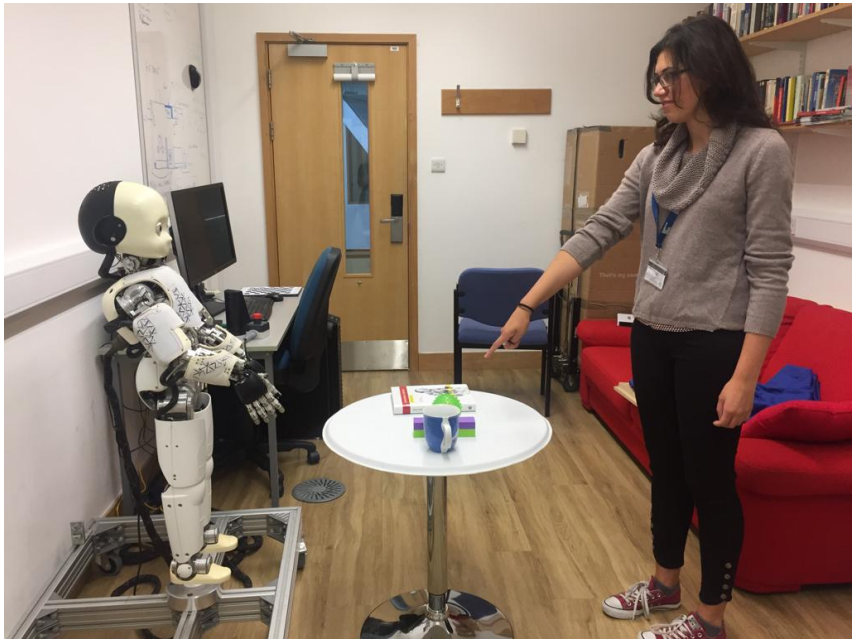
The gesture creates a two-unit construction by **adding a new semantic element** to the meaning.

Early Language Development-An Overview





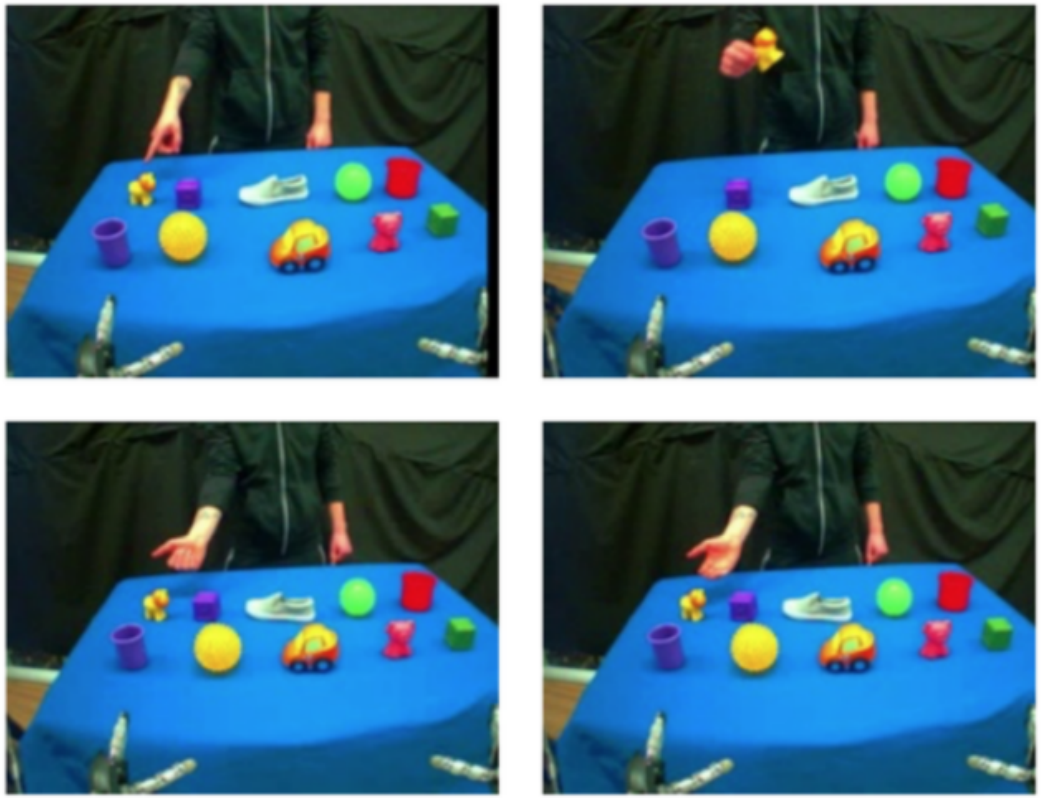
Problem Definition



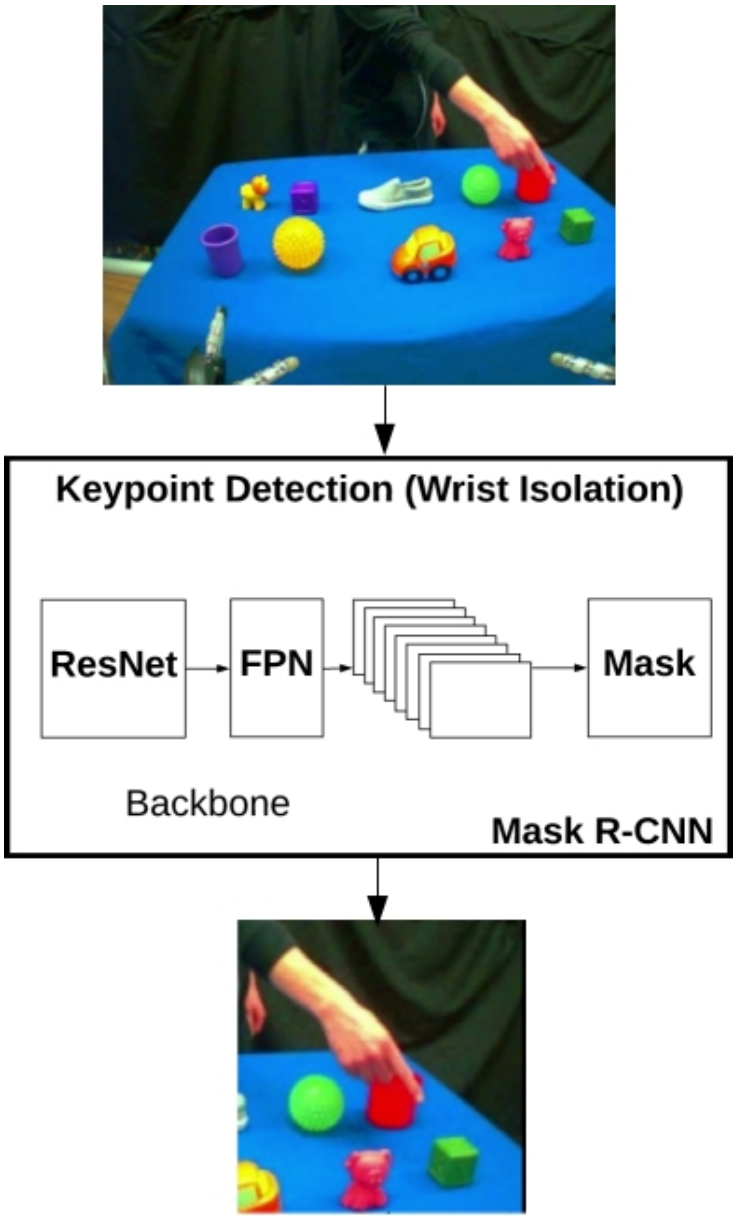
- 2-stage cascaded Mask-RCNN as the core of the model (wrist keypoint and object detection)
- Deploy model on the iCub for complementary gesture-word comprehension

iCub's Gesture-Object Dataset

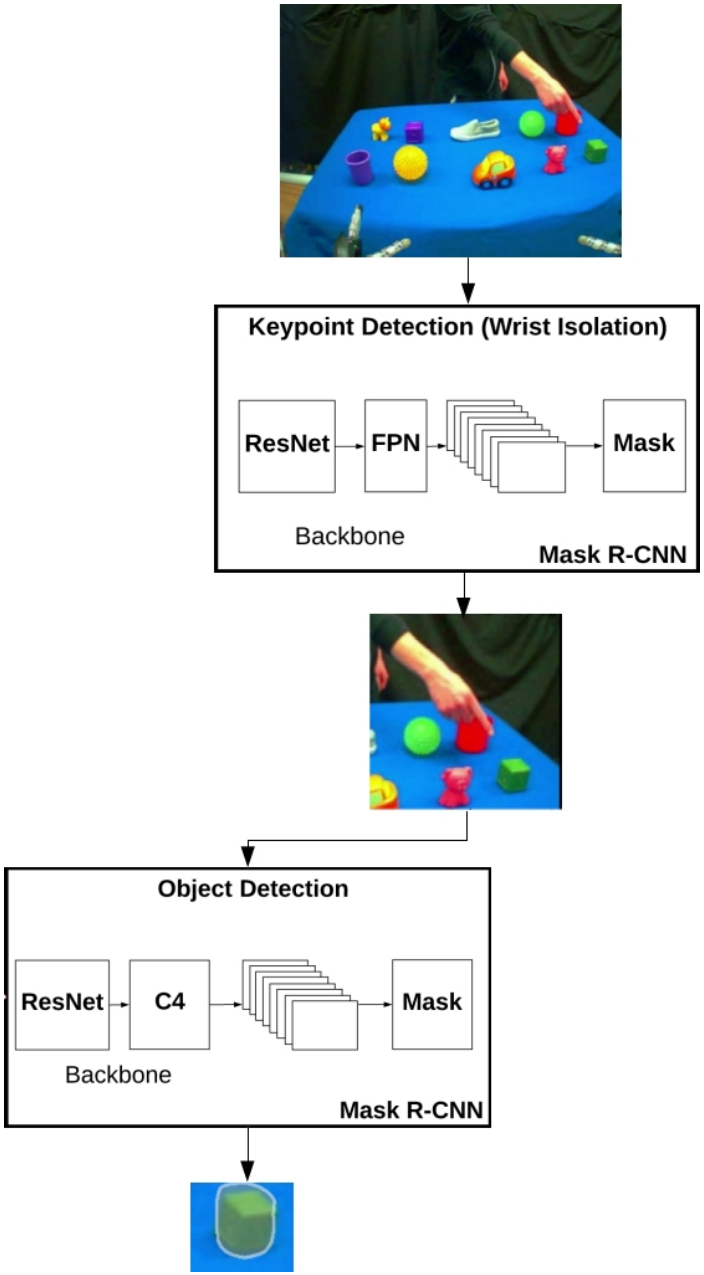
- Dataset collection with the iCub (20 participants)
- Partially annotated for RCNN



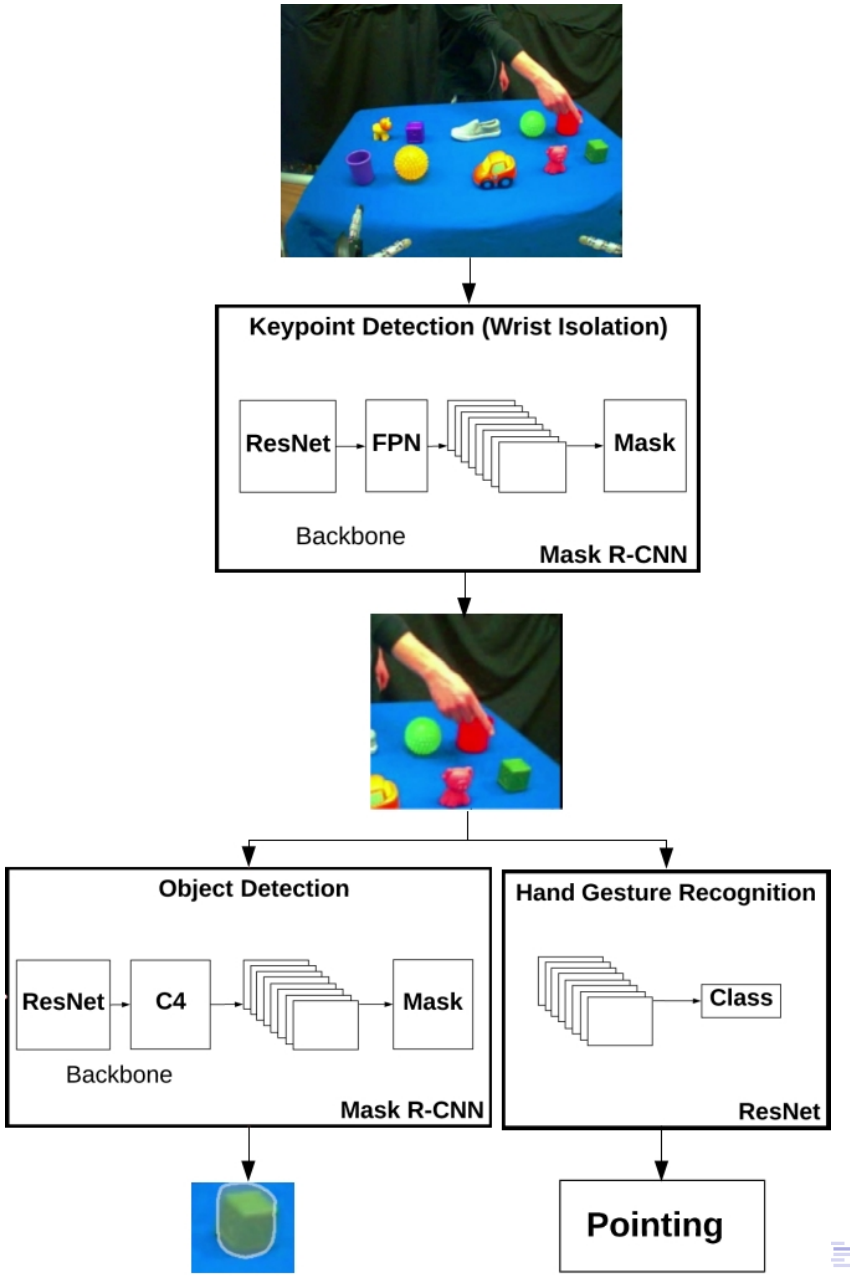
The Gesture-Object Detection System



The Gesture-Object Detection System



The Gesture-Object Detection System

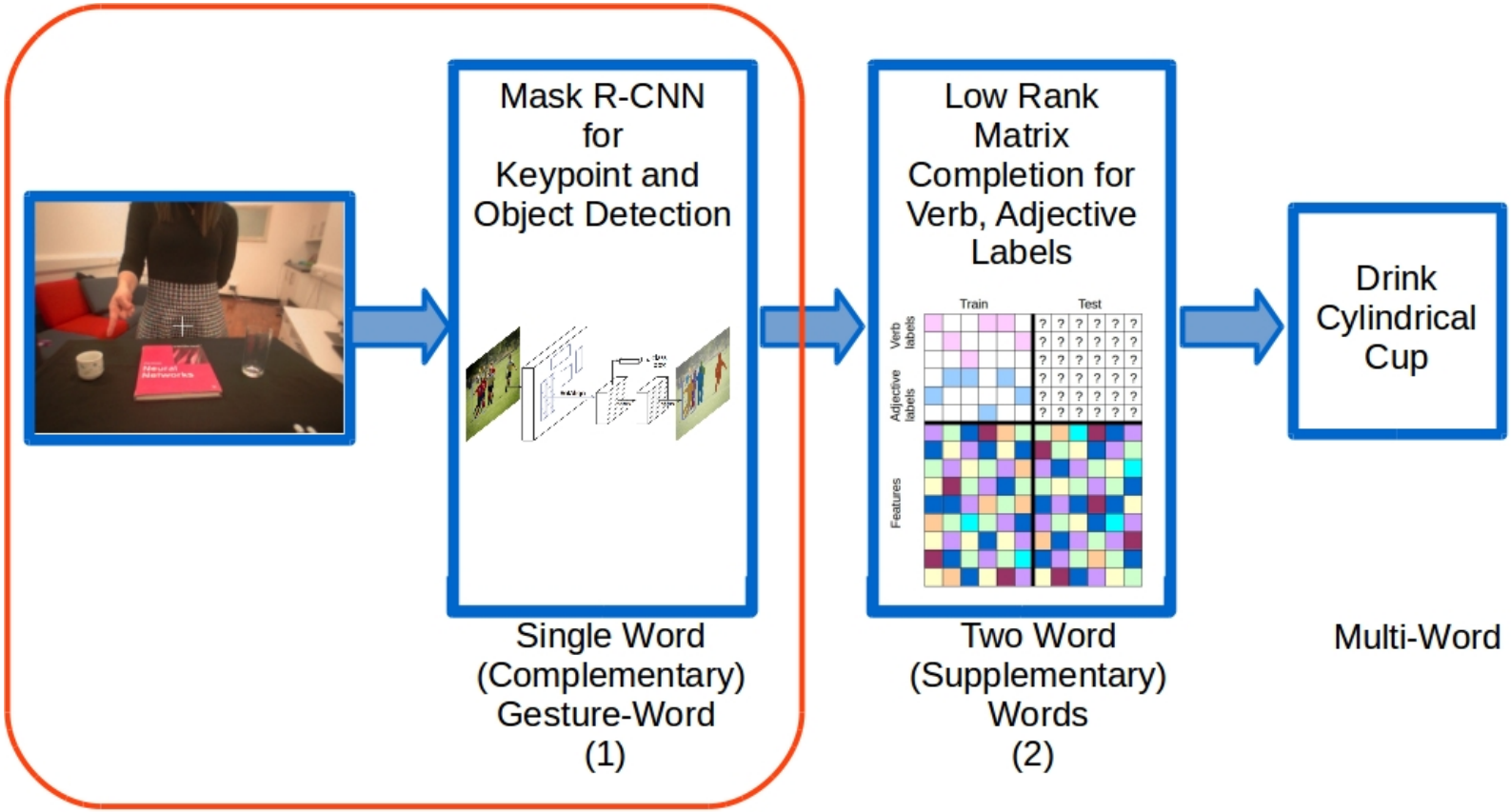


Deployment of the Gesture-Object Detection System on the iCub

Aim: To illustrate the implementation of the gesture-object detection system on the iCub, focusing on having the iCub understand complementary gesture-word combinations.

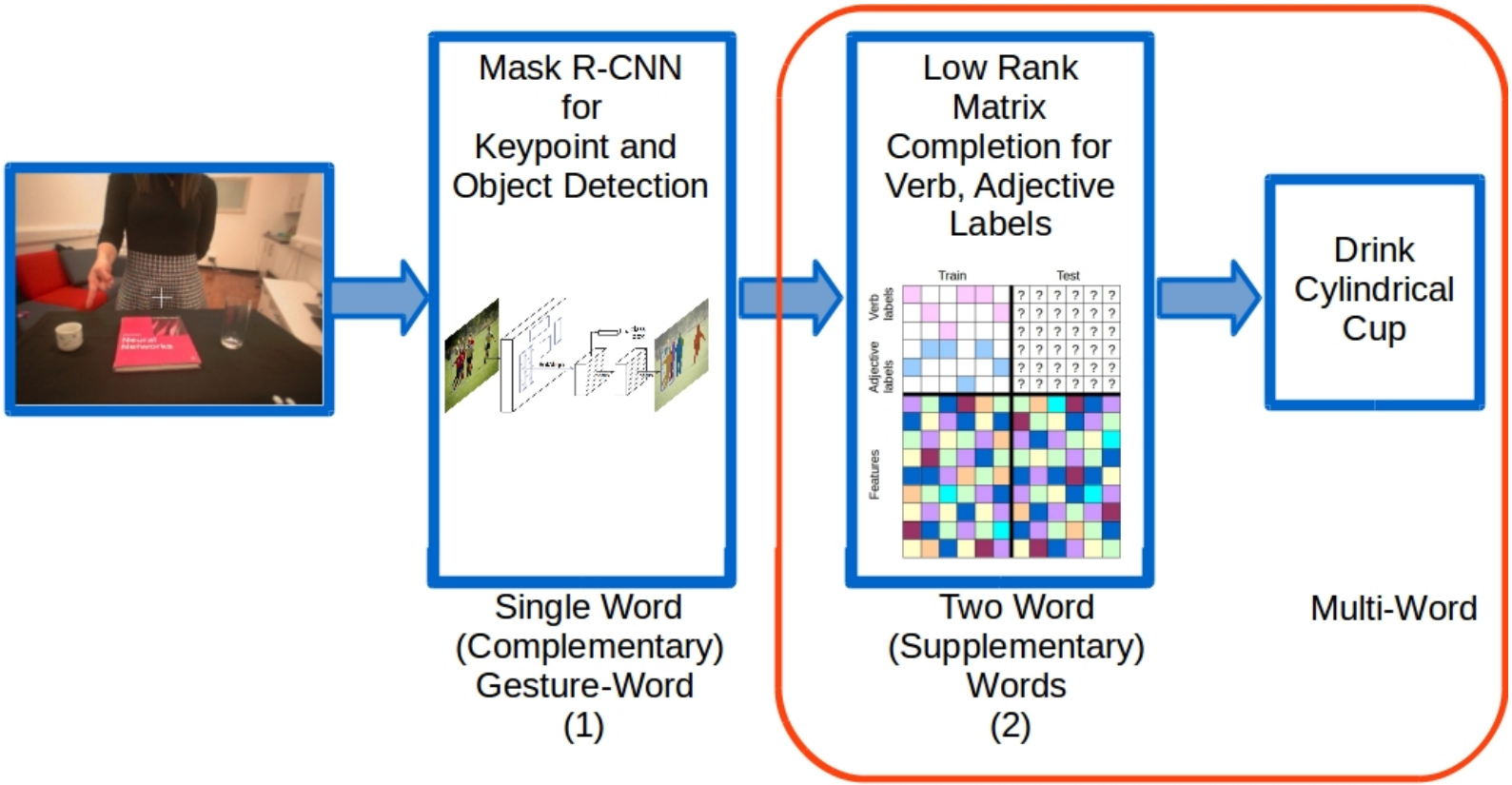
- The deictic gesture helps to obtain the mask for the object of interest
- Even with previously unseen objects, the system generalises well

A Computational Model For Gesture-Word Combinations



(1) G. Pizzuto and A. Cangelosi, *"Exploring Deep Models for Comprehension of Deictic Gesture-Word Combinations in Cognitive Robotics"*, IJCNN 2019.

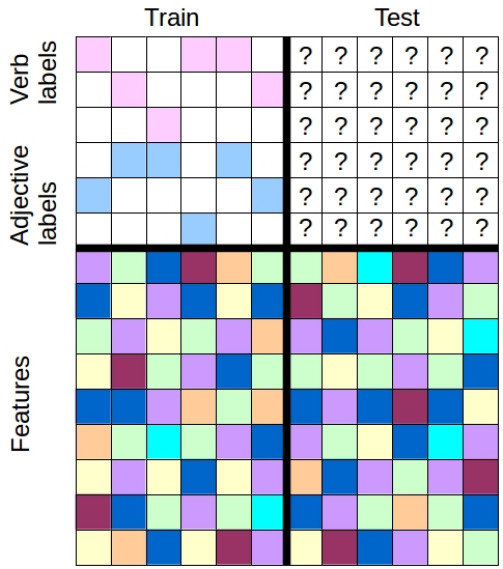
A Computational Model For Gesture-Word Combinations



(1) G. Pizzuto and A. Cangelosi, *"Exploring Deep Models for Comprehension of Deictic Gesture-Word Combinations in Cognitive Robotics"*, IJCNN 2019.
(2) G. Pizzuto, T. Hospedales, O. Capirci and A. Cangelosi, *"Modelling the Single Word to Multi-Word Transition Using Matrix Completion"*, under review.

Matrix Completion For Supplementary Word Generation

Idea: ***Missing supplementary labels of objects can be inferred from those with similar features.***



$$\underset{\mathbf{Y}_{tst}}{\text{minimize}} \quad \text{rank}(\mathbf{Z})$$

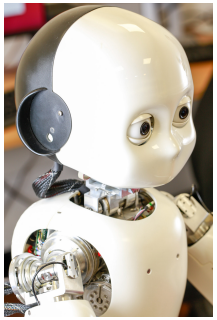
$$\text{subject to} \quad \mathbf{Z} = \begin{bmatrix} \mathbf{Y}_{tr} & \mathbf{Y}_{tst} \\ \mathbf{X}_{tr} & \mathbf{X}_{tst} \\ \mathbf{1}^T & \end{bmatrix} \tag{1}$$

Takeaway Message

(1) **Infant developmental stages** as an inspiration for **computational and robotic models**.
(2) **Machine learning** methods can be applied to computational **models of early language acquisition**.

The Team



Related Papers

(1) G. Pizzuto and A. Cangelosi, *"Exploring Deep Models for Comprehension of Deictic Gesture-Word Combinations in Cognitive Robotics"*, IJCNN 2019.
(2) G. Pizzuto, T. Hospedales, O. Capirci and A. Cangelosi, *"Modelling the Single Word to Multi-Word Transition Using Matrix Completion"*, ICDL-EPIROB 2019.