

POETICON

The Poetics of Everyday Life: Grounding Resources and Mechanisms for Artificial Agents



Katerina Pastra







Cognitive Systems Research Institute (CSRI: www.csri.gr)

& Institute for Language & Speech Processing (ILSP),

ATHENA Research Center, Athens, Greece

POETICON

January 2008 – December 2010 (ext. April 2011)
www.poeticon.eu

Participant organisation name		Country
Institute for Language and Speech Processing / "Athena" Research Centre	ILSP 	Greece
University System of Maryland Foundation, University of Maryland College Park - Institute for Advanced Computer Studies	USMF 	USA
Univerza v Ljubljani – Visual Cognitive Systems Laboratory	UL 	Slovenia
Max Planck Society – Max Planck Institute for Biological Cybernetics	MPG 	Germany
Istituto Italiano di Tecnologia – Department of Robotics, Behaviour & Cognitive Science	IIT 	Italy
University of Ferrara – Department of Neuroscience	UNIFE 	Italy

POETICON Research Questions



How is language related to perception and action?

Do they share a common 'syntax'?



Experiments on the neuroscience of action (perception and generation) and language (Fadiga et al. 2009, Fazio et al. 2009, Clerget et al. 2009);

If language, perception and action share a common syntax, what is this common structure like, and how do they interact based on this common structure?

The minimalist grammar of action (Pastra and Aloimonos 2012)

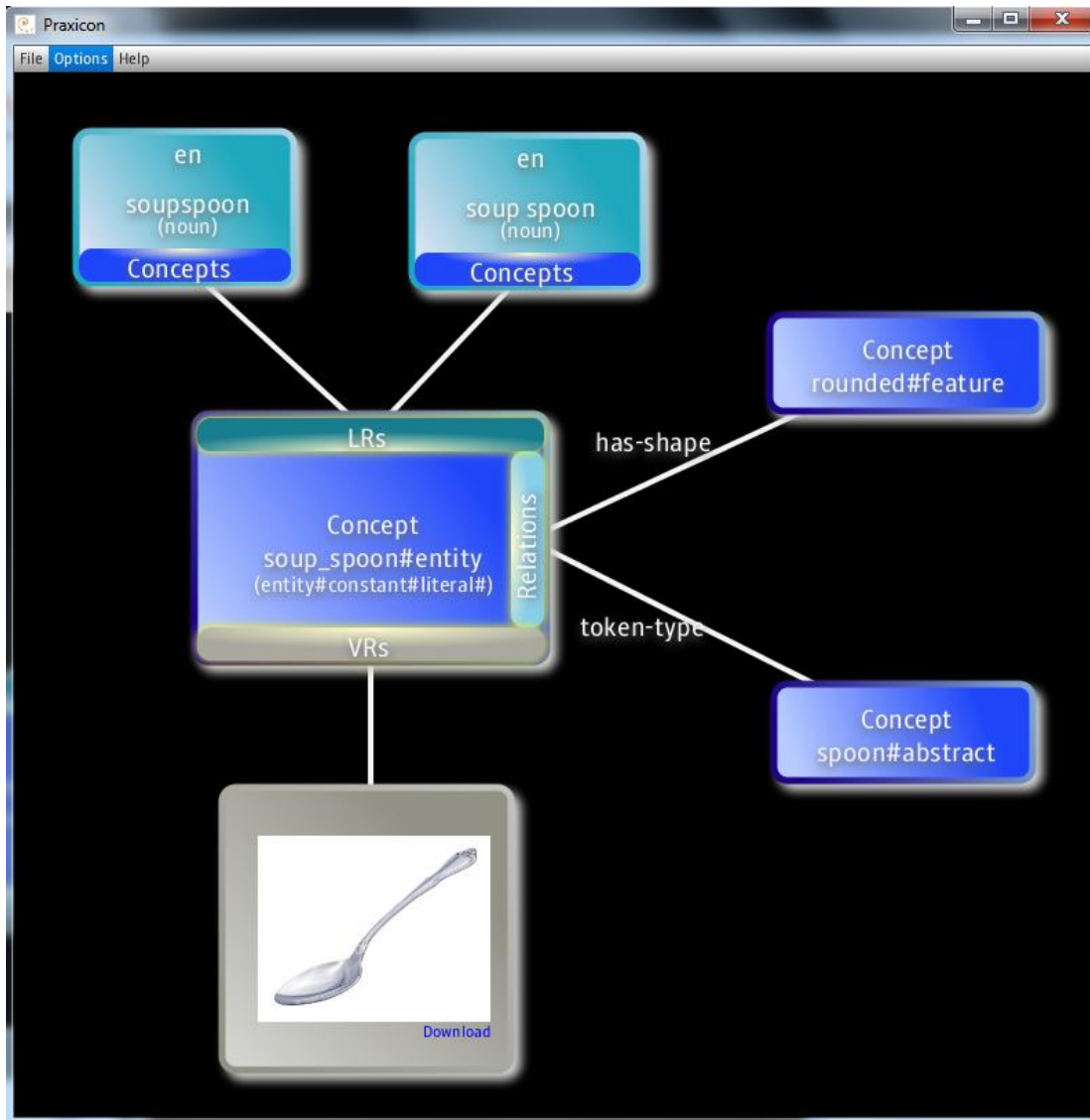
How could one achieve such integration computationally?

POETICON basic tools: Visual, motoric, language, reasoning and manipulation skills; POETICON everyday interaction scenarios corpus and PRAXICON

Can the POETICON tools and knowledge base assist an agent to generalise over learned schemas and set behaviors?

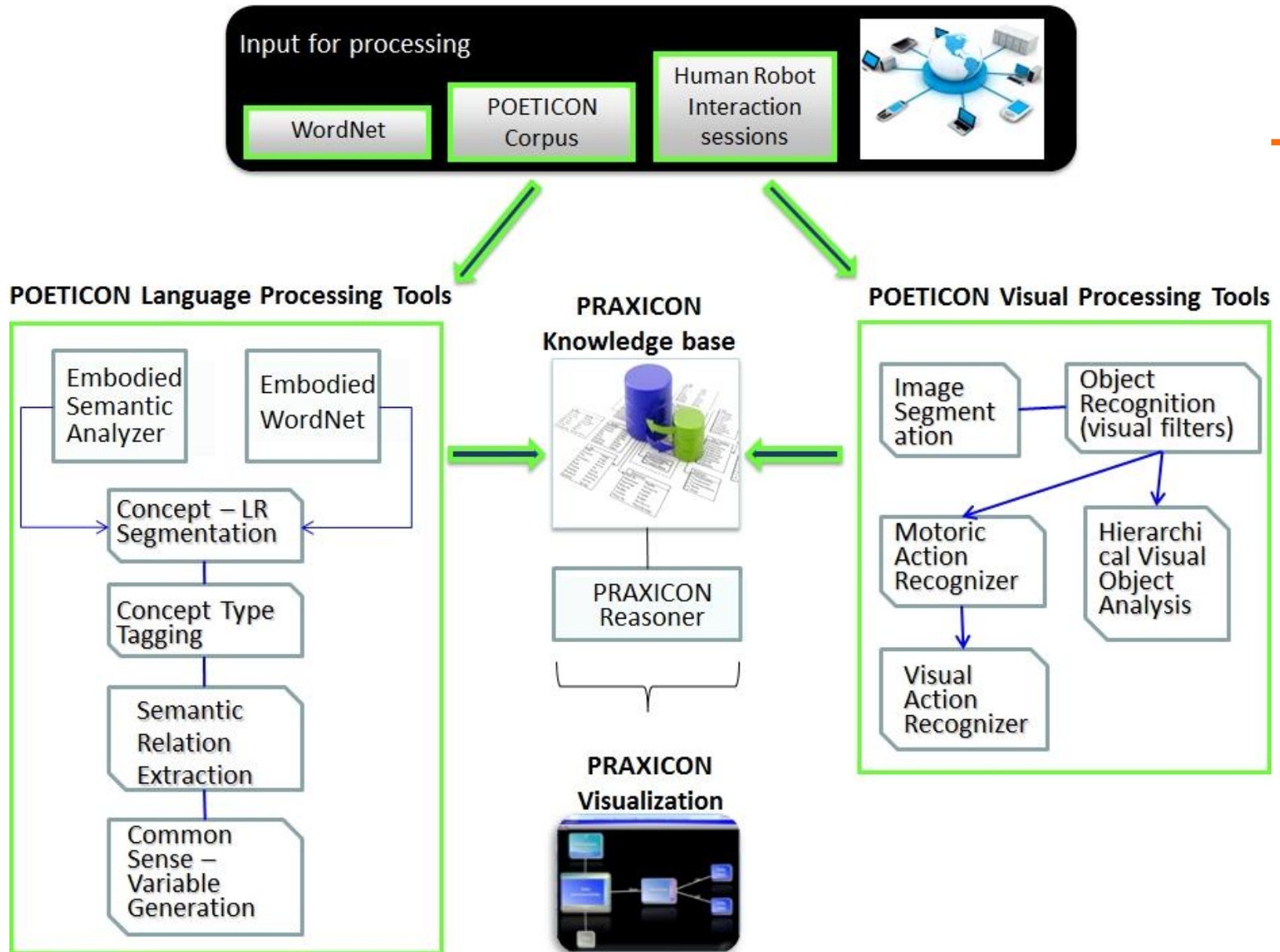
POETICON robotic demonstration

The objective: a computational PRAXICON



- PRAXICONs: From Liepman's (1908) input/output motor representations stored in memory, to...
- ...embodied-concept representations of perceptual, motoric and/or linguistic/symbolic nature, perceived and stored in memory for behaviour generation and understanding

The POETICON Tools



POETICON Final Integrated Demo

From verbal instructions to action generation

The image is a composite illustrating the POETICON system's process. At the top, a conceptual diagram shows the flow from a verbal instruction to action generation. It features several blue boxes representing concepts and their relationships:

- A central box: `Concept stir_with_dummy_object_the_coffee#movement (movement#)`
- A box to its right: `Intersection`
- A box above the intersection: `Concept dummy_object#entity (entity#)`
- A box to the right of the intersection: `Concept spoon#abstract`
- A box to the right of the intersection: `Concept stir#abstract`
- A box below the intersection: `Concept java#abstract`

Relationships are indicated by lines and labels:

- `stir_with_dummy_object...` is **Inherent** to `Intersection`.
- `Intersection` is **use-artifact** of `Concept dummy_object#entity`.
- `Intersection` has an **action-purpose** of `Concept stir#abstract`.
- `Intersection` has an **action-object** of `Concept java#abstract`.
- `Concept dummy_object#entity` **has_value** `Concept spoon#abstract`.

Below the diagram, a robot is shown holding a white coffee cup. The text **"Stir the coffee..."** is displayed next to it. To the left, a sign reads **poeticon** and *The Poetics of Everyday Life*. In the foreground, a wooden table holds a coffee-making scene with a moka pot, a silver shaker, and two coffee cups.

POETICON Public Release (1)

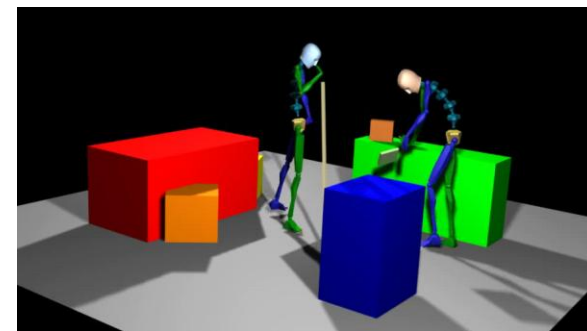
The POETICON Multimodal and Multisensory Corpus

<http://poeticoncorpus.kyb.mpg.de/>

Language use and sensorimotor experience recordings in everyday human:human interaction:

- spontaneous language communication (EN)
- multiview video recordings (5 cameras)
- recording of 3D full body kinematics and 3D tracking of objects in focus.

The corpus comprises 6 everyday interaction scenes between couples, 12 hours in total. It has been enhanced with integrated 3D animations of the full body kinematics and object tracking.



POETICON Public Release (2)

The POETICON Audiovisual Cognitive Experiments Corpus

Corpus of the POETICON cognitive experiments on grounding, i.e. experiments that explore the type of information, and the verbalization strategies used by humans for naming objects and actions:

- Visual and tactile stimuli
- ‘unknown objects’ (lithic tools)
- verbal reports (EL)
- 120 participants
- 75 hours of recordings
- Fully transcribed and semantically annotated
- Fully processed automatically for extraction of relational information and making inferences



POETICON Public Release (3)

The PRAXICON Knowledge Base, its visual exploration interface (GUI) and the integrated language analysis and reasoning tools

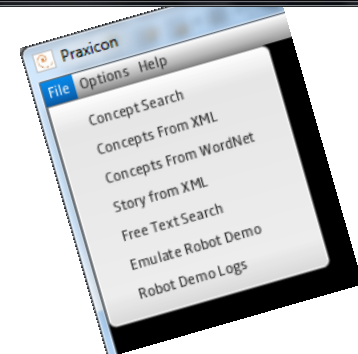
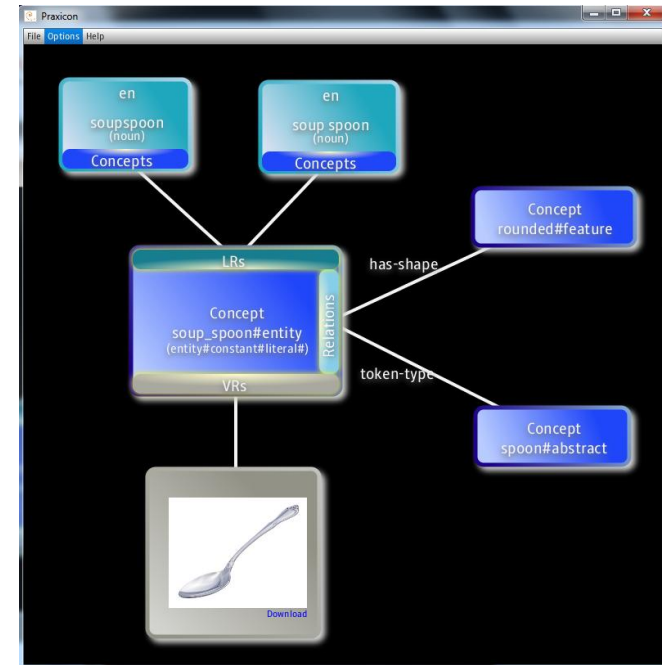
In two forms:

- as a web service (database and game)
- as a downloadable, standalone application for local installation.

Contents:

- Part of the Embodied WordNet Lexical Database and
- Corresponding visual representations from the ImageNet database.

This version of the PRAXICON will contain more than 5.000 concepts and more than 4.500 relations among them.



POETICON++







Robots need Language: A computational mechanism for generalisation & generation of new behaviours in robots



Katerina Pastra

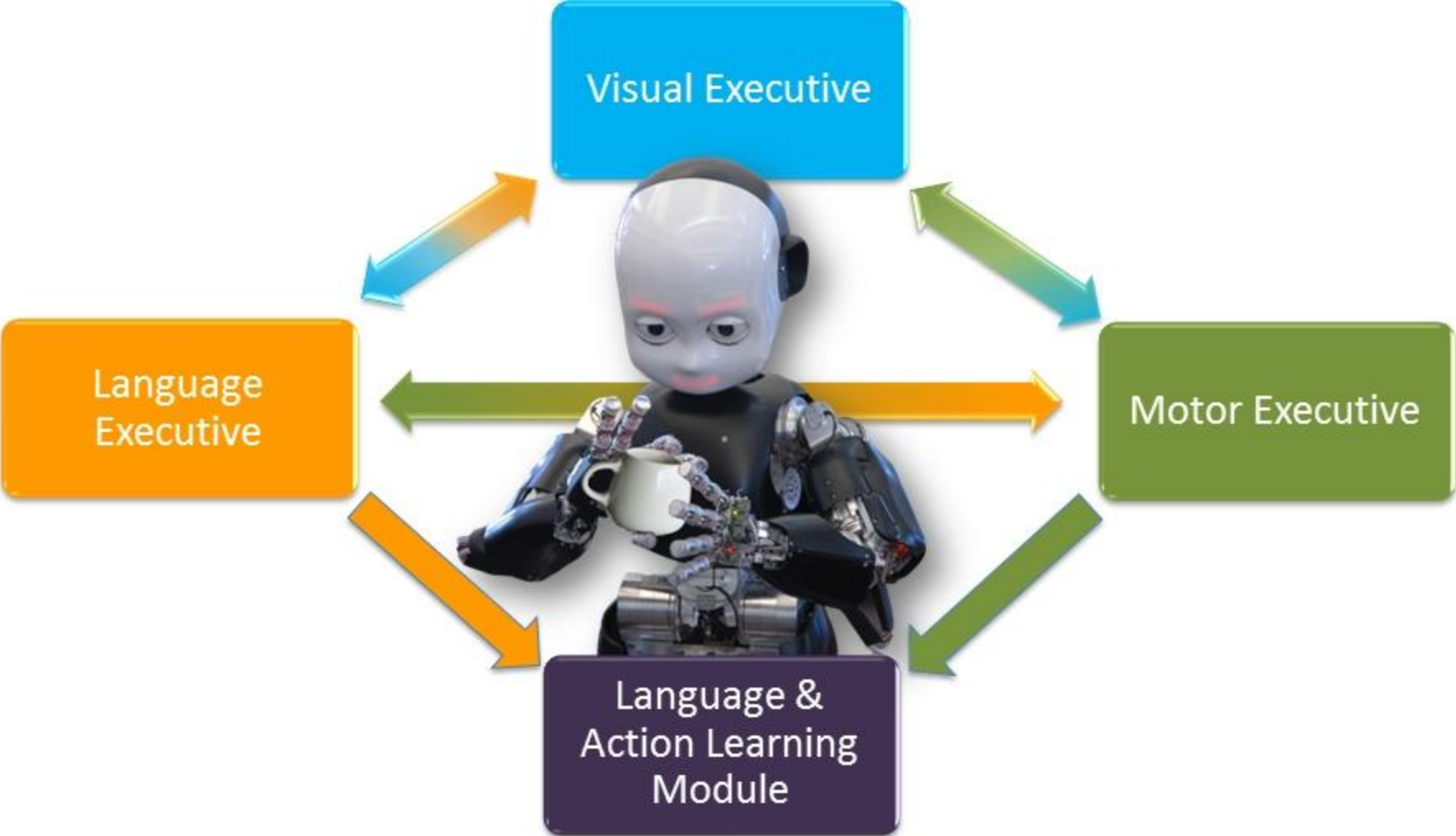
Cognitive Systems Research Institute (CSRI) &
Institute for Language & Speech Processing (ILSP),
ATHENA Research Center, Athens, Greece

POETICON++ Consortium

Participant organisation name	Part. short name		Country
Istituto Italiano di Tecnologia Dept. of Robotics, Brain & Cognitive Science	IIT		Italy
ATHENA Research Center Institute for Language and Speech Processing	ATHENA R.C	 Ερευνητικό Κέντρο Αθηνά Ερευνητικό Κέντρο Καινοτομίας στις Τεχνολογίες της Πληροφορίας, των Επικοινωνιών, της Γλώσσας	Greece
Cognitive Systems Research Institute Embodied Language Processing Group	CSRI		Greece
University System of Maryland Foundation University of Maryland College Park - Computer Vision Lab	USMF		USA
University of Plymouth Center for Robotics & Neural Systems	UoP		UK
Instituto Superior Tecnico Computer and Robot Vision Lab	IST		PT

POETICON ++

The active Interpreter



POETICON++



Hypothesis & Objective

Natural language may be used as a **learning tool** for **generalisation** of learned behaviours, and generation of new ones (creativity); it may allow an agent to choose the optimal solution to be followed given a particular situation, and know when ingenuity stops and wrong generalisation starts.

Objective:

development of a computational mechanism for generalisation and creative generation of new behaviours and perceptions in artificial agents through interaction in dynamic real-life environments.

POETICON++



Objectives (in detail)



- To **develop** a cognitively-plausible computational learning model and mechanism for language-mediated behaviour generalisation and creativity in interactive cognitive systems;
- To **demonstrate**, through humanoid robotic experiments, the feasibility and validity of this model in two dynamic, real-life scenarios: behaviour generation through verbal instruction, and visual scene understanding.
- To **advance our understanding** of the neuroscientific, cognitive and linguistic phenomena and mechanisms supporting the generalisation and creation of new behaviour through the essential role of language and action-language hierarchical representations;

POETICON++



Experimental and Theoretical Investigation



- (a) Exploration of the *neural activation of the motor system during generation of structured action*
- (b) Exploration of *the role of the motor system in speech perception*
- (c) Exploration of the optimal developmental training protocol for action-language learning in humanoids
- (d) Experimentation with employing a neuroanatomical model of language and action learning.
- (e) Exploration of the use of language for object, action and attribute naming

POETICON++



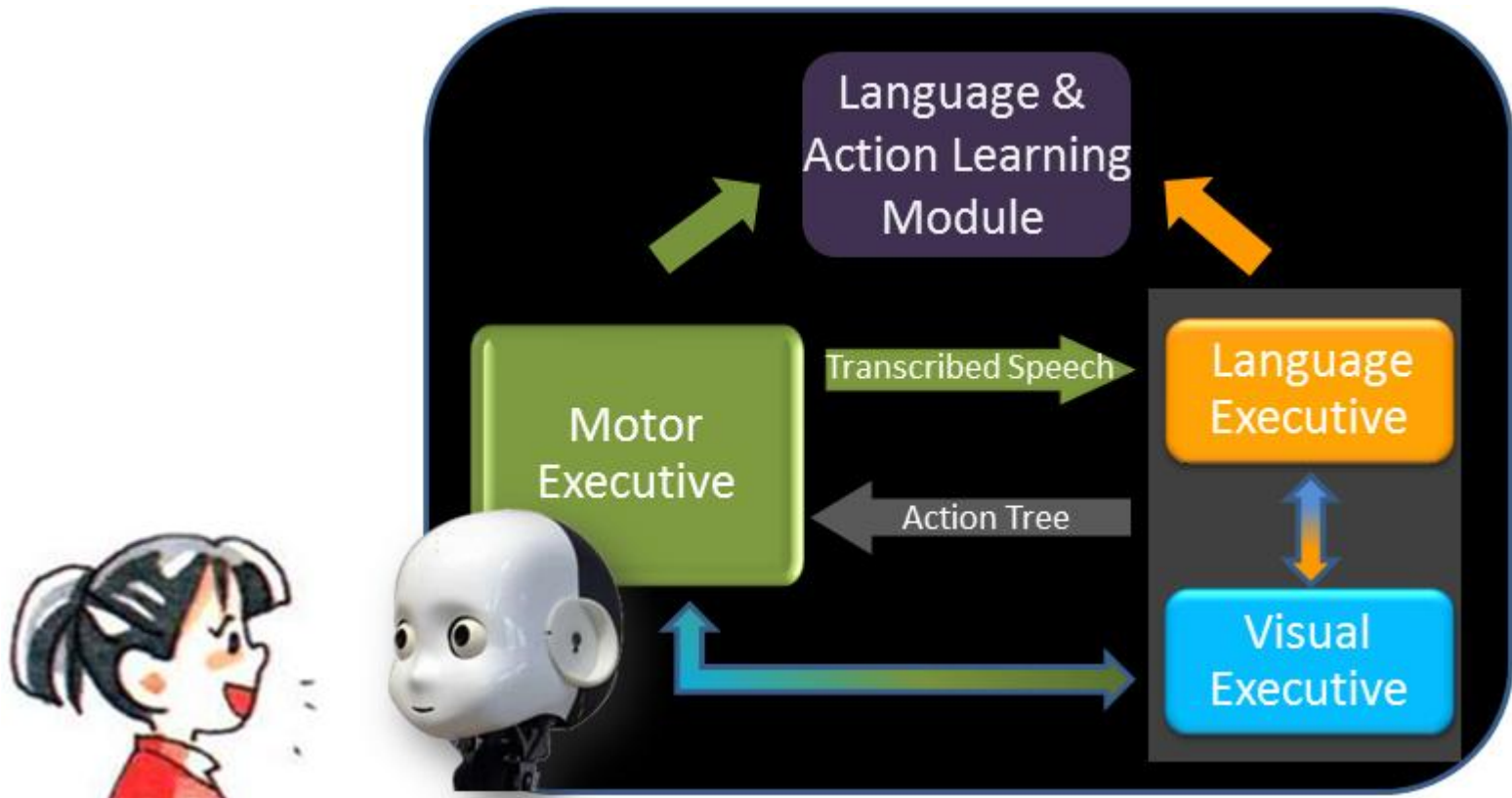
Tools



- A set of *embodied, generative language processing tools*
- A set of *generative visual object and action analysis tools*
- A *self-exploration model* that will integrate motoric skills, multisensory perception skills (visual and tactile) and verbal labelling of self-acquired sensorimotor experiences for artificial agents
- A *neuroanatomical model of language-action learning*
- Improved *grasping skills* for a humanoid *via learning and affordances*.
- *A word-level articulation-based automatic speech recognition system*; this module will allow an agent to go beyond traditional speech recognition abilities using motor information.

POETICON++

Demonstrators (1)



POETICON++

Demonstrators (2)

