The Inner Speech of a Cognitive Architecture

Antonio Chella
Department of Engineering
University of Palermo, Italy
antonio.chella@unipa.it

Università degli Studi di Palermo

Virtual International Symposium on Cognitive Architecture (VISCA 2021)
June 8-11, 2021: 10/11am-1:30pm EDT
Hosted by the Soar Research Group, University of Michigan

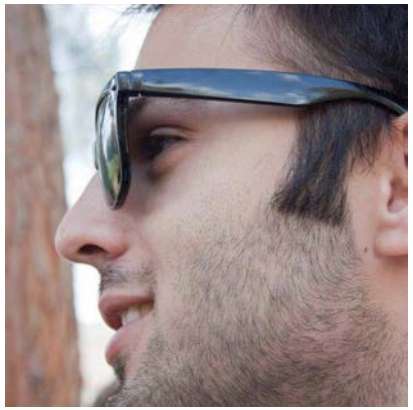
Joint work

Antonio Chella

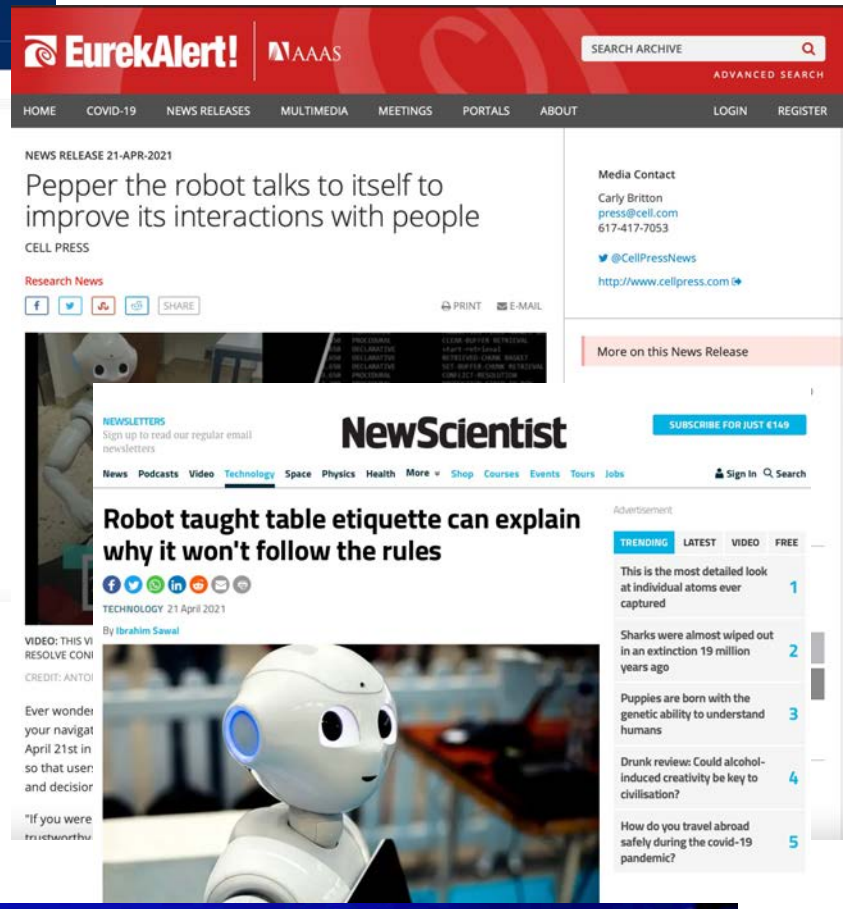
Arianna Pipitone

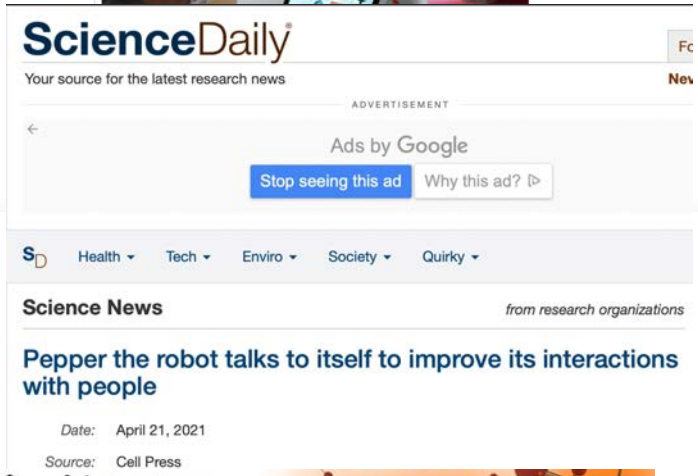
Valeria Seidita

Antonella D'Amico



Alessandro Geraci





Pepper the robot can now 'think out loud': Scientists modify the popular bot so users can hear its thought process and better understand its motivations and decisions

- Pepper is a 4-foot-tall assistance robot created by Japanese company SoftBank
- Researchers specially created 'an extension' of the robot for their experiments
- Pepper was betting at solving dilemmas when it could say its 'thoughts' out loud

Air Force Office of Scientific Research (AFOSR)
20.450 follower
1 giorno •

Researchers at the Università degli Studi di Palermo have designed a robot that 'thinks out loud' so that users can hear its thought process and better understand the robot's motivations and decisions.

"If you were able to hear what the robots are thinking, then the robot might be more trustworthy. The robots will be easier to understand for laypeople, and you don't need to be a technician or engineer. In a sense, we can communicate and collaborate with the robot better." - Professor Antonio Chella

https://lnkd.in/dFr3Jqx

The research is supported by the AFRL/AFOSR Life Sciences and Human Performance program, managed by Dr. Nandini Iyer.

Air Force Research Laboratory #AFOSR_IO #AFOSRBoldResearch #HighRiskHighReward #HumanMachineTeaming #EOARD

Vedi traduzione

20 • 1 commento

Why inner speech for robot?

Inner speech:

- is linked to self-consciousness [Morin, 2005]:
  - comes to play important roles in self-regulation and planning
  - focusing attention and self-attention

it enables high-level cognition (inferences, planning…)

it makes the robot operations more transparent

it allows to internalize human’s explanaition (task learning)

Models of inner speech

- ***Morin***

  inner speech is linked to self-consciousness

- ***Baddeley***

  inner speech as a component of the phonological loop of the working memory

- *Vygotskij*

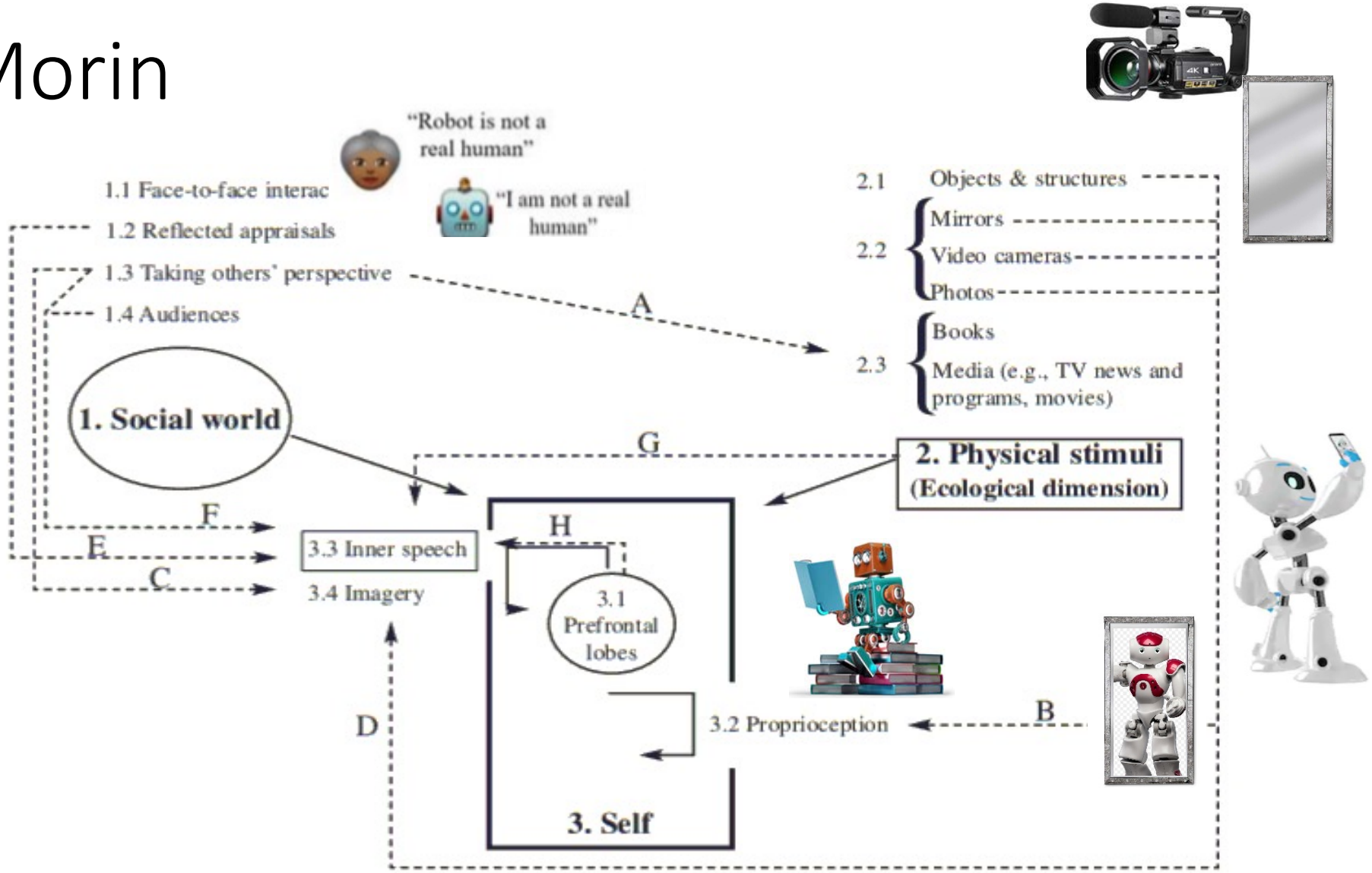
  inner speech is the result of a developemental internalization process

- *Martínez-Manrique and Vicente*

  veichle or activity forms of inner speech

…and others

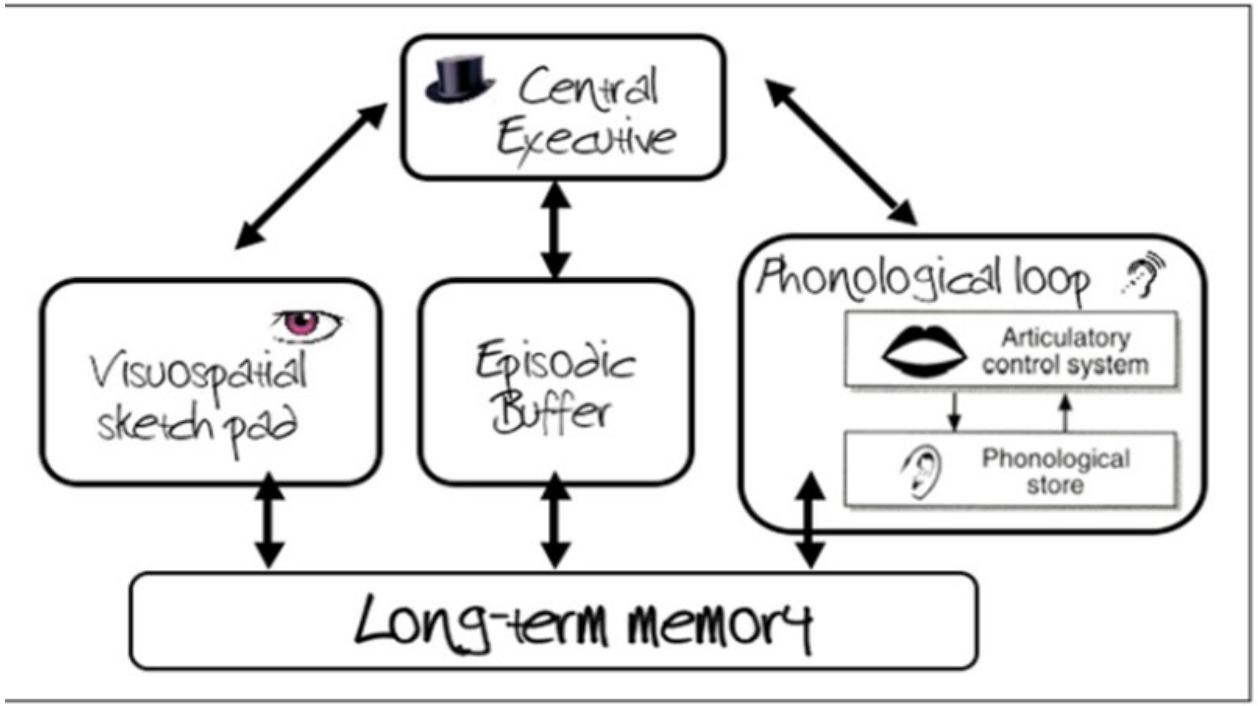
Morin



Morin, Alain. (2005). A Neurocognitive and Socioecological Model of Self-Awareness. Genetic, social, and general psychology monographs. 130. 197-222. 10.3200/MONO.130.3.197-224.

Baddley

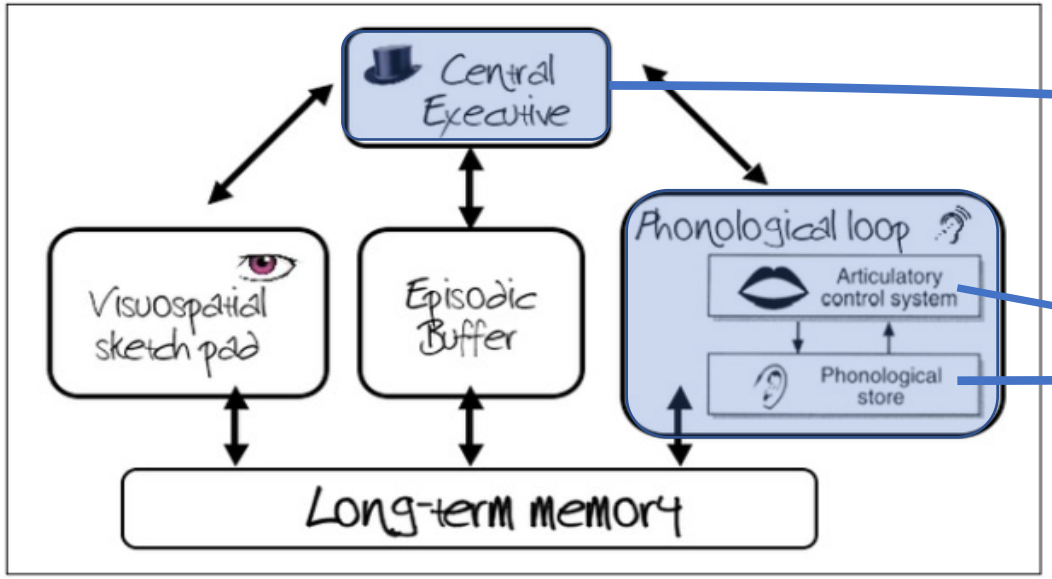
- *Baddeley, A. (1992), 'Working memory', Science 255, 556--559.*



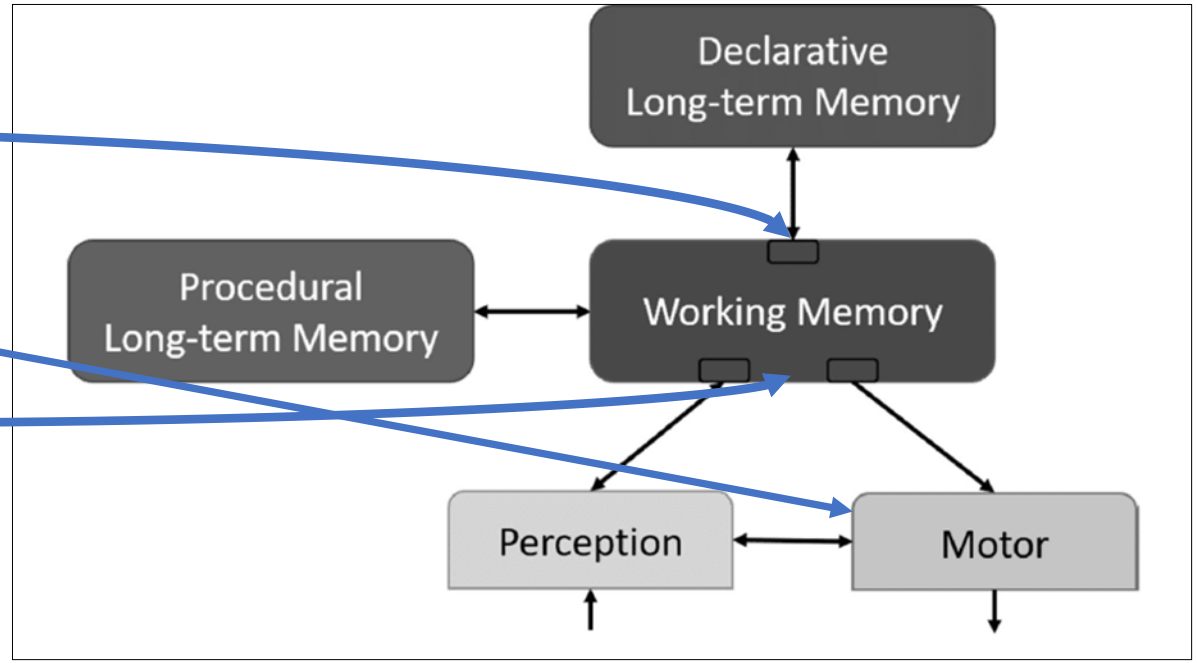
Baddley and the Common Model of Cognition

Integrating Baddley's model of inner speech and the Common Model of Cognition

Baddley's model

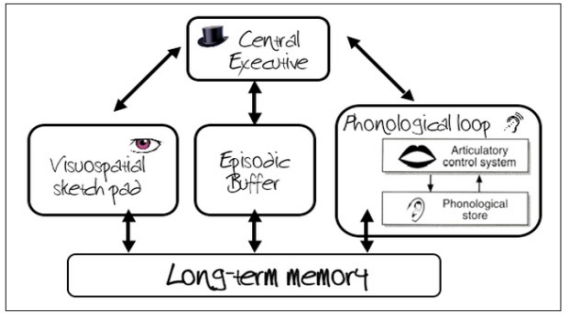


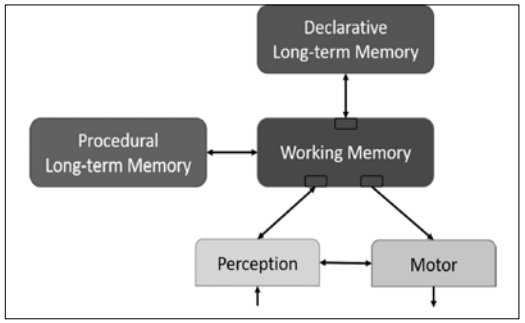
Common Model of Cognition

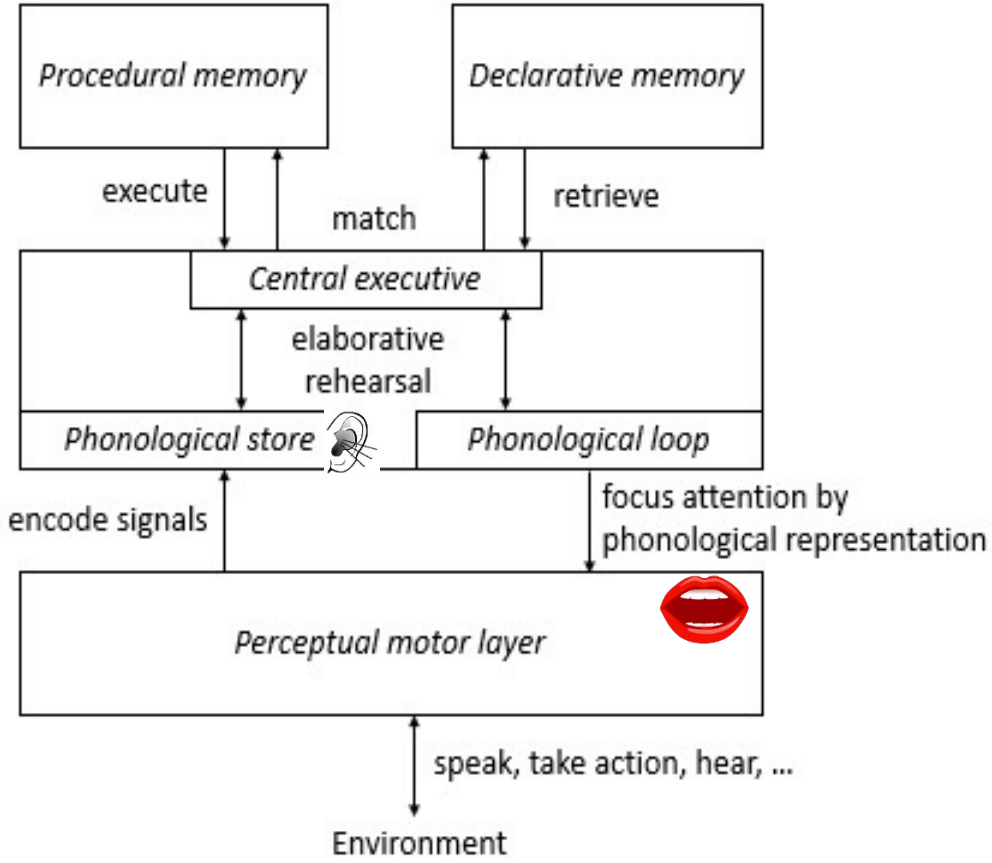


- *Baddeley, A. (1992), 'Working memory', Science 255, 556--559.*
- *Laird, J.E., Lebiere, C., and Rosenbloom, P.S.: A Standard Model of the Mind: Toward a Common Computational Framework across Artificial Intelligence, Cognitive Science, Neuroscience, and Robotics. AI Magazine, Winter, 13–26 (2017).*

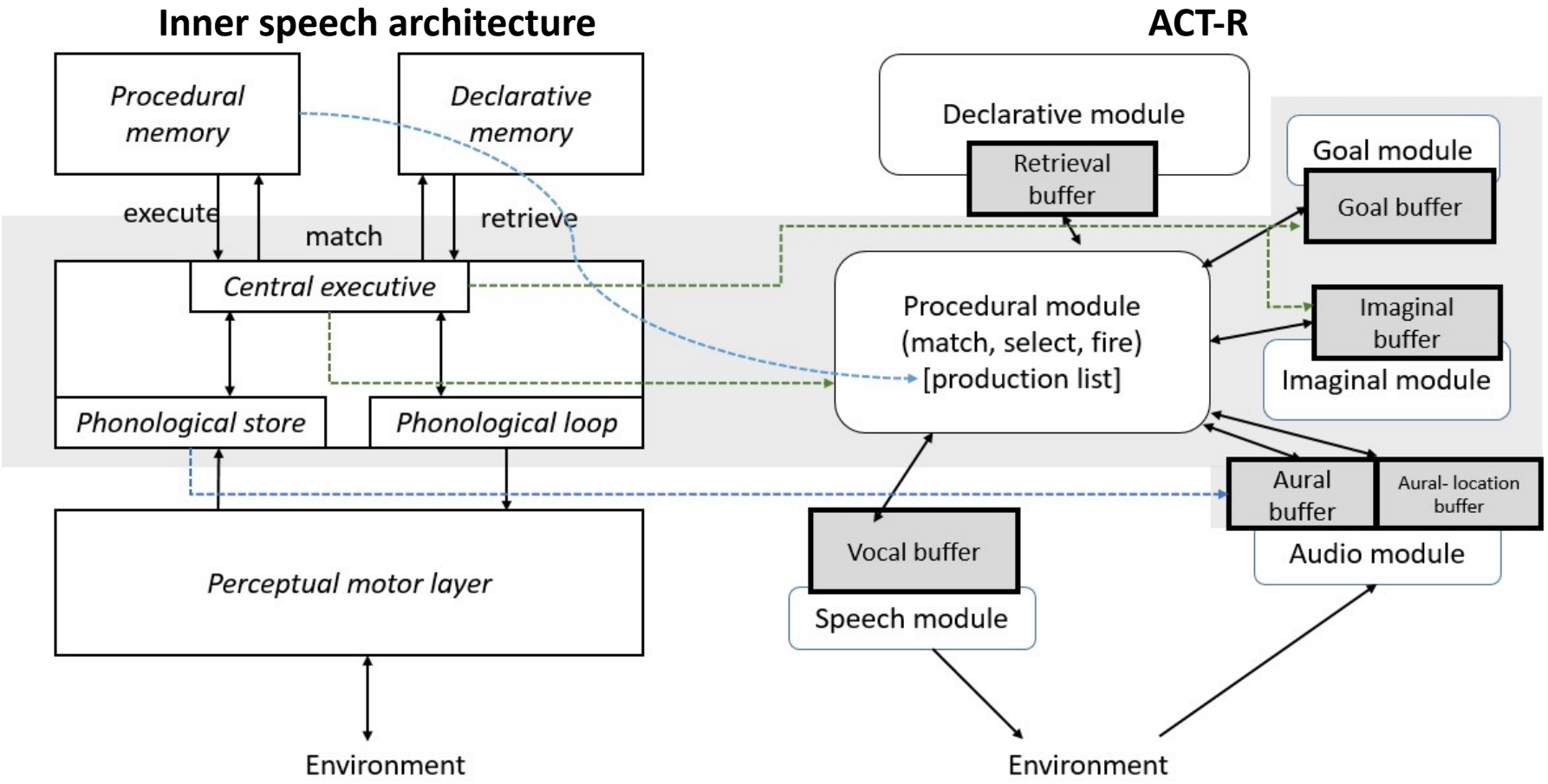
A cognitive architecture for inner speech







Implementation of the architecture



ACT-R inner speech model

Modeling perception:

```
(chunk-type perception channel objects self-reflect)
```

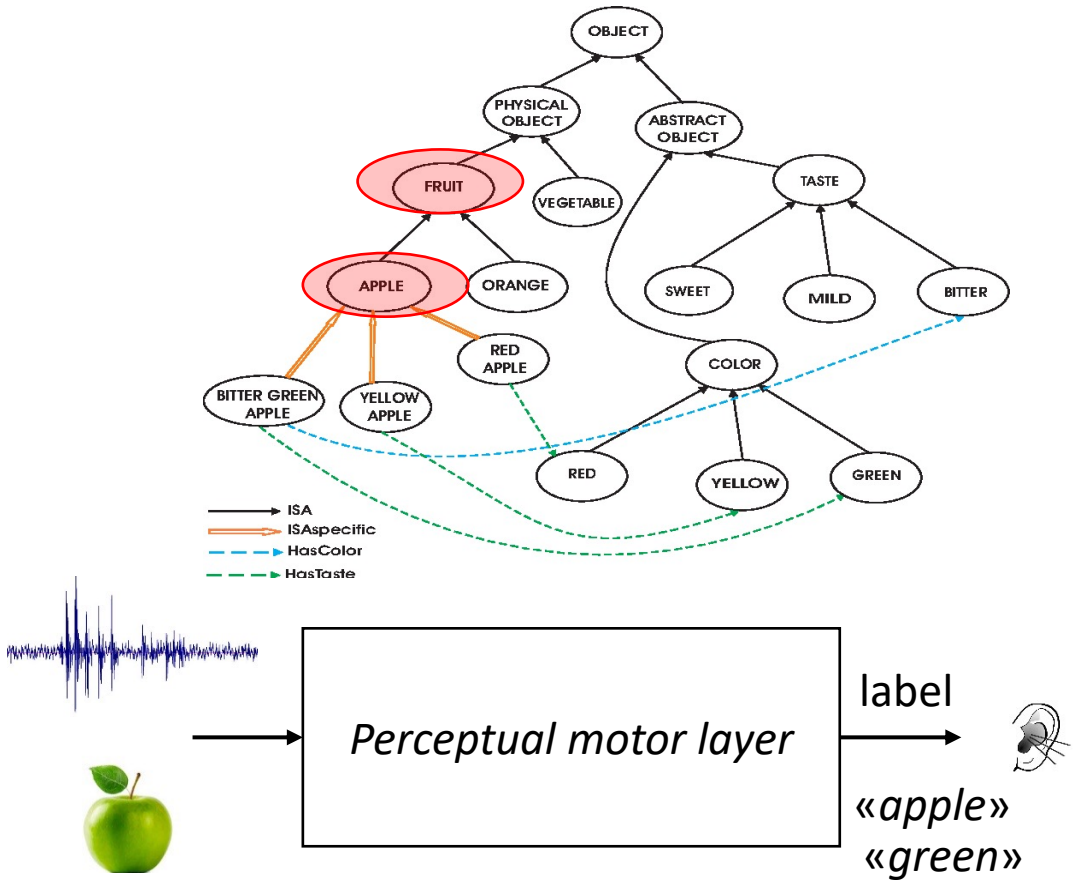
Modeling inner dialogue turn:

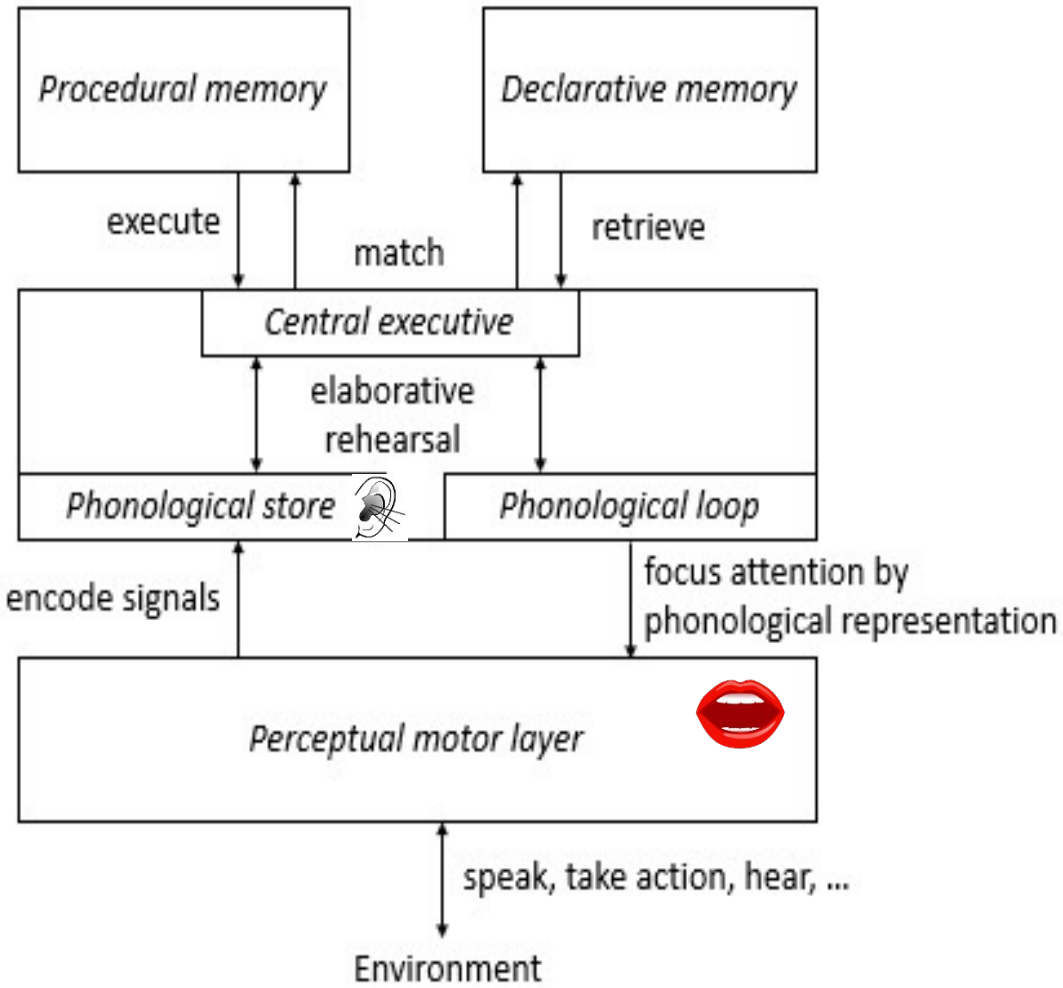
```
(chunk-type link-turn turn1 turn2)
```

Turns for [Morin, 2005]:

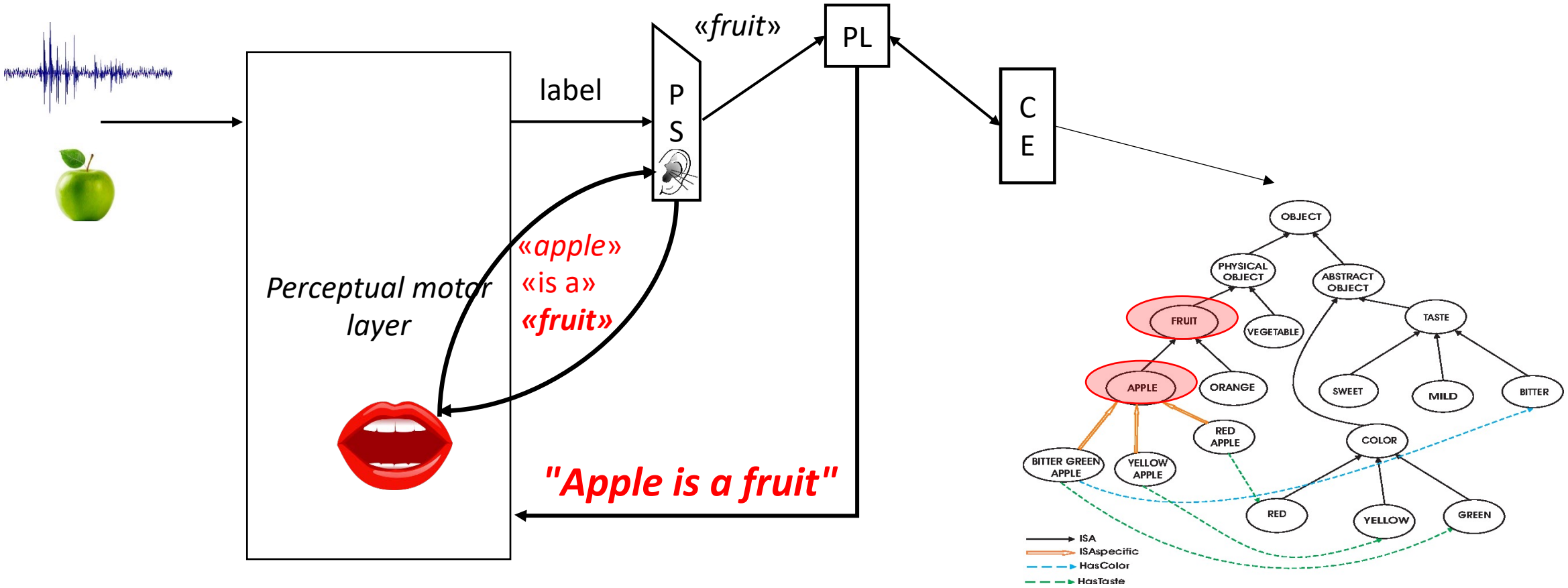
- Social milieu («What do I see in a video?», «What do the other thinl about me?»)
- Self-direct question («What am I doing?»)
- Self-control («I move my hand»)
- Self-focus («My arm hurts me»)

Operation of the architecture

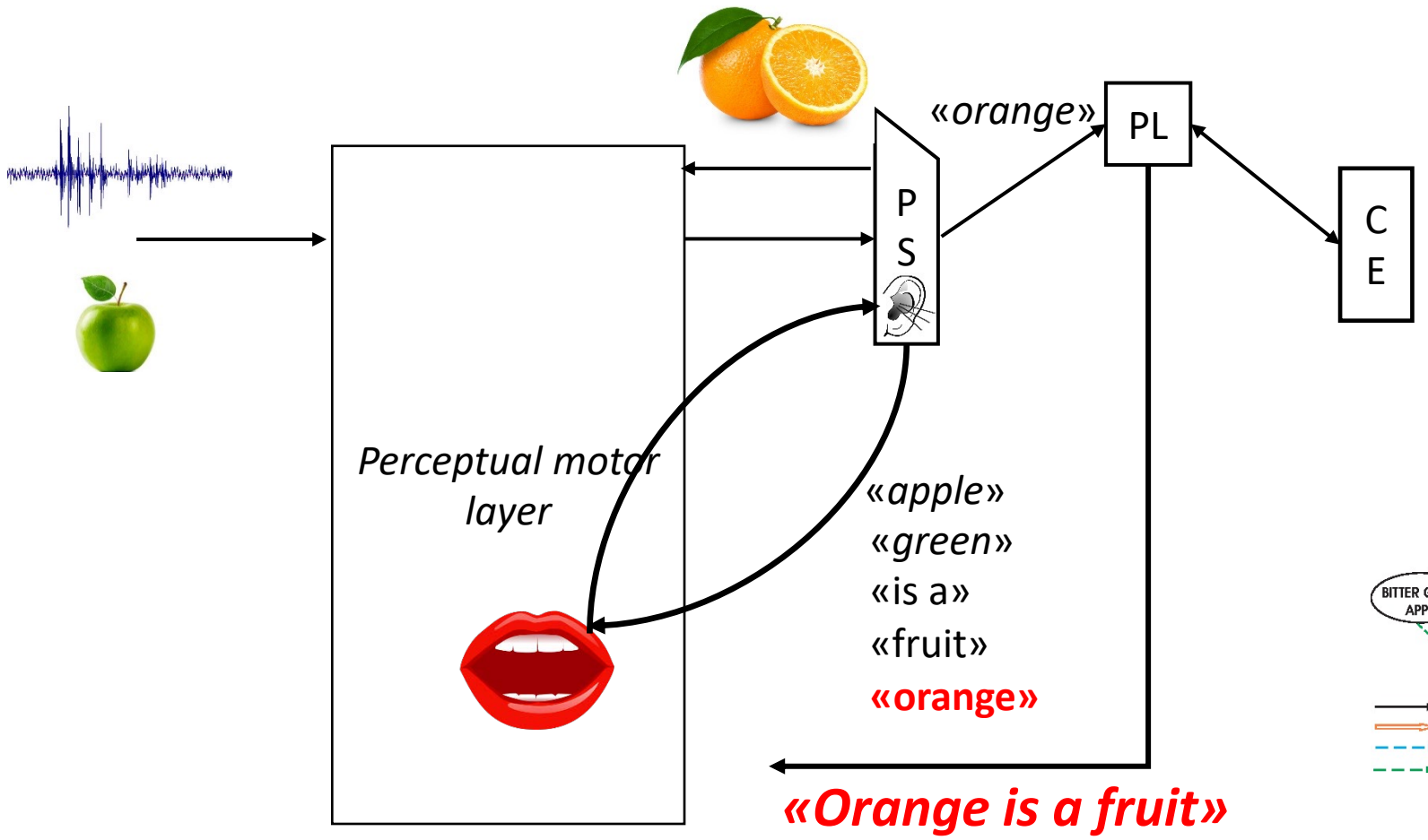


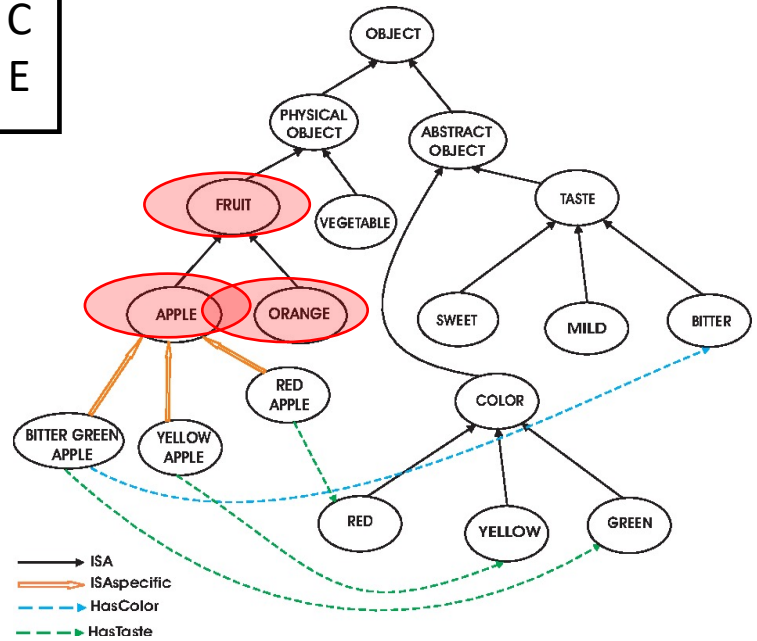


The covert articulator

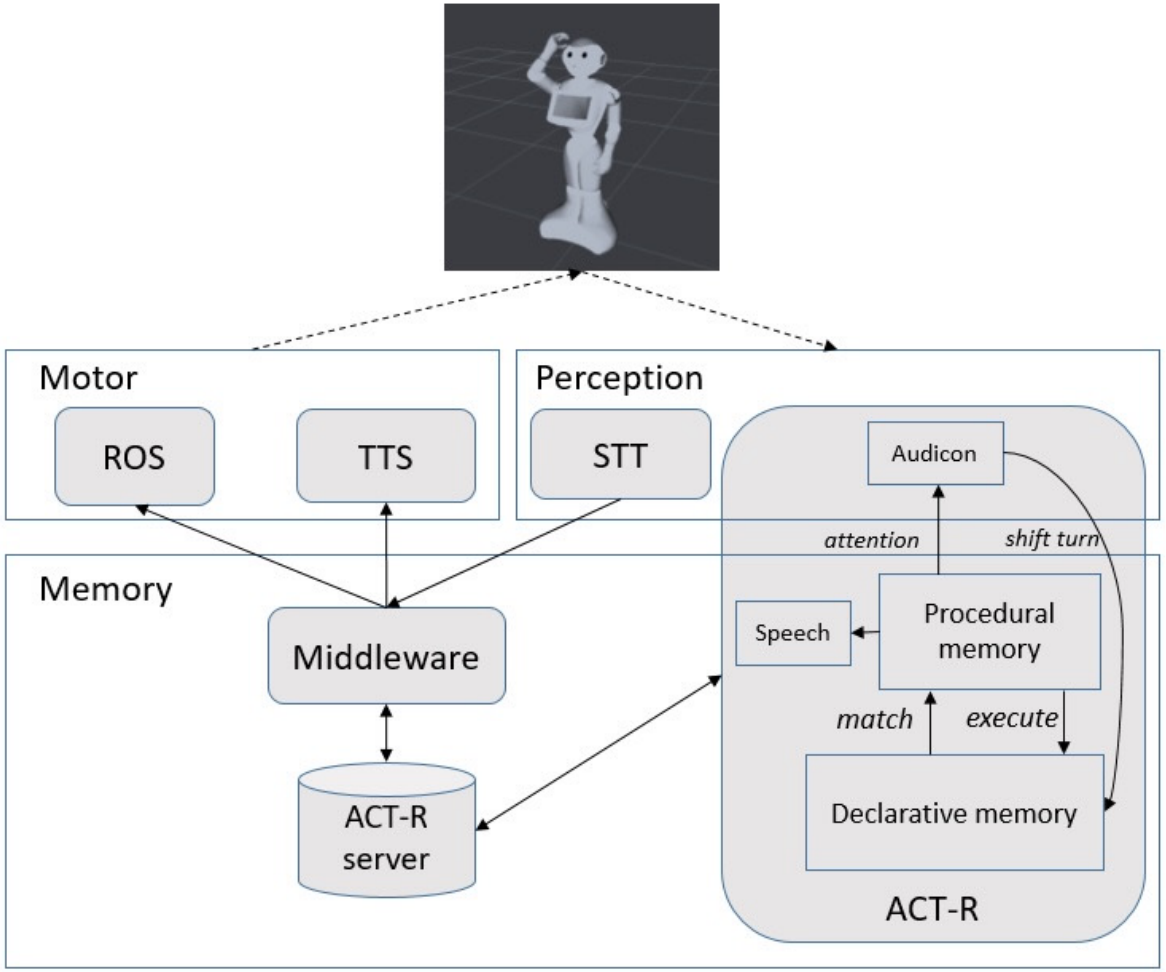


Searching for an orange…





The inner speech implementation in Pepper



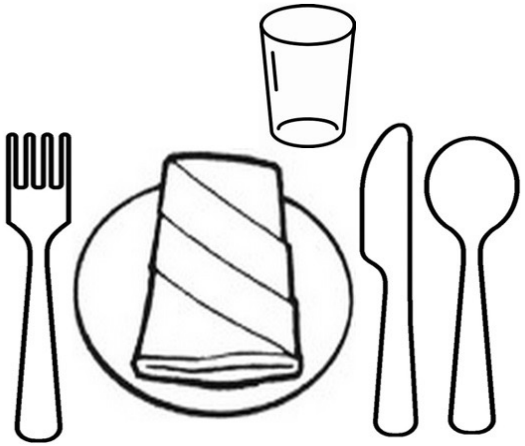
Inner speech for transparency

- The robot talks about its inner reasoning process, making the interaction more transparent.

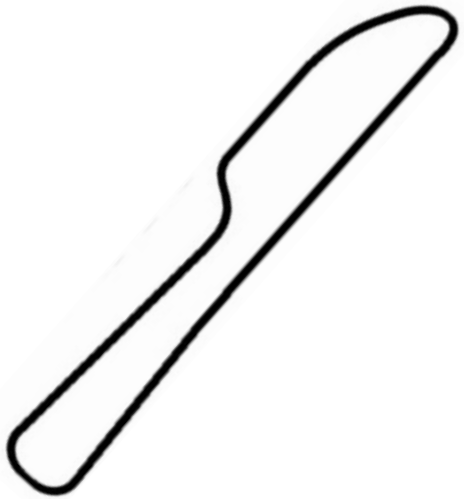
- Self-direct question («What am I doing?» «Did I pick the object before?»)
- Self-control («I move my hand»)
- Self-focus («My arm is overheated»)

How inner speech improves interaction

- By inner speech, facts and information are retrieved from the knowledge base.
- Inner speech improves the interaction with human, enabling the partner to focus on problems otherwise not considered.

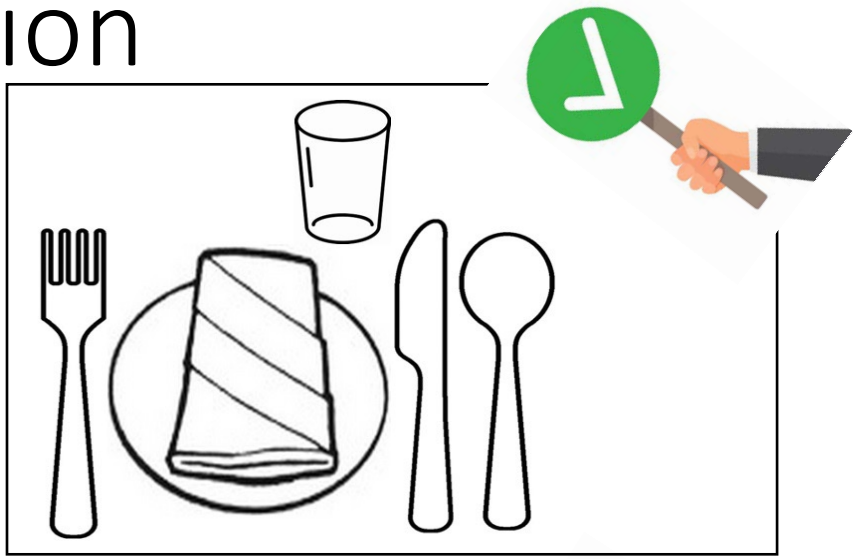


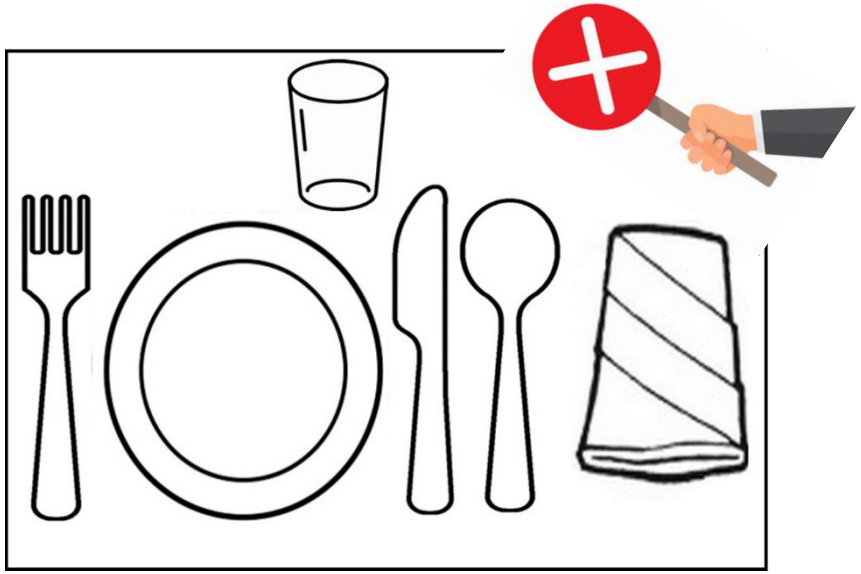
Attention! The knife cuts!



Inner speech for conflict resolution

- What to do when human requires to break a rule? It is a dilemma to solve for robot.
- To place object in a location which is different from those in the etiquette schema, falls in conflict situation.
- inner speech and conflicts (Tappan, 1962)



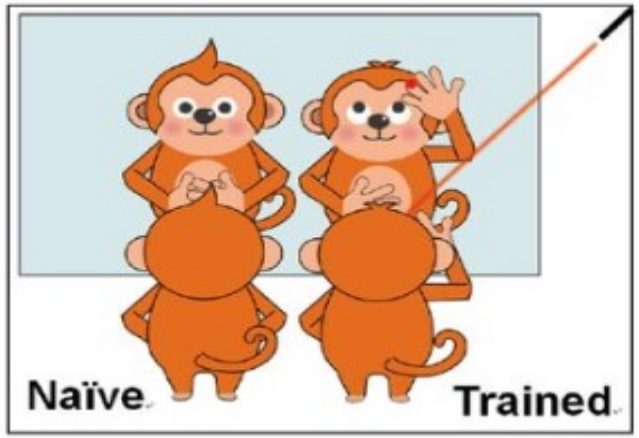


Mirror Self-Recognition Test

- Mirror self-recognition (MSR) test
- ...by using inner speech [Morin, 1989]

- Social milieu («What do I see in a video?», «What do the other thinl about me?»)
- Self-direct question («What am I doing?»)
- Self-control («I move my hand»)
- Self-focus («My arm hurts me»)

Passing test



Morin, Deblois
Gallup's Mirrors: More Than An Operationalization Of Self-awareness In Primates,
Psychological Reports, 65, 287-291.(1989)

UNIVERSITÀ
DEGLI STUDI
DI PALERMO

Prof. **Antonio Chella**
antonio.chella@unipa.it

DEPARTMENT OF ENGINEERING

Head of RoboticsLab

dipartimento
di ingegneria
unipa

Viale delle Scienze, building 6
90128 Palermo, Italy
Tel. +39 09123862601

Thank you for your attention!

frontiers in Robotics and AI

ORIGINAL RESEARCH
published: 19 February 2020
doi: 10.3389/frobt.2020.00016

Developing Self-Awareness in Robots via Inner Speech

Antonio Chella[1,2], Arianna Pipitone[1], Alain Morin[3] and Famira Racy[4]*

[1] RoboticsLab, Dipartimento di Ingegneria, Università degli Studi di Palermo, Palermo, Italy, [2] Istituto di Calcolo e Reti ad Alte Prestazioni (ICAR), Consiglio Nazionale delle Ricerche, Palermo, Italy, [3] Department of Psychology, Mount Royal University, Calgary, AB, Canada, [4] Researcher, Mount Royal University, Calgary, AB, Canada

The experience of inner speech is a common one. Such a dialogue accompanies the introspection of mental life and fulfills essential roles in human behavior, such as self-restructuring, self-regulation, and re-focusing on attentional resources. Although the underpinning of inner speech is mostly investigated in psychological and philosophical fields, the research in robotics generally does not address such a form of self-aware behavior. Existing models of inner speech inspire computational tools to provide a robot with this form of self-awareness. Here, the widespread psychological models of inner speech are reviewed, and a cognitive architecture for a robot implementing such a capability is outlined in a simplified setup.

Keywords: inner speech, self-awareness, robot, human-robot interaction, cognitive cycle

OPEN ACCESS

Edited by:
Salvatore Maria Anzalone, Université Paris 8, France

Reviewed by:
Tony Belpaeme, University of Plymouth, United Kingdom
Soheil Keshmiri, Advanced Telecommunications Research Institute International (ATR), Japan

****Correspondence:***
Antonio Chella
antonio.chella@unipa.it

Specialty section:
This article was submitted to Human-Robot Interaction, a section of the journal Frontiers in Robotics and AI

Received: *27 September 2019*
Accepted: *30 January 2020*
Published: *19 February 2020*

Citation:
Chella A, Pipitone A, Morin A and Racy F (2020) Developing Self-Awareness in Robots via Inner Speech. Front. Robot. AI 7:16. doi: 10.3389/frobt.2020.00016

INTRODUCTION

The idea of implementing self-awareness in robots has been popular in science-fiction literature and movies for a long time. This quest is also becoming increasingly prevalent in scientific research, with articles, special topics, books, workshops, and conferences dedicated to it.

It is widely assumed that there are two dimensions of awareness (see Dehaene et al., 2017), and namely, awareness as experience and awareness as self-monitoring, i.e., self-awareness. In essence, awareness as experience occurs when an agent perceives the environment and experiences it from within in the form of images, sensations, thoughts, and so on (see Block et al., 2019); as such, awareness (or consciousness) exists when an organism can focus attention outward toward the environment (Duval and Wicklund, 1972). Instead, self-awareness takes place when the agent focuses attention inward and apprehends the self in its diverse manifestations, like emotions, thoughts, attitudes, sensations, motives, physical attributes, which frequently involves a verbal narration of inner experiences (Morin, 2011).

Models of awareness and self-awareness are being proposed, each with idiosyncratic views of what the aforementioned concepts constitute, as well as different suggestions on how to implement them in artificial agents (see among others, Tononi and Edelman, 1998; Gray et al., 2007; Seth, 2010; Edlund et al., 2011; Oizumi et al., 2014; Tononi et al., 2016; Juel et al., 2019). For reviews, see Reggia (2013) and Chella et al. (2019).

The proposed approach focuses on implementing a form of robot self-awareness by developing inner speech in the robot. Inner speech is known to importantly participate in the development and maintenance of human self-awareness (Morin, 2018); thus, self-talk in robots is an essential behavioral capability of robot self-awareness.

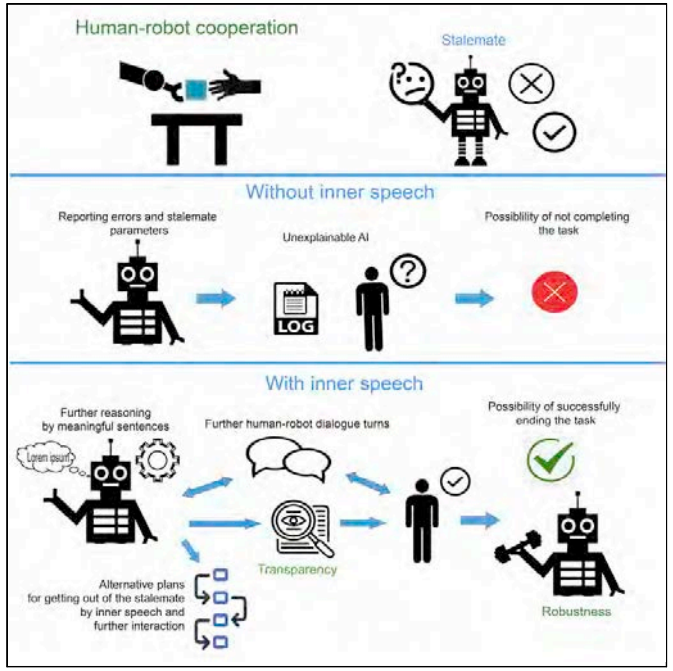
More in detail, the paper discusses a computational model of inner speech. The proposed model is based on the cognitive architecture described by Laird et al. (2017). Therefore, the approach

iScience

CellPress OPEN ACCESS

Article

What robots want? Hearing the inner voice of a robot



Arianna Pipitone, Antonio Chella

arianna.pipitone@unipa.it

Highlights

An inner speech cognitive architecture enables robots for rehearsing and self-talk

Robot's inner speech affects functionality and transparency in human-robot cooperation

To self-talk enables the robot to further reasoning and plans in accomplishing the task

The inner speech is applicable in many robotics contexts as learning and regulation

Pipitone & Chella, iScience 24, 102371
April 23, 2021 © 2021 The Author(s).
https://doi.org/10.1016/j.isci.2021.102371

frontiers in Robotics and AI

HYPOTHESIS AND THEORY
published: 23 April 2021
doi: 10.3389/frobt.2021.620026

Automation Inner Speech as an Anthropomorphic Feature Affecting Human Trust: Current Issues and Future Directions

Alessandro Geraci [1,2], Antonella D'Amico [2], Arianna Pipitone [1], Valeria Seidita [1] and Antonio Chella [1]*

[1] Robotics Lab, Department of Engineering, University of Palermo, Palermo, Italy, [2] Department of Psychology, Educational Science and Human Movement, University of Palermo, Palermo, Italy

This paper aims to discuss the possible role of inner speech in influencing trust in human–automation interaction. Inner speech is an everyday covert inner monolog or dialog with oneself, which is essential for human psychological life and functioning as it is linked to self-regulation and self-awareness. Recently, in the field of machine consciousness, computational models using different forms of robot speech have been developed that make it possible to implement inner speech in robots. As is discussed, robot inner speech could be a new feature affecting human trust by increasing robot transparency and anthropomorphism.

Keywords: inner speech, trust, anthropomorphism, human-automation interaction, human-robot interaction, robot, automation

OPEN ACCESS

Edited by:
Marco Nørskov, Aarhus University, Denmark

Reviewed by:
James A. Reggia, University of Maryland, United States
Felix Lindner, University of Ulm, Germany

Correspondence:
Antonella D'Amico
antonella.damico@unipa.it

Specialty section:
This article was submitted to Ethics in Robotics and Artificial Intelligence, a section of the journal Frontiers in Robotics and AI

Received: *21 October 2020*
Accepted: *26 February 2021*
Published: *23 April 2021*

Citation:
Geraci A, D'Amico A, Pipitone A, Seidita V and Chella A (2021) Automation Inner Speech as an Anthropomorphic Feature Affecting Human Trust: Current Issues and Future Directions. Front. Robot. AI 8:620026. doi: 10.3389/frobt.2021.620026

INTRODUCTION

In the past years, robots and automation development and implementation have increased exponentially in every context, leading to growing interactions with humans (Merritt and Ilgen, 2008). Robots are now used in different contexts, such as military, security, medical, domestic, and entertainment (Li et al., 2010). Robots, compared with other types of automation (e.g., machines, computers), are designed to be self-governed to some extent to respond to situations that are not prearranged (Lewis et al., 2018). Therefore, the greater the complexity of robots, the higher the importance to focus on factors that influence human–automation interaction (HAI) as their collaboration increases over time (Lee and See, 2004; Schaefer et al., 2016). In this paper, we aim to start the exploration of the role of inner speech in HAI and, in particular, on its role in improving human trust toward automation. For this purpose, we first focus on the concept of inner speech in psychological literature, also examining the first results of its implementation in automation. Then, we discuss the possible role of inner speech as one of the anthropomorphic automation features that may affect human trust in HAI.

INNER SPEECH

Inner speech is an everyday covert inner monolog or dialog with oneself, which is essential for human psychological life and functioning because it is linked to reasoning, self-regulation, and self-awareness (Morin, 2012).

ELSEVIER

Available online at www.sciencedirect.com

ScienceDirect

Cognitive Systems Research 59 (2020) 287–292

Cognitive Systems RESEARCH

www.elsevier.com/locate/cogsys

A cognitive architecture for inner speech

Antonio Chella [a,b], Arianna Pipitone [a,*]

[a] Dipartimento di Ingegneria – Universitá degli Studi di Palermo, Italy
[b] C.N.R., Istituto di Calcolo e Reti ad Alte Prestazioni, Palermo, Italy

Received 4 June 2019; accepted 12 September 2019
Available online 4 October 2019

Abstract

A cognitive architecture for inner speech is presented. It is based on the Standard Model of Mind, integrated with modules for self-talking. Briefly, the working memory of the proposed architecture includes the phonological loop as a component which manages the exchanging information between the phonological store and the articulatory control system. The inner dialogue is modeled as a loop where the phonological store hears the inner voice produced by the hidden articulator process. A central executive module drives the whole system, and contributes to the generation of conscious thoughts by retrieving information from long-term memory. The surface form of thoughts thus emerges by the phonological loop. Once a conscious thought is elicited by inner speech, the perception of new context takes place and then repeating the cognitive loop. A preliminary formalization of some of the described processes by event calculus, and early results of their implementation on the humanoid robot Pepper by SoftBank Robotics are discussed.

Keywords: Inner speech; Cognitive architecture

1. Introduction

Daily, human beings are engaged in a form of inner dialogue, which enables them to high-level cognition, including self-control, self-attention and self-regulation. By inner dialogue, a person plans tasks, finds problem's solution, self-reflects, critical thinks, feels emotions, and restructures the perception of the world and of himself. Obviously, the inner dialogue cannot be directly observed, thus making empirical studies difficult. However, psychological and philosophical perspectives were developed during the last decades, and are recognized in research communities.

Alderson-Day and Fernyhough (2015) stated that to talk to oneself makes a person able to retrieve memorized facts, learn new knowledge and, in general, simplify otherwise demanding cognitive processes. The self-dialogue is closely related to thought, and therefore is an essential component in the dynamics of information thinking; Carruthers (1998), Jackendoff (1996), among many others, claimed that genuine conscious thoughts need language. Vygotsky (2012) considers inner language as the result of an internalization process during which linguistic explanations by a caregiver to a children become an inner conversation of the children with the self when he is engaged in similar task. Morin (2005) stated that inner dialogue can be linked to self-consciousness. Self-attention on internal resources triggers inner speech and generates self-awareness on these resources.

Baddeley (1992) described the inner speech phenomena by a working memory architecture, which is suitable for the automation of the processes into artificial agent. In particular, Baddeley claimed that the inner voice is the re-entrance of a sentences covertly produced by an articula-

* Corresponding author.
E-mail addresses: antonio.chella@unipa.it (A. Chella), arianna.pipitone@unipa.it (A. Pipitone).

https://doi.org/10.1016/j.cogsys.2019.09.010
1389-0417/