



**HAL**  
open science

# Spoken Language Interaction with Virtual Agents and Robots: Personal statement

Joseph J Mariani

► **To cite this version:**

Joseph J Mariani. Spoken Language Interaction with Virtual Agents and Robots: Personal statement. Spoken Language Interaction with Virtual Agents and Robots (SLIVAR), Jun 2020, Dagstuhl, Germany. hal-04413415

**HAL Id: hal-04413415**

**<https://hal.science/hal-04413415>**

Submitted on 23 Jan 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Spoken Language Interaction with Virtual Agents and Robots: Personal statement.

*Joseph Mariani*

Université Paris-Saclay, CNRS, LIMSI, 91400, Orsay, France

## **Background**

I'm not a specialist in robotics, nor in human-robot interaction. My interest in participating is to have a better view of the state-of-the-art in those fields, and to encourage the two communities of specialists in robotics and human-machine (spoken language) interaction to work more closely, using best practices, and especially evaluation.

## **State-of-the-art**

Where are we now regarding (spoken) interaction with robots? I recently asked Roger Moore at the LT4All conference, who was to say "stone age"?

## **Mutual understanding**

For many years, researchers thought that spoken interaction with robots was simply adding speech input/output to robots. Things have now slightly progressed, but the two communities still have to better understand each other, their research agenda and their constraints.

## **HRI Evaluation metrics and protocols**

Objective and quantitative evaluations have been decisive in the development of workable language technologies, starting in the 80s. It seems that the situation is not comparable in the robotics area. How could the evaluation paradigm be also used in robotics? What would be the evaluation metrics and possible protocols? Conducting evaluation tests with embodied robots appears difficult and costly. Can we conduct them in simulated environments, including virtual agents? To what degree is it similar to using actual robots? Similarly, conducting spoken dialog systems evaluation with human subjects is costly and time consuming. How can we automatize this process with proper data, metrics and protocols?

## **Multi-Party Human(s)-Robot(s) interaction**

Humans are often able to understand whether they are concerned by a statement issued by other humans. How can we handle this when one or several humans are communicating with one or several robots?

## **Language portability**

Most of the developments concern the English language, while robots will be used by humans who want to keep on speaking their native language. What is the size of the effort to port a spoken dialog system to another language? Do we have to devote the same effort, or is there a way to adapt the system to a different language? It is usually necessary to benefit from large quantities of speech to develop an application, while they are difficult to gather in spoken dialog situations. How can we address this difficulty?

## **Ethics of robots**

Who is responsible for the robots behavior, especially given that it is based on machine learning? Is it the designer, the programmer, the trainer, the seller, the owner?