connected everything.
industrial systems in the digital age

Investigating spoken dialogue to support manufacturing processes

Project team

University of Sheffield

James Law
Robotics

Rob Gaizauskas
Computer Science

Emma Barker
Computer Science

Industry partners

AMRC ADVANCED MANUFACTURING RESEARCH CENTRE / The University Of Sheffield.

ELG Utica Alloys

Contact

Rob Gaizauskas
University of Sheffield

r.gaizauskas@sheffield.ac.uk

Will factory workers soon talk with robots?

Digital technologies are bringing about revolutionary changes in manufacturing. Robots, for example, have great potential to assist with manufacturing tasks in factories throughout the UK. There is much interest in how and when humans can work together with robots to deliver a performance which is better than either one working alone. This is known as human-robot co-working. Another emerging technology is intelligent decision support; here, factory workers' decision-making is supported by information-rich systems.

In both human-robot co-working and intelligent decision support, communication between humans and machines/systems is central. Speech Dialogue Systems have developed over the last 20 years for various applications (e.g. for booking a flight), but not for a factory setting.

Maintenance, Repair and Overhaul (MRO) is an activity common to industries such as aerospace, automotive and civil engineering. It could be transformed through the use of spoken language interfaces. Imagine a human and robot working together to disassemble and recover materials from a spent jet engine. This scenario requires situated interaction in which both parties refer to objects in their shared setting. The human perhaps instructing the robot to use its gripper to remove a specific part; in deciding how to proceed the human might wish to consult external information sources to determine, for example, whether removing the part might release a hazardous chemical.

The EPSRC-funded Network Plus *Connected Everything: Industrial Systems in the Digital Age* aims to identify the key challenges we face as digital technologies transform our industrial systems.

EPSRC
Engineering and Physical Sciences Research Council

connected everything.

industrial systems in the digital age

There are obvious benefits of using speech to communicate with machines and systems in factories:

- It is 'hands-free' and 'eyes-free', thus allowing the user to carry on with physical tasks.
- It is natural for the worker, requiring minimal training.
- It is highly flexible, allowing communication at varying levels of detail and the capacity to switch quickly between multiple task contexts.

Could developing this technology allow work to be carried out more effectively, whilst also gaining a positive response from the workforce? First, there are challenges to overcome:

- Providing current 'Spoken Dialogue Systems' with appropriate manufacturing tasks is hard.
- Factories often have significant background noise.

The project is the first to assess the potential for Spoken Dialogue Systems to be integrated into human-robot co-working and intelligent decision support in real manufacturing settings. The project has demonstrated that this is both feasible and likely to be a highly fruitful line of work.

Exploring the potential of Spoken Dialogue Systems in manufacturing settings

Observing current practice

We visited our industry partners' factories to identify tasks that could benefit from using human speech to communicate with machines/systems. Following discussions with partners, employers and workers, a number of tasks were selected to develop into typical 'use' cases.

The two factories provide examples of contrasting environments, one traditional and the other highly digitised, both carrying out Maintenance, Repair and Overhaul (MRO) activities.

Identifying use cases

Spoken dialogue technologies would be of most benefit in tasks requiring: (A) Lifting and Positioning, (B) Decision Support, (C) Data entry. Typical dialogues for (A) and (B) are depicted. Immediate gains could result from implementing spoken dialogue technologies for (C). Gains from spoken language interfaces for (A)-type tasks will depend on more factories using robots.

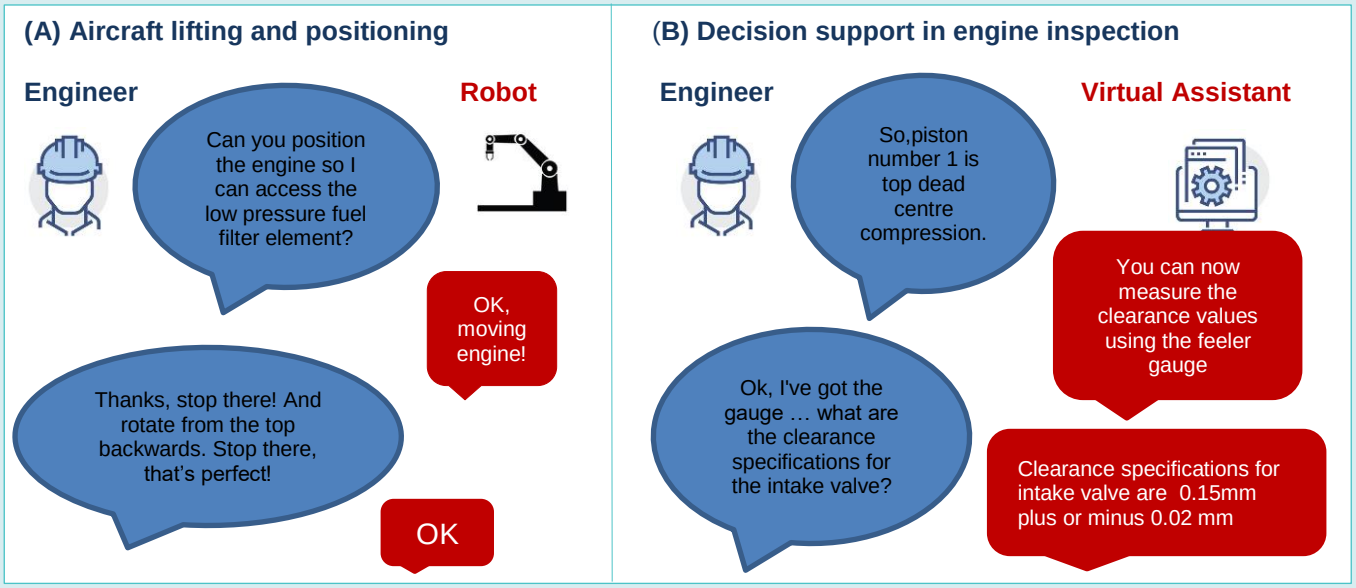
(A) Lifting and Positioning: ELG Utica need flexible mechanical support during the initial stages of breaking down a large jet engine. We explored the scenario of a robotic arm holding a jet engine and positioning it, under voice control, to enable a worker to access and remove parts. This process is currently carried out by many people using fork-lift trucks to position the engine. Automation would require workers to be able to direct the robot using a hands-free interface. We received generally positive feedback.

connected everything.

industrial systems in the digital age

(B) Decision Support: We explored the inspection practice of one of AMRC's partners. We looked at the kinds of information required for successful Maintenance, Repair and Overhaul tasks in the aerospace and automotive industries and found that spoken dialogue systems would be of use when a worker needs to check and set valve clearance or rivet diameters.

(C) Data-entry: This would be useful to ELG Utica to replace their current system in which weights and process settings are recorded on paper, which is then taken to a remote computer location for data-entry.



Testing Automatic Speech Recognition (ASR) systems

We devised a dialogue based on the Data Entry use case. A researcher and an ELG Utica Alloy employee role-played the dialogue in key locations in the very noisy factory and their conversations were recorded. The recorded data were run through two Automatic Speech Recognition systems: the Google Cloud ASR system and the Kaldi open source ASR system. Both were robust enough to use in practice in real manufacturing settings. Both would need some adaptation for each specific use case.

Exploring use case (A) 'Lifting and Positioning' in the laboratory

We have created a demonstrator of human-robot co-working using a speech interface to direct the actions of the robot. Mapping commands directly to motor actions, such as "move right arm up", resulted in an interface that was awkward. We found more abstract commands such as "more/less", "back a bit", "towards me" and instructions such as "pass the spanner" to be more user-friendly.

The EPSRC-funded Network Plus *Connected Everything: Industrial Systems in the Digital Age* aims to identify the key challenges we face as digital technologies transform our industrial systems.

EPSRC Engineering and Physical Sciences Research Council

connected everything.

industrial systems in the digital age

Scenario specification

Use Cases (A) and (B) have provided a basis for the design of two novel experimental scenarios and protocols. The aim is to capture human to human task-based dialogues that will help to establish the requirements for human to computer dialogues in the manufacturing tasks described below (Scenarios 1 and 2).

Scenario 1: A human-robot interaction in which the robot is lifting and positioning an artifact under voice control to support a human in carrying out maintenance work on it.

Scenario 2: Human-intelligent agent interaction in which the human needs advice and instructions on how to carry out a maintenance task.

Key finding 1

Automatic Speech Recognition (ASR) systems are now robust enough to use in practice in manufacturing settings.

Key finding 2

Multiple potential applications for ASR in manufacturing and construction settings have been identified, for shop floor data entry and querying and MRO assembly tasks.

Wider applications

The project highlighted the possibility of using spoken language to transform the experience of using Augmented Reality (AR) headsets. The use of headsets is receiving much attention from the manufacturing sector. The project also found that spoken dialogue would be beneficial in the construction industry, where there is an increasing requirement for access to information systems by several groups of workers who require use of their hands.

What next?

In further work we intend to:

- Conduct the experiments designed in this project, based on use cases (A) and (B). AMRC Apprenticeship trainees will be taking part.
- Build a demonstrator of spoken language dialogue combined with an AR headset display for an application in Maintenance, Repair and Overhaul.
- Explore the possibilities for extracting task models from equipment manuals and how-to documents.
- Investigate commercialisation of this work.