Recent innovations in robotics technology have drastically reduced the size and production costs of robots, leading to a natural growth in multi-robot systems. While multi-robot systems have been commonplace throughout the world for some time now, particularly in the manufacturing industry, the potential of swarms of semi-autonomous robots has sparked interest in a new field of study - swarm-robotics. This field of study encompasses systems comprised of a large number of robots that collaborate to achieve a desired goal. The field presents a number of challenges relating to the effective command and control of a large number of semi-autonomous robots. These challenges have prompted study into Human-Swarm Interaction (HSI) - the effective integration of human operators/collaborators with robotic swarms. This project will investigate the prospect of speech-based HSI in the command and control of a simulated robotic swarm.

Contents

I. Introduction ................................................................................................................................................. 1
II. Research Objectives ................................................................................................................................. 1
   A. Aim .......................................................................................................................................................... 1
   B. Requirements ......................................................................................................................................... 1
III. Literature Review ...................................................................................................................................... 1
   A. Automatic Speech Recognition ......................................................................................................... 2
   B. Swarm Robotics and Human-Swarm Interaction ................................................................................. 2
   C. Command and Control Methodologies .............................................................................................. 4
IV. Research Significance ............................................................................................................................... 4
V. Speech-Based HSI ......................................................................................................................................... 5
   A. Speech-Based Command Sets ............................................................................................................... 5
   B. Simulation and Evaluation .................................................................................................................... 6
VI. Experimental Results and Discussion ...................................................................................................... 7
   A. Characterisation .................................................................................................................................... 7
   B. Command Set Evaluation ..................................................................................................................... 9
VII. Conclusion and Future Work .................................................................................................................. 10
Acknowledgements .......................................................................................................................................... 11
References ......................................................................................................................................................... 11
Appendix A. ....................................................................................................................................................... 12
I. Introduction

Recent research into swarm robotics has focused on Human-Swarm Interaction (HSI), which is the effective integration of human operators/collaborators with robotic swarms. This research becomes increasingly important as robotic swarms are utilised in a broader range of applications, and as the size of the swarms increase. While an argument may be made that, in some scenarios, a fully autonomous swarm without human interaction could achieve higher levels of performance, human involvement is highly desirable if not essential in many conceivable scenarios including: search and rescue, military operations and humanitarian relief. Related to this research are a number of challenges and potential solutions in the context of swarm communication, state estimation and state visualization, as well as human command and control of swarms [1]. The focus of this project is speech-based HSI, investigating the feasibility of enabling a single human operator to effectively control a robot swarm solely via speech.

Within the scope of this project it is important to make a distinction between speech recognition, voice recognition and speaker recognition/diarization. Speech recognition is the conversion of an audio recording of a person’s speech into a string, or discrete words, whereas speaker recognition usually refers to the identification or validation of the person speaking in the recording without regard to what they are saying. While speaker recognition could also be applied in the context of HSI in order to distinguish the operator from intermixed voices, the scope of this project is limited to the application of speech recognition. It should also be noted that the term voice recognition is ambiguous and can be used to refer to both speech and speaker recognition.

The Boids (short for ‘bird-oid’) model proposed by Craig Reynolds in 1987 [2] will be utilised within the scope of this project to simulate the behaviour of robotic swarms. The simulated swarm will be controlled through proposed speech command sets that form the focus of this project. The research focus of this project will be the evaluation of speech-based HSI in its ability to allow a human operator to assert effective control over a simulated robotic swarm.

II. Research Objectives

A. Aim

The aim of this research project is to investigate the feasibility of using speech-based for the control of robotic swarms. This will be assessed through a swarm simulation based on Craig Reynolds’ Boids model.

B. Requirements

The implied requirement of this project is to develop a system that can perform Automatic Speech Recognition (ASR) and produce the desired effect in the Boids simulation. This system should then provide the means through which a speech command structure can be implemented. First of all, this system needs to perform some form of Automatic Speech Recognition (ASR) in order to receive speech input from the human operator. Building on this, some form of parsing will be required through which the command structure will be utilised to convert the string received from the ASR into a desired command for the simulated swarm. The final requirement will be the development of the swarm simulation through which the effectiveness of speech-based HSI will be evaluated.

III. Literature Review

A literature review was conducted in order to establish the state of recent research relevant to this project. Areas of research reviewed included: automatic speech recognition, swarm robotics, human-swarm interaction and command & control methodologies.
A. Automatic Speech Recognition

The focus of the speech recognition literature review was on Automatic Speech Recognition (ASR), which functionally consists of automatic computational conversion of received audio recordings into a word sentence. Automatic speech recognition (ASR) presents one of the biggest challenges in Artificial Intelligence (AI) with the greatest challenges arising when changing speaker or language. Over the past 50 years, the best technology in this area relied on Hidden Markov Models (HMM) [3], with best results obtained using mel-frequency cepstral coefficients (a representation of a sound’s short-term power spectrum) or other coefficients based on perceptual linear prediction [4].

In laboratory environments, high levels of speech recognition accuracies have been achieved, however, accuracy tends to degrade in real-world environments. The majority of problems faced by ASR in real-world environments have predominantly been due to a lack of noise robustness in these systems. While machine learning is starting to be employed in the field of ASR, allowing for further advances to be made in terms of ASR performance and accuracy, the majority of current ASR applications make use of pattern matching models combined with language models and a dictionary to constrain decoding outcomes [5].

The basic speech recognition model (Figure 1) is described as having a front-end which decodes an audio signal into usable frames and then extracts what are known as features from these frames. These features are then passed to a back-end for pattern matching against a language model and dictionary in conjunction with an acoustic model. The result from this pattern matching process is then provided to a ‘decision rule’ based on confidence scoring which then interprets what words have been spoken [6].

![Figure 1 – Basic block representation of a speech recognition system [6]](image)

B. Swarm Robotics and Human-Swarm Interaction

Swarm Robotics (SR) is a broad term that encompasses multi-robot systems consisting of autonomous ‘swarming-robots’ which imitate organisms that exhibit swarm behaviour. The initial concept aimed to mimic natural swarm phenomena such as the flocking of birds, swarming of insects and movements of schools of fish. The emergent behaviour of these so called ‘swarms’ can be described and modelled as the aggregation of many simple interactions between each of the organisms [7]. Humans have imitated these interactions for many years in a military context from formation marching to aircraft formations; however, implementing these behaviours computationally for a swarm of robots has numerous challenges.
The challenges presented by mimicking these behaviours has led to the growth of a number of related fields including Swarm Intelligence (SI) and Adaptive Behaviour (AB). Swarm robotics and its related fields have recently become some of the most popular research areas in wider robotic technology, with cross-pollination from experts in closely related fields such as: artificial intelligence, control theory, robotics, systems engineering and biology [8].

While many systems incorporating swarm robotics are expected to function autonomously and without human interaction, the presence of a human operator can be beneficial or even essential in some scenarios. The field of Human-Swarm Interaction (HSI) investigates methods through which effective human supervisory control of swarms is possible [1]. HSI can be considered a closed loop with the human operator closing the loop by providing commands to a robotic swarm based on information provided by the swarm itself. Research into HSI aims to answer questions relating to communication between the operator and swarm, feedback methods from swarm to operator, and swarm command and control methods.

A key part of HSI research is being able to effectively model the subject robotic swarm. Generally speaking, there are four main categories of swarm models which vary in their inspiration and the environment they seek to model. These categories are: Bio-Inspired, Control Theory, Amorphous Computing and Physics-Inspired. Originally the bulk of research into swarm robotics focussed on the study of biologically-occurring swarms, leading to bio-inspired swarm models. However, additional modelling methods were needed for a range of different applications and subsequent research identified the other three categories of swarm models.

A common example of a bio-inspired model is the Boids model conceived by Craig Reynolds [2] as previously mentioned. This model exhibits flocking behaviour based on the aggregation of three interactions or rules: cohesion – movement towards the centre of a flock, alignment – synchronisation to the average velocity of the flock, and separation – the avoidance of collision with other flock agents. Describing this model in vector form we consider each agent, $A_j$ at a given time, $t$, with a vector position of $\vec{x}_j^t$ and velocity of $\vec{v}_j^t$. On each time-step, the model updates every agent’s velocity based on the following equation:

$$\vec{v}_{j+1}^t = \vec{v}_j^t + W_c \vec{c}_j^t + W_a \vec{a}_j^t + W_s \vec{s}_j^t + \cdots \tag{9}$$

In this equation we consider: $\vec{c}_j^t$, the cohesion unit vector towards the average position of all agents; $\vec{a}_j^t$, the alignment unit vector in the average direction of all agents; and $\vec{s}_j^t$, the separation unit vector away from the average position of all targets. These vectors are scaled by their corresponding magnitudes: $W_c$, $W_a$ and $W_s$. This equation expands as other conditional rules are applied in the simulated environment including: boundary, attract and repel rules.

In some studies, HSI is evaluated in relation to the operator’s cognitive effort in controlling a multi-robot system [10]. If this system consists of multiple robots conducting independent tasks and the operator allocates attention equally between these robots then the resulting cognitive complexity is of linear complexity, $n$, denoted $O(n)$. An example of this could be search and rescue where each robot searches independently of other robots. In many scenarios the cognitive complexity remains constant as instructions are issued to the swarm of robots and then carried out autonomously by this grouping of robots. In such a scenario the resulting cognitive complexity is denoted $O(1)$. It is also conceivable that the cognitive complexity could be worse than both of these cases. For example a system where the interaction between robots is not autonomous, but the task requires coordination of the robots on an individual basis, denoted as $O(> n)$. These relationships are shown in Figure 2.

![Figure 2 - Graphical representation of an operator’s cognitive effort in multi-robot system control [1]](image-url)
C. Command and Control Methodologies

Research into command and control of robot swarm formations has attracted a wide array of attention ranging from military applications [11] to space exploration [12]. In most cases the ability to control the formation of a body of robots with swarm behaviour allows for the system to adapt to accomplish a diverse range of tasks. There are three generally accepted methods of formation control in robot swarms: behaviour-based formation [13, 14], leader-follower [15], and virtual structure methods [16]. Behaviour-based control utilizes a set of basic behaviours such as point-to-point movement and obstacle avoidance which is suitable to highly autonomous swarms. The leader-follower method utilises one or many designated leaders that are followed by the remaining robots. While useful in some situations where some robots form the communications hub for the swarm, this method has been shown to cause collisions if not properly implemented. The final method is virtual structure, which treats the swarm as an inflexible entity and sets strict requirements in maintaining the formation.

From a human operator’s perspective, the key to a control method that ensures a cognitive complexity of \( O(1) \) is the classification of the robot swarm as a single entity from a command perspective, often meaning high levels of inter-swarm autonomy. This may not always be possible in scenarios where the swarm must be split into smaller sub-swarms, or where some split in swarm functionality is required; however, such a view of the swarm of robots as a single, more capable entity caters for a different approach to command and control methodologies. The main limitation to this approach is the level of automation of the robot swarm (its ability to deal with the complex interactions between robots) with the swarm’s Level Of Automation (LOA) defining the level of human interaction required [17].

While this research provides some guidance on the implementation of a command & control methodology, research directly related to the implementation of a speech-based methodology in a HSI context was not identified in the literature review undertaken.

IV. Research Significance

Effective methods of Human-Swarm Interaction (HSI) are becoming more critical as robotic swarms become more common and technology increases the possible size (being number of robots) of robotic swarms. The majority of research within this field has focussed on investigating effective communication and swarm state estimation/visualisation [1, 18, 19]; however, the demands on human operators of multi-robot systems remains an area that has not been adequately addressed.

Many methods and permutations of HSI have already been researched and developed including joystick, keyboard/mouse and touch screen [9, 20, 21], however, based on the literature review carried out as part of this project, little apparent research and/or experimentation has involved the use of speech-based command and control. One paper written by a team from the University of Málaga, Spain [22], investigated the speech teleoperation of a mobile robot supplemented by a keyboard and mouse; however, this command and control method has not yet been applied to a robotic swarm.

Speech command is a natural method through which humans convey instructions and information to other human beings, with the potential command complexity of this method only limited by the dictionary of the ASR system and the actual command structure implemented. With the state of ASR technology such that the dictionary of these systems is sufficiently large to not limit the complexity of commands, the main limiting factor of this method is the implemented command structure.

While an argument can be made that physical methods of control such as joystick, keyboard/mouse etc. allow for more precise control of a group of robots, most applications of swarm robotics rely on a level of automation on the part of the individual robots, such that the human operator only provides a goal for the swarm to achieve, not necessarily precise instructions. This is particularly true in terms of the movement of a robot swarm in models such as Craig Reynolds’s Boids model where simulated robots are constrained by the set of accumulative rules that define the formation and movement of the swarm. The significance of this project will be in investigating the feasibility of speech-based HSI as a potential method of command/control of a robotic swarm.
V. Speech-Based HSI

A. Speech-Based Command Sets

Within the scope of this project, two different speech-based command methods were proposed and evaluated, comparing their relative performance: Lateral Movement Control (LMC) and Point-Effect Control (PEC).

**Lateral Movement Control** (LMC) provides the operator with basic control over the lateral movement of the simulated swarm. LMC was implemented through four basic commands which moved the swarm laterally through the simulation space. This lateral movement is induced through the addition of a directional bias vector with magnitude scaled by the constant \( c \) and direction defined by the associated unit vector in a two-dimensional space. The proposed command set is described in Table 1:

<table>
<thead>
<tr>
<th>Command Word</th>
<th>Directional Bias Vector</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘NORTH’</td>
<td>+c( \hat{y} )</td>
<td>Moves swarm laterally upwards in the simulation space.</td>
</tr>
<tr>
<td>‘SOUTH’</td>
<td>−c( \hat{y} )</td>
<td>Moves swarm laterally downwards in the simulation space.</td>
</tr>
<tr>
<td>‘EAST’</td>
<td>+c( \hat{x} )</td>
<td>Moves swarm laterally rightwards in the simulation space.</td>
</tr>
<tr>
<td>‘WEST’</td>
<td>−c( \hat{x} )</td>
<td>Moves swarm laterally leftwards in the simulation space.</td>
</tr>
<tr>
<td>‘CLEAR’</td>
<td>N/A</td>
<td>Clears current direction.</td>
</tr>
</tbody>
</table>

The grammar for LMC allows for up to two of these commands to take effect at any one point in time. Speaking a single command activates the single effect, with the next command spoken replacing the previous spoken command. Speaking two commands without a pause combines the effects of the two commands, allowing for diagonal movement of the swarm. It should be noted that combining two inverse commands (e.g. ‘NORTH’ and ‘SOUTH’) results in no net effect on the simulation. Some examples of this grammar are listed in Table 2:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>‘NORTH’</td>
<td>‘NORTH’ &lt;PAUSE&gt; ‘WEST’</td>
<td>‘NORTH WEST’</td>
<td>‘NORTH SOUTH’</td>
</tr>
<tr>
<td>Effect: Swarm moves laterally upwards.</td>
<td>Effect: Swarm moves laterally upwards and then laterally left on receipt of the second command.</td>
<td>Effect: Swarm moves diagonally up-left.</td>
<td>Effect: No effect on swarm.</td>
</tr>
</tbody>
</table>

This command method was the first implemented to assess the effectiveness of speech-based control with relatively low complexity in terms of command set and grammar.

**Point-Effect Control** (PEC) allows the operator to place points within the simulation which have a specified effect on any Boids within their effective radius. Two effects are available for these points: attraction and repulsion. The attraction effect has an effective radius which covers 100% of the simulation space, attracting the swarm from any position, whereas the repulsion effect has an effective radius of 20% of the simulation space. PEC was implemented through a set of commands that allowed for the selection of the desired effect and for the placement of an effect point at a specified co-ordinate within the simulation. The command set is listed in Table 3:
Table 3. PEC Command Set

<table>
<thead>
<tr>
<th>Command</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘BRAVO’</td>
<td>Select ‘Attract’ Effect</td>
</tr>
<tr>
<td>‘INDIA’</td>
<td>Select ‘Repel’ Effect</td>
</tr>
<tr>
<td>‘DELETE’</td>
<td>Select ‘Delete’ Mode</td>
</tr>
<tr>
<td>‘&lt;INTEGER&gt; &lt;INTEGER&gt;’</td>
<td>Place Point at co-ordinate [INTEGER,INTEGER]</td>
</tr>
<tr>
<td>‘CLEAR’</td>
<td>Clear All Points</td>
</tr>
</tbody>
</table>

The grammar of this control method allows for switching between point effects, placement of these points and for the clearing of all previously placed points. Selecting an effect by using commands such as ‘BRAVO’ or ‘INDIA’ switches the effect mode, meaning that points placed after that command have the selected effect. Any points placed will remain active with their originally placed effect until the ‘CLEAR’ command is called, at which point all currently active points are removed, or until the delete mode is selected and a particular point’s position is spoken. Some examples of this grammar are listed in Table 4:

Table 4. PEC Command Grammar Examples

<table>
<thead>
<tr>
<th>Example</th>
<th>Words Spoken</th>
<th>Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Example 1.</td>
<td>‘BRAVO’ &lt;PAUSE&gt; ‘Three Three’</td>
<td>‘Attract’ Effect is selected and an ‘Attract’ point is placed at co-ordinates [3,3].</td>
</tr>
<tr>
<td>Example 2.</td>
<td>‘INDIA’ &lt;PAUSE&gt; ‘Four Three’</td>
<td>‘Repel’ Effect is selected and a ‘Repel’ point is placed at co-ordinates [4,3].</td>
</tr>
<tr>
<td>Example 3.</td>
<td>‘Five Eight’</td>
<td>A point is placed at co-ordinates [5,8] with the point’s effect determined by the currently selected effect.</td>
</tr>
<tr>
<td>Example 4.</td>
<td>‘INDIA’ &lt;PAUSE&gt; ‘Two Seven’ &lt;PAUSE&gt; ‘BRAVO’ &lt;PAUSE&gt; ‘Eight One’</td>
<td>‘Repel’ Effect is selected and a ‘Repel’ point is placed at co-ordinates [2,7]. ‘Attract’ Effect is then selected and an ‘Attract’ point is placed at co-ordinates [8,1].</td>
</tr>
</tbody>
</table>

This command method represents a different control method which is aimed at enabling more precise control of the simulated swarm, at the cost of a more complex command set and grammar.

B. Simulation and Evaluation

The simulation utilised in evaluation of the proposed command sets integrated: ASR, command parsing and Boid simulation. This simulation allowed for the evaluation of speech-based HSI in a simulated scenario which required the operator to navigate the swarm from a starting point to an objective, solely utilising speech as the method of control. The simulation was fully developed in Java, making use of Carnegie Mellon University’s Sphinx-4 package for ASR and Drew Heavner’s Universal Java Flocking Engine for swarm simulation and GUI. Code to interface the two packages and to implement the command sets was developed and the GUI/flocking engine was heavily modified for the purposes of evaluation.

In order to evaluate the command sets, environmental complexity was simulated in the form of ‘no-fly zones’. These zones destroyed any Boids that entered them and allowed for the ability of the command set to handle environmental complexity to be assessed in the form of survival rate, calculated from the ratio of Boids completed the task to Boids destroyed. In this evaluation, environmental complexity was iterated in the form of difficulty levels, ranging from one to three as detailed in Table 5:

Table 5 - Environmental Complexity (Difficulty) Levels

<table>
<thead>
<tr>
<th>Difficulty Level</th>
<th>Number of Zones</th>
<th>Maximum Zone Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>One</td>
<td>5</td>
<td>4% of total simulation space</td>
</tr>
<tr>
<td>Two</td>
<td>10</td>
<td>4% of total simulation space</td>
</tr>
<tr>
<td>Three</td>
<td>10</td>
<td>9% of total simulation space</td>
</tr>
</tbody>
</table>
The performance of the command sets was also assessed based on their ability to navigate the swarm in a timely manner, with the total task completion time recorded as the second performance metric. Evaluating performance against these two measures allowed for a speed versus accuracy assessment of the command sets. An example of the experimental setup is shown in Figure 3.

![Speech-Based HSI Simulation Screenshot (labelled)](image)

**Figure 3 - Speech-Based HSI Simulation Screenshot (labelled)**

**VI. Experimental Results and Discussion**

A. Characterisation

The first phase of experiments focussed on the characterisation of the simulation in terms of speech recognition and Boid response to commands.

The Sphinx-4 ASR package was first assessed to determine its speech recognition accuracy for the set of words to be utilised in the command sets. Initially the recognition accuracy of spoken integers between one and nine was characterised. Initial trials suggested that the unmodified Sphinx-4 package recognition accuracy was not sufficient to allow for accurate and timely control of the swarm in the simulation. In order to improve this accuracy homonym detection logic was added at the output of the ASR package to detect words of similar phonetics. An example of this is the word ‘Two’ which is often mistaken for the words ‘To’, ‘Too’ and even ‘Tooth’. Utilising this technique, suitable recognition accuracies were achieved as shown in Figure 4:
This technique was applied to all other commands utilised in this project to achieve the accuracies shown in Figure 10 (Appendix).

The reaction of the Boid swarm to the two command sets: PEC and LMC were then characterised through case studies. The ‘Attract’ and ‘Repel’ effect points were first characterised by plotting the average Boid distance from the effect points recorded over time. The PEC Attract response is shown in Figure 5.

The PEC ‘Repel’ effect and LMC case studies can be found as Figure 11 and Figure 12 of the appendix, respectively.

The recorded data from these case studies was then utilised to determine an end-to-end response time including ASR, command parsing and swarm response. This data provided a total estimated response time from the first utterance of a command to the desired ‘final’ state of the swarm, and is plotted in Figure 6.
This data suggests that while the total reaction time is upwards of ten seconds, the majority of that time is the swarm actively reacting to the command and the actual time from the first utterance to the detection of the command and simultaneous placement is less than three seconds.

B. Command Set Evaluation

The two command sets were evaluated in a series of experiments in which swarm size and environmental complexity were varied to determine their effect on command set performance. The results were averaged across ten trials to obtain higher levels of accuracy:

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Swarm Size</th>
<th>Difficulty Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>20</td>
<td>1</td>
</tr>
<tr>
<td>1.2</td>
<td>20</td>
<td>2</td>
</tr>
<tr>
<td>1.3</td>
<td>20</td>
<td>3</td>
</tr>
<tr>
<td>2.1</td>
<td>40</td>
<td>1</td>
</tr>
<tr>
<td>2.2</td>
<td>40</td>
<td>2</td>
</tr>
<tr>
<td>2.3</td>
<td>40</td>
<td>3</td>
</tr>
<tr>
<td>3.1</td>
<td>60</td>
<td>1</td>
</tr>
<tr>
<td>3.2</td>
<td>60</td>
<td>2</td>
</tr>
<tr>
<td>3.3</td>
<td>60</td>
<td>3</td>
</tr>
</tbody>
</table>

On completion of these experiments the results of each experiment were averaged and plotted in terms of completion speed and survival rate between the two command methods. The results are shown in Figure 8.

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Swarm Size</th>
<th>Difficulty Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>20</td>
<td>1</td>
</tr>
<tr>
<td>1.2</td>
<td>20</td>
<td>2</td>
</tr>
<tr>
<td>1.3</td>
<td>20</td>
<td>3</td>
</tr>
<tr>
<td>2.1</td>
<td>40</td>
<td>1</td>
</tr>
<tr>
<td>2.2</td>
<td>40</td>
<td>2</td>
</tr>
<tr>
<td>2.3</td>
<td>40</td>
<td>3</td>
</tr>
<tr>
<td>3.1</td>
<td>60</td>
<td>1</td>
</tr>
<tr>
<td>3.2</td>
<td>60</td>
<td>2</td>
</tr>
<tr>
<td>3.3</td>
<td>60</td>
<td>3</td>
</tr>
</tbody>
</table>
Both completion time plots indicated an increasing trend in completion time for increasing environmental complexity with no evidential trend in completion time for increasing swarm size. These plots also reveal a significant difference in completion time between the LMC and PEC command sets with LMC completing the scenario, on average, 35 seconds faster or, expressed as a percentage, the PEC command set took almost twice as long.

These results are not surprising considering that with the LMC command set, the swarm is not ‘stationary’ around a single point at any given time, but are instead continuously moving towards the objective based on the directional bias. The increasing completion time with environmental complexity is intuitive with a greater amount of time required to navigate through obstacles.

Both survival rate plots (Figure 9) indicate a decreasing trend in survival rate for increasing environmental complexity with no evidential trend in survival rate for increasing swarm size. The plots show a noticeable difference in survival rate with the PEC command set offering a 15% increase in survival rate, on average, compared to the LMC command set.

The results obtained are not surprising considering that the PEC command set allows for more precise control of the swarm via the placement of effect points as opposed to laterally moving the swarm. The decreasing survival rate with increasing complexity is not surprising either considering the presence of more obstacles of greater size is likely to result in larger proportions of the swarm encountering them.

**VII. Conclusion and Future Work**

This paper explored the feasibility of speech-based HSI in a simulated environment. Two speech-based command sets were proposed and evaluated in terms of scenario completion speed and survival rate for varying environmental complexity and swarm size. The result of this evaluation suggested that the Lateral Movement Control (LMC) method enabled faster completion of the scenario at the cost of a reduced survival rate whereas the Point Effect Control (PEC) method allowed for a higher survival rate at the cost of completion speed.

This paper offers a preliminary assessment on the feasibility of a solely speech-based approach to swarm control. Future work related to this subject could investigate the refinement of the PEC method into a waypoint-based command potentially allowing for faster completion times or a combination of both command sets to harness the benefits of each individual set. Future work could also focus on a cognitive load or general performance comparison of speech-based HSI to existing methods such as keyboard/mouse or touchscreen.
Acknowledgements

I would like to acknowledge my project supervisor, Kathryn Merrick, for guiding me through my final year project and ensuring I kept a steady pace throughout the year. I would also like to acknowledge Des Boorman and Holly Woodford for their support and assistance in writing this paper.

References

Appendix A.

Figure 10 - Speech Recognition Accuracies of Utilised Commands

Figure 11 - Repel Effect Response Case Study

Figure 12 - LMC Response Case Study