

5.4. Speech and Person Recognition

The robot has to identify unknown people and answer questions about them and the environment.

5.4.1. Focus

This test focuses on human detection, sound localization, speech recognition, and robot interaction with unknown people.

5.4.2. Setup

1. **Location:** One room of the arena is used for this test³.
2. **Crowd:** There is a crowd of 5 to 10 people in the designated room. People may be standing, sitting, lying, and in any pose.
3. **Doors:** All doors of the apartment are open, except for the entry door.

5.4.3. Task

1. **Start:** The robot starts at a designated starting position and announces it wants *to play riddles*.
2. **Waiting and turn:** After stating that it wants *to play a riddle game*, the robot waits for 10 seconds while a crowd is merged on it's back. When the time elapses, the robot must turn around (about 180°) and find the crowd.
3. **Requesting an operator:** After turning around, the robot must state the size of the crowd (including male and female count⁴) and request for an operator (e.g. *who want to play riddles with me?*). The crowd will move and surround the robot, letting the operator to stand in front of the robot.
4. **The riddle game:** Standing in front of the robot, the operator will ask 5 questions. The robot must answer the question without asking confirmation. Questions will only be asked only once; no repetitions are allowed.
5. [DSPL only] **Blind man's bluff game: Crowd line-up.** The crowd will reposition, lining up in front of the robot. A random person from the crowd standing in front of the robot will ask a question. The robot may
 - Turn towards the person who asked the question and answer the question
 - Directly answer the question without turning
 - Turn towards the person and ask them to repeat the question

This process is repeated with 10 (possibly) different people. The game will end when the 10th question has been made, following a similar distribution of questions as in the riddle game. The robot must answer the question without asking confirmation. Questions may be repeated once.

³This test may also be held outside the arena

⁴It is possible to state the number of people whose gender couldn't be determined by the robot, therefore stating correctly the size of the crowd and, possibly, one of the gender groups.

5. [OPL & SSPL] **Blind man’s bluff game: Circling Crowd.** The crowd will reposition, making a circle around the robot. A random person from the crowd surrounding the robot will ask a question. The robot may
 - Turn towards the person who asked the question and answer the question
 - Directly answer the question without turning
 - Turn towards the person and ask them to repeat the question

This process is repeated with 5 (possibly) different people. The game will end when the 5th question has been made, following the same distribution of questions as in the riddle game. The robot must answer the question without asking confirmation. Questions may be repeated once.
6. **Leave** The robot must leave the arena/test area after all questions have been asked or when instructed to do so.

5.4.4. Additional rules and remarks

1. **Bypassing ASR:** Bypassing Automated Speech Recognition via the CONTINUE rule (see Section 3.9.1) is not allowed during this test.
2. **Asked questions:** The distribution of questions to be randomly asked is as follows:
 - One is a predefined question
 - Between one and two are about the arena and its status
 - Between one and two are about the crowd
 - Between one and two are about the list of official objects

Question examples see Appendix Section A. Questions are generated by the generator which is made publicly available at <https://github.com/kyordhel/GPSRCmdGen>.

3. **Distance to the robot:** The distance between each person and the robot must be between 0.75 and 1.0 meters away from the robot position (See Figure 5.3). In the *riddle game* the operator shall be between -60° and 60° from the robot’s center (front range).
4. **Precise turning:** When the robot finishes turning toward an operator, it must be clear that the robot is facing the person who made the question.
5. **Question repetition:** In the *blind man’s bluff game*, if the robot asks for repetition, it should be done clear and loud, and after the robot has ended turning.
6. **Question timeout:** If the robot does not answer within 10 seconds, the question is considered as *missed*, and referee will proceed with the next one.
7. **Standing still operators** Operators are not allowed to move to or turn towards the robot or shout to the robot.
8. **Water-clear answers:** If the referee is unable to hear or understand the robot’s answer, the question is considered as *incorrect*. Single-word and short answers should be avoided

5.4.5. Referee instructions

The referee needs to

- avoid shouting to the robot
- avoid getting closer to the robot (or even move)

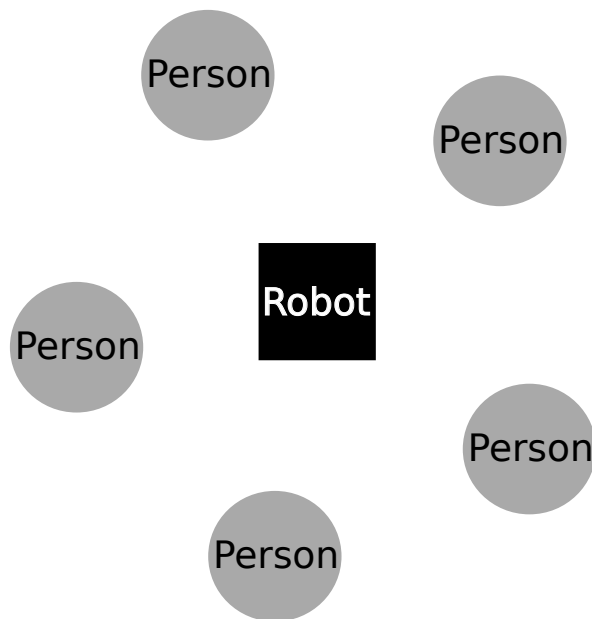


Figure 5.3.: Speech recognition test: person setup around the robot for 2nd part.

- speak to the robot loud and clear with plain standard English
- avoid repeating questions for the same robot
- distribute the questions among the volunteers

5.4.6. OC instructions

1 day before the test

- Provide the set of predefined questions

2 hours before the test

- Announce the placement of the robots
- choose the volunteers for the second part of the test, and clearly explain the procedure to them.
- When the test is held outside the arena, announce the (way)point through which the robot shall leave

5.4.7. Audio Data Recollection

Teams are encouraged to submit to the TC the audio data recorded during the test, specially that which was captured during speech recognition. If so, teams are urged to provide it with annotation of what the user said and what was recognized. Audio files are expected to be mono, one per microphone (in the case array recordings), of a sample rate equal to or higher than 16 kHz, and with a sample size of at least 16 bits. Depending on the quality of the recordings and their annotations, an official certificate that formalizes these efforts may be provided to submitting teams.

5.4.8. Score sheet

The maximum time for this test is 5 minutes.

Action	Score
<i>Crowd</i>	
State crowd's size	5
State crowd's male/female count	10
<i>Riddle game</i>	
Understanding question	5 × 5
Correctly answered a question	5 × 5
Answering all 5 riddle game question	5
<i>[DSPL only] Blind man's bluff game</i>	
Understanding question on the first attempt	10 × 5
Understanding question on the second attempt	10 × 2
Correctly answered a question	10 × 2
Turned towards person asking the question	10 × 5
Answering all 10 blind man's bluff questions	10
<i>[OPL & SSPL] Blind man's bluff game</i>	
Understanding question on the first attempt	5 × 10
Understanding question on the second attempt	5 × 5
Correctly answered a question	5 × 5
Turned towards person asking the question	5 × 10
Answering all 5 blind man's bluff questions	5
<i>Special penalties & standard bonuses</i>	
Not attending (see sec. 3.10.1)	-50
Outstanding performance (see sec. 3.10.3)	20
Total score (excluding penalties and bonuses)	200