SOCIAL ROBOTICS

First Time Meeting With A Robot

Using Backchannel Head Nodding

Group mta141038

27th of June, 2014
Master's Thesis

AALBORG UNIVERSITY
STUDENT REPORT

Department of Architecture, Design, and Media Technology
Sofiendalsvej 11
DK-9200 Aalborg S
http://create.aau.dk

Title:
Social Robotics - First Time Meeting With A Robot Using Backchannel Head Nodding

Theme:
Master's thesis

Project Period:
09 2013 – 06 2014

Project Group:
mta141038

Participant(s):
Daniel Grønkjær Jensen

Supervisor(s):
Matthias Rehm

Copies: 2

Page Numbers: 59

Date of Completion:
27/06 - 2014

Abstract:

In HRI backchannel head nodding have been found to be an important part when a human speaker is talking to a robot listener. So in this project the focus will be on first time meetings betweeen a human speaker and a robot listener only responding with non-verbal backchannels (head nodding). A replication of an earlier study is conducted with Danish participants, and the results contradicts the original study. The importants of embodiment is also investigated, and the results indicate that the embodiment will effect the perception of the agent. An evaulation of culturally appropriate head nodding from different agents was also conducted, and the perception of a humanoid robot using backchannel head nods is perceived as more intelligent.

Contents

Preface

This report is the product of the Master's Thesis by group mta141038 covering the 9^{th} and 10^{th} semester in Medialogy at Aalborg University from the period September 2013 to June 2014. The report is aimed for the supervisor, the censor and others who may be interested in reading the following. The report covers an area of social robotics (HRI) with focus on first time conversation between a human speaker and a artificial listener giving feedback in the form of head nodding only.

The source of references are given by a number enclosed by brackets – eg. [1] which can be found in the Bibliography in the end of the report.
In the bibliography the source are given by the authors(s) name, title, publisher, publication year, edition. - e.g.:

[19] V. H. YNGVE. "“On getting a word in edgewise"”. In: Chicago Linguistics Society 6th Meeting (1970)

On the attached DVD the software for My Keepon, and the data from the two experiments at Aalborg University and SOSU Nord. The AV production can be found on the DVD, too.

The group would like to thank Jens Vilhelm Dinesen from SOSU Nord Future Lab for the collaboration with experiment at SOSU Nord with the Telenoid. A thanks to Xtel for the development of the software used for pilot study at SOSU Nord for the Telenoid. Lastly a big thanks to all the participants in the two pilot studies and the two main studies.

Aalborg University, June 26, 2014

Daniel Grøkjær Jensen
<dgje08@student.aau.dk>

Chapter 1

Introduction

In this project the focus will be on human-robot interaction (HRI) - more specifically on the effect of culturally appropriate feedback signals from a variety of robot embodiments in first time conversations with a human where the robots will act as the listener.

As we see robots move from restricted environments that are specially designed for them (e.g. industrial robots and robotic surgery) into the homes of ordinary people and being used in eldercare the interest for social robotics is gaining speed.

People are already familiar with the idea of robots acting in a social manner from movies e.g. *2001: A Space Odyssey* and *Star Wars* - movies that are more than 35 years old. In these movies humans are able interact with robots through everyday conversation. This is an indication that we as humans picture a future with robots that we are going to interact with as trusted entities in our everyday life. One of the reason that the audience might not be put off by humans communicating with robots using everyday conversations could be due to the human-like adapted behaviour shown by the robots. It would therefore be interesting to investigate the effect of having a robot that is able to abide by the conventions from human-human conversation in real time conversations with a human.

But these conventions in conversation varies based on peoples cultural background, and it will therefore be interesting to investigate whether or not a robot will have to adapt to the cultural setting that the robot is placed in.

The idea of robots being able to act in a socially acceptable manner is not the only issue social robotics is facing; the embodiment of robots is also a key aspect if we wish to achieve a sustainable relationships between humans and robots. In the healthcare system we have seen a broad variety of different embodiments designed to interacted with a specific target group, and this opens up for the opportunity investigate whether certain features in a robot will effect peoples opinion when used in a different scenario, in this case it would be first time face-to-face conversation.

One of the bigger challenges with social robotics, and thereby the ability to isolate and replicate studies, is found in the method used to investigate social behaviour in robots - *The Wizard of Oz experiment*. This method requires a human facilitator to directly control the robots actions, and it is therefore prone to variation from sessions. So it is naturally intriguing to make an autonomous system, and see how it compares to the case where the robot is controlled by a human facilitator. An autonomous system could be based on the information retrieved from human-human interaction and other studies of similar nature.

Another argument in favor for the development of autonomous systems for human-robot interaction is the emergence of cheap and reliable hardware (Microsoft Kinect, Tobii Eye-trackers, high resolution cameras) allowing developer the opportunity to create software capable of detecting human features e.g. body movement, gestures, facial expressions, etc. All of this combined with the mentality of open source communities is forging the path for low priced systems that can be used in the social scientific field which earlier was of limited due to the technical expertise that was required for development of autonomous systems.

So the research question(s) for this project will be based on the prior understanding of the field of social robotics just covered.

1. *Can culturally appropriate feedback signal help to maintain first time face-to-face conversation with an artificial listener for a longer duration of time?*
   (a) *And will the speakers perception of the robot change when using cultural appropriate feedback signals?*
2. *Will the embodiment of the artificial listener matter in relation to maintaining first time face-to-face conversation for a longer duration of time?*
   (a) *And will the speakers perception differ based on embodiment of the robot?*

3. *Is it possible to develop an autonomous system capable of generating real-time feedback signals during a face-to-face conversation that will help to maintain it in longer time?*

Chapter 2

Related work

One of the objectives with this project is to replicate and extend the research done by Koda et al. [11] where the effect of cultural appropriate head nodding from a virtual agent during human-robot communication were investigated and Koda et al. concluded that it was important to consider the cultural differences based on the findings which indicated that the participants of the study had longer interaction sessions with the agent that had the same cultural background as the participants. In the study by Koda et al. all the participants where from Japan and it would therefore be interesting to extend the research by doing the same experiment in other countries. A pilot study was conducted before the main study in order to determine the parameters for a head nod, that people perceived as giving the highest sense of listening [11].

Krosager et al. [23] investigated the role of the physical embodiment of the listener agent in a similar setup as Koda et al. [11] done with danish participants. In the experiment by Krosager et al. [23] a humanoid robot (Nao) was used instead of a virtual agent. Unlike the findings by Koda et al. [11] the physical agent elicited shorter speech duration when culturally appropriate head nodding was used. Based on this a second study were employed to investigate if the physical present of a robot might be the cause of the contradicting findings. It was found that the physical present of the robot being in the same room as the speaker influenced the speech duration.

Since the focus will be on first time conversations with an listener agent utilizing culturally appropriate head nodding information regarding the following areas will be presented in the next part of the report. A review of the social signals found during human-human conversation and how this might vary across cultural differences.

2.1 Conversation - form and function

During human-human interaction e.g. conversation the listener consciously and subconsciously produce a variety of different feedback signals that are used to communicate to the speaker that you (cf. the listener) you are listening without interrupting the speaker. Theses social signals produced by the listener are called backchannels, and they can refer to vocal (e.g. 'yes', 'mm-mm', and 'uh-huh') or visual (e.g. nodding, gaze, gestures, and body language) signals that are minimal and non interruptive [24] [9]. Since the terms for theses social signals where coined in the 70s [24] it has been investigated by a variety of different branches of sciences in order to better understand human-human interactions, and it has during the last few years entered the research area of human-robot interaction.

One of the reason that backchanneling is interesting for human-robot interaction is that findings indicate that discourse is a joint activity ([3], [9], [10]), and it is therefore important that the robots abide by the rules of face-to-face conversation [10] that we as humans expects.

A general division of listener responses is made by Bavelas et al. They discuss two different kinds of listener responses in 'Listeners as Co-Narrators' [3]; *generic* and *specific* listener responses. Generic responses are not specifically connected to the utterance of the speaker whereas specific responses are linked directly to what the speaker is saying. So in order for a specific listener response to come at the correct time during a conversation the context of what the speaker is saying is relevant. This is something that most people can do without much effort, but for a robot it is a completely different story. It would require a system that is able to interpret the spoken language, and generate an appropriate feedback signal. The generic response on the other hand follows some more, as the name implies, generic rules e.g. pauses between sentences. In order to generate generic responses some general guidelines have to follow if it is not going to be completely random. The quantity and timing of backchannels has been investigated ([18], [17]) in a setup with the participants rating video fragments of an human speaker and an artifical listener based on how human-like the backchannel behaviour is perceived. It was found that the quantity will have an effect on the perception of the listener, and in [18] it was stated at more backchannels are better for a human-like listener, but in [17] it was concluded that more are not necessarily better. In both cases it was found that the timing was important for the perception of the listener, and that random backchanneling was perceived as less human-like. This supports the idea of having culturally appropriate head nodding since this would control the quantity and timing.

Another reason to have social signals incorporated is the ability to build rapport between a virtual or physical agent and the speaker with the use of *simple* backchannels [10], and through the increased engagement getting longer interaction sections ([21], [22]). This can be useful in healthcare where robots are being used more often than earlier in order achieve a higher life quality for the patients [8].

As mentioned in section 2 one of the aims is to replicate the study by Koda et al. [11], and in their experimental setup the only type of backchannel were non-verbal in the form of head nodding. A pilot study is conducted to determine the parameters for the head nod giving the participants the greatest sense of listening. The parameters regarding the head nod can be used to distinguishing between different types of head nods [9]. If the head nod's parameters are not adjusted to give the speaker a sense of listening the speaker might interpret the feedback incorrectly seeing that a head nod can also be used by the listener to signal a floor change (the speaker and the listener changes roles) [9]. Though it seems simple the head movement can convey a great deal of information based solely on the direction such as side to side movement in most case signals negation and up/down movement signals affirmation [14] and other distinguishes for the head movement covers terms as jerk, tilt, waggle which all conveys different information [1] [16]. For this project the up/down head movement will be used, and the parameters repetitions, angle, and speed will be examined before the main study.

2.2 Cultural diversity

It is common knowledge that human behavior varies based on cultural background, and this is also the case regarding the use of backchannel head nodding in first time meetings. In the study [11] the cultural appropriate head nodding is based on the conversation analysis by Maynard [13] based on Japanese dyadic interaction. Through the analysis it is found that Japanese perforn backchannel head nodding 10.4 nods/minute or a head nod on average every 5.75 seconds where Americans only have 2.7 nods/minutes or a head nod on average every 22.5 seconds. In order to replicate the study [11] with participants from Denmark some information regarding Danish behavior in first time conversation meetings is needed. The NOMCO corpus [15] and CUBE-G corpus [19] is included for this reason as this is video of first time meetings between Danish participants. The CUBE-G corpus reveals a frequency of 7.1 nods/minute or a head nod on average every 8.45 seconds where the NOMCO corpus has a head nod on average every 5.82 seconds [16]. This will be used as the basis in the replication of the study [11].

2.3 Robot embodiments

Within the research field of social robotics there are now a variety of different kinds ranging from humanoid robots (e.g. Nao and Telenoid) to more animal-like (e.g. Keepon Pro and My Keepon), and robots that completely resembles animals (e.g. PARO and Pleo). The most common use of social robots are for people with disabilities or elderly people.

Chapter 3

Experimental systems

For this project four different listener agents will be used

- The virtual agent developed by Koda et al. [11] (will be referred to as *Cat*),
- the physical robot *My Keepon*,
- the physical humanoid robot *Nao* and
- the physical humanoid robot *Telenoid*.

The Cat is included in order to replicate the study by Koda et al. only with Danish participants to compare it to the results they got. My Keepon is added to have a physical agent that is not a humanoid robot. It also has some resemblance to the Cat with the anime-style looks to them. Lastly the Nao and Telenoid is included to have to different kinds of physical humanoid agents that are currently being used in research and aimed for the commercial marked (education and healthcare). This variety of agents will be used to determine if the embodiment of the agent will have an effect on the results found in the studies.

The Cat being the system to mimic as closely as possible to get comparable results from the experiments will be used as the bases for the UIs developed for the other systems that are to be used in the pilot study I will get into in the next chapter. But before any studies can be conducted the systems for the agents has to be developed. As mentioned the Cat is already fully functional for the experimental setups and the Nao has been used earlier [23], but UIs will be needed for My Keepon and Telenoid. So the next part will be a break down of the Cat system and the UIs developed for My Keepon and Telenoid.

3.1 Virtual Agent (Cat)

The user input for manipulating the virtual agent can be seen in figure 3.1 as frequency, nod speed, and nod angle. In the list below the range and keyboard input are shown:

Change frequency: 1, 2, or 3

Keyboard input: [+] and [-]

Change nod speed: 0.4, 0.6, or 0.8 sec.

Keyboard input: [→] and [←]

Change nod angle: 1, 2, or 3

Keyboard input: [↑] and [↓]

The reason for having similar UIs for My Keepon and the Telenoid is that it will make it faster and easier for the participants to move between the UIs during the pilot study without having to learn a new way of controlling the different agents. A different approach was tried for My Keepon, which will be presented in the next section (see section 3.2).

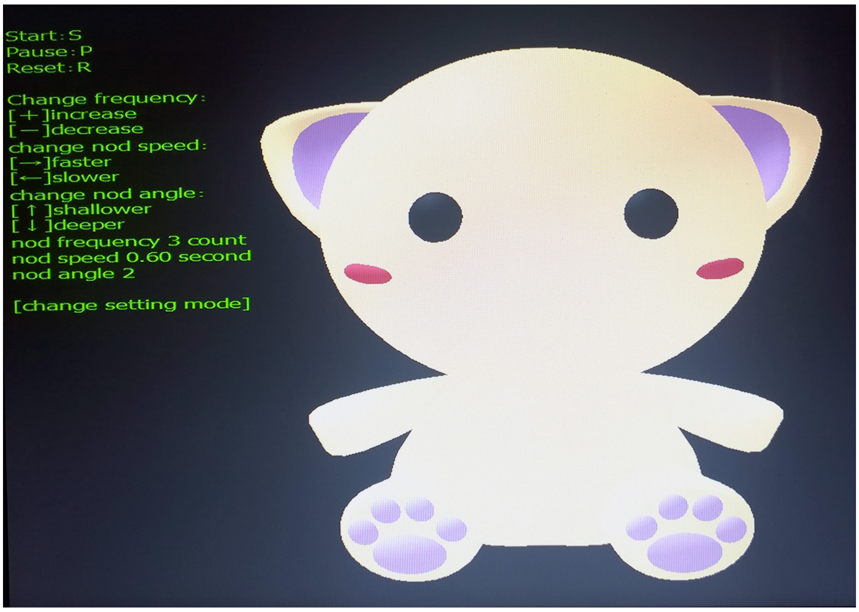


Figure 3.1: An overview of the virtual agent developed by Koda et al. [11]

The Cat will be running on a laptop connected to an external screen, and having only the screen visible for the participants in order to minimize the feeling of talking to a computer

Figure 3.2: The first UI for choosing My Keepon's head nod

for the main experiment, but during the pilot study a keyboard will be in front of the participants in order to change the parameters.

3.2 My Keepon

For My Keepon two different UIs were developed and both have been tried out with participants. The first one were developed with the idea of presenting all the different head nods in a grid grouping it using the three parameters range of motion, velocity, and angle, as shown in figure 3.2.

This UI turned out to be unreliable and would lose connection to My Keepon while participants were using it. The second reason why another UI were developed is the issue of having two very different UIs for manipulating the same thing, as mentioned earlier, which will only add extra time and more explanation for the participants. The second UI for My Keepon therefore took the inspiration from the UI developed by Koda et al. [11], since changes for the interface for the virtual agent would require more time compared to development for My Keepon. The second UI can be seen in figure 3.3 and it uses the same keyboard inputs as the virtual agent (Cat). For a full guide of how to develop for My Keepon see appendix A.

The final setup for My Keepon can be seen in figure 3.4. The main reasoning behind

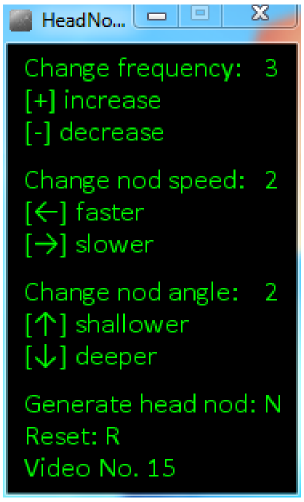


Figure 3.3: The second UI for choosing My Keepon's head nod with inspiration from Koda et al. [11]

placing My Keepon on a table with the black part with all the motors hidden was, as with the Cat, to give the participants the feeling of the robot being independent of a computer controlling it and thereby strengthening the illusion of the robot actually being *alive*.

3.3 Nao

For the Nao the system and data collected by Krosager et al. [12] were used for the experimental setup. The head nod with the highest perceived sense of listening for the Nao was determined with an online video survey and was found to be 3 repeated nods with a nod angle of 8 degrees executed in 2 seconds.

3.4 The Telenoid

The Telenoid was added to the study through collaboration with Jens Vilhelm Dinesen from SOSU Nord Future Lab.

The UI for the Telenoid were developed by Xtel (a software company located in Aalborg). The reason that an external software company was brought in is due to some paper work SOSU Nord has regarding the development of new software for the Telenoid made with the developer of the Telenoid. Through e-mail correspondence and a meeting the

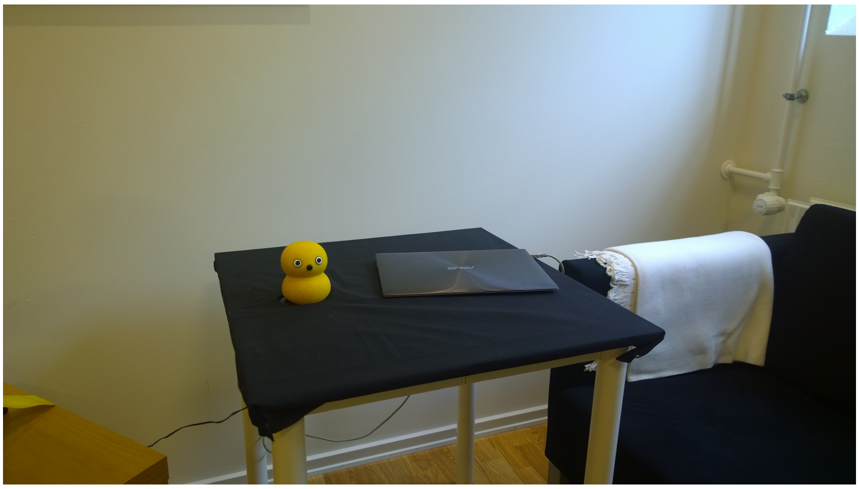
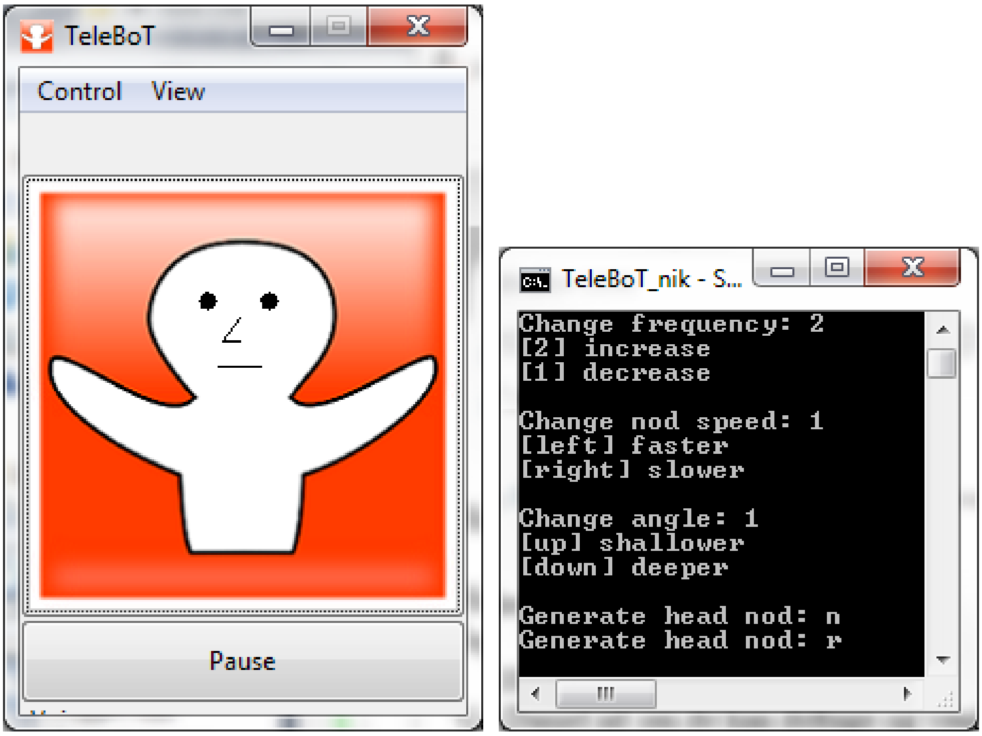


Figure 3.4: My Keepon's setup for the pilot study with the computer visible.

requirements for the UI was established. The software can be seen in figure 3.5 and it consistes of two parts. The original software is still running, but with an added window showing the settings for the head nod just like the Cat and My Keepon. The controls are the same except for the part the [+] and [-] which is changed to [1] and [2] due to some issues of mapping for [+] and [-] on the keyboard.

Unlike the other systems the Telenoid is controlled through a wireless connection which means that no computer is need to be in the room when using the Telenoid, but for the pilot study the computer will be there so the participants can manipulate the head nod, and it can be seen in figure 3.6.



(a) (b)

Figure 3.5: (a) A part of the original software for the Telenoid (b) This part of the UI shows the settings

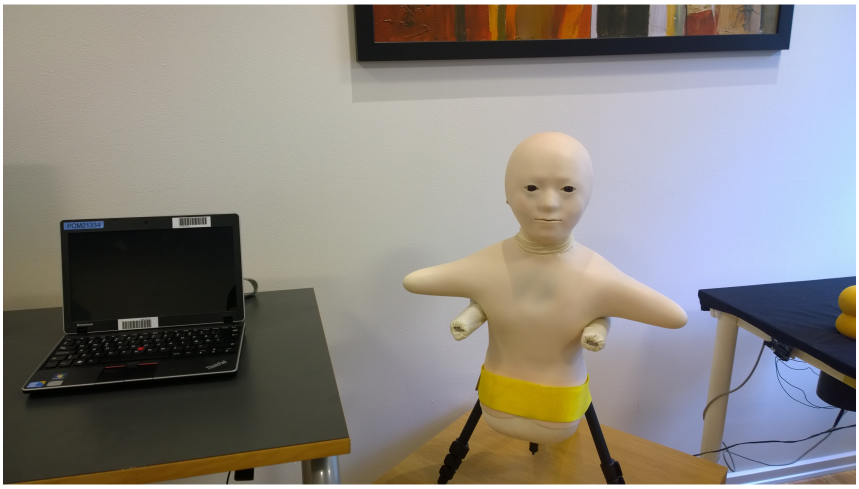


Figure 3.6: The Telenoid's setup for the pilot study with the computer visible.

Chapter 4

Experiment(s)

In this chapter the experiment designs and results will be presented. The structure of each of the experiments will be loosely based on the book *How to Design and Report Experiments* [6] in order to make it an easy and structured read.

The experiments for this project were carried out twice once in December 2013 at Aalborg University and once in June 2014 at SOSU Nord in the Futurelab department. The experiments consists of a pilot study and the main study, and it will in the following sections be presented with the two pilot studies and the two main studies in the same sections, and I will use the place it was held when referring to one of the two studies in an effort to make it clearer.

One of the focuses for this project is to investigate whether cultural appropriate context will have an influence when working with HRI. Participants throughout all of the experiments will therefore be from Denmark.

4.1 Pilot Study - Finding the head nod with the highest perceived sense of listening.

In this part of the study the focus will be on determining the parameters regarding forward head nods as feedback that will influence the perceived sense of listening from different agents during a conversation with a human. For this pilot study My Keepon and the Cat (virtual agent) was used at Aalborg University and My Keepon, Cat, and the Telenoid was used at SOSU Nord.

The participants task in this pilot study is to determine the head nod's *range of motion*, *velocity*, and *repetitions/frequency* for the two agents based on the participants perceived sense that the agent is listening to a conversation and giving feedback through head nodding.

The participants will interact with the agents through a PC interface and thereby manipulate the three parameters. Using the same interface the particpants can tell the agent to perform a head nod, and the participant can then evaluate the head nod physically. To create some context for the head nod 27 small videos of the different combinations of the three parameters will available with the the agent nodding to a sound recording of a person talking for the participant to see. The participant is given all the time needed to make the choice, and asked to write down which combination at the end.

The reason for doing this pilot study is to eliminate the possibility that the head nod performed as feedback by the different agents will affect the result in the experiments where the agents will be using the head nods as the primary feedback for the participants as a head nod can signal a varity of information during a conversation (see section 2.1).

4.1.1 Participants

Aalborg University A total of eight participants participated in this pilot study (five male and three female) with ages from 24 to 30 ($M = 27$, $SD = 2.8$). The participants were informed of the purpose of this pilot study.

SOSU Nord A total of 21 participants participated in this second pilot study (4 male and 17 female) with ages from 18 to 58 ($M = 30.5$, $SD = 12.2$). The participant were all recruited at SOSU Nord and were students and employees. The participants were informed of the purpose of this pilot study.

4.1.2 Apparatus

As mentioned the agents used for this pilot study were My Keepon, Cat, and Telenoid. To be able to use My Keepon it will have to be connected to a PC through an Arduino, and how that it is done can be seen in appendix A. A UI was developed for My Keepon for the pilot study at Aalborg University which was changed for the pilot study at SOSU Nord (see chapter 3 for more information).

4.1.3 Procedure

The participant is seated in front of the agent that he/she will be determining the three parameters for first. The order of the task is altered between each participant so that for participant No. 1 it will be My Keepon first and for the No. 2 the Cat will be first and so on. The participant is then explain that the task is to determine the head nod that gives him/her the greatest perceived sense of listening imagining that the agent is part of an everyday conversation as only the listener and the only feedback is head nodding. The participant is explain how to manipulate the parameters and that there is a short video available for each of the combinations of parameters. The participant is given the time needed to make the choice and the result is written down when the participant is done. In figure 4.2 the setup at Aalborg University is shown and in figure 4.1 the setup at SOSU Nord is shown.

(a) (b)

Figure 4.1: (a) All three agents in the setup at SOSU Nord (b) The laptop with the videos of the different settings

4.1.4 Results

Aalborg University The results from the pilot study at Aalborg University is summarized in table 4.1.

SOSU Nord The results from the pilot study at SOSU Nord is summarized in table 4.2.

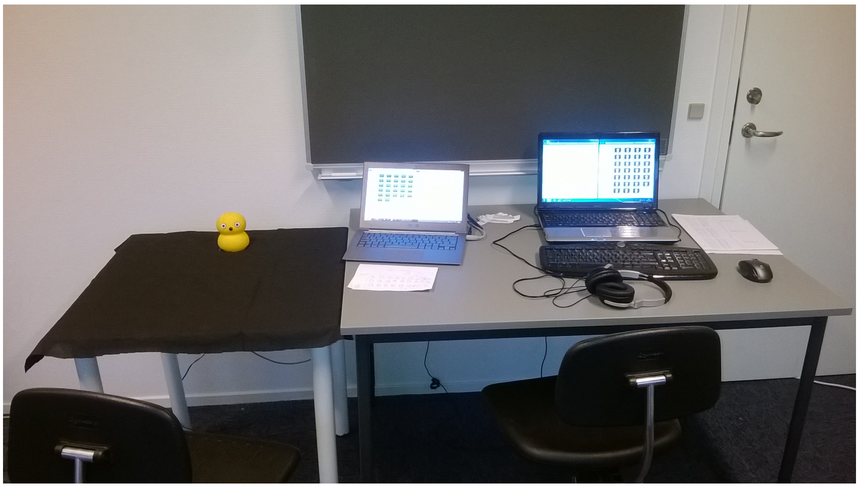


Figure 4.2: The setup of My Keepon and Cat for the pilot at Aalborg University

	My Keepon	**Cat**
	Frequency / Repetitions	
No. 1	50%	50%
No. 2	50%	50%
No. 3	0%	0%
	Speed / Velocity	
Slow	**75%**	**75%**
Medium	12.5%	25%
Fast	12.5%	0%
	Range of motion	
Short	**87.5%**	**62.5%**
Medium	12.5%	25%
Long	0%	12.5%

Table 4.1: Summarized results from pilot study for perceived sense of listening through head nodding

4.1.5 Discussion

Aalborg University The preferred setting is **(1 - 2 repetitions, slow speed, short range of motion)** for both agents. Due to issues with both of the software systems the pilot study at Aalborg University were stopped after eight participants.

	My Keepon	Cat	Telenoid
	Frequency / Repetitions		
No. 1	**66.7%**	**44.9%**	**57.1%**
No. 2	28.6%	33.3%	38.1%
No. 3	4.8%	23.8%	4.8%
	Speed / Velocity		
Slow	**52.4%**	**95.2%**	**57.7%**
Medium	42.9%	0%	23.8%
Fast	4.8%	4.8%	19%
	Range of motion		
Short	**47.6%**	**71.4%**	**61.9%**
Medium	42.9%	19%	38.1%
Long	9.5%	9.5%	0%

Table 4.2: Summarized results from pilot study for perceived sense of listening through head nodding

SOSU Nord The preferred setting is **(1 repetition, slow speed, short range of motion)** for all three agents. During the experiment the participants were told that if they had any problems/issues with selecting the parameters they should just ask, and three general comments and observations made by the participants during the experiments were;

- The virtual agent (Cat) were still a bit too fast at the slowest setting.
- My Keepon nods with its entire body which is different from the two other robots.
- One type of head nod do not seem natural. It depends on the context of what the speaker is speaking about.

These observations has to be kept in mind in regards to the results found during the main study.

4.2 Main Study - Culturally appropriate head nodding from an artificial listener

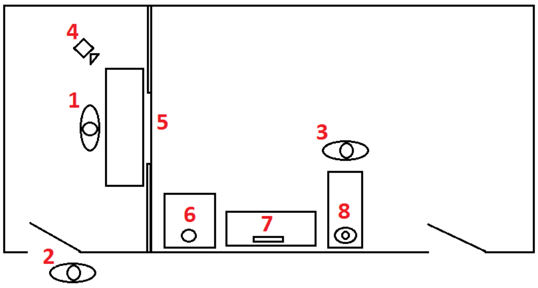
The main study will be divided into several sections determined by the hypothesis being examined. As with the pilot study it was carried out twice at different locations; Aalborg University and SOSU Nord. The two experiments will be going through the same hypotheses and they will therefore be group together based on the hypothesis being investigated.

This is done in order to reduce the need for repeating sections that are the same for each experiment. The difference between the studies will be mentioned, and the reason for the change. Both experiments was conducted using a Wizard of Oz method. Compared to the earlier study by Krosager et al. [23] the facilitator was not presented in the room during the interaction between the participant and the robots. This is to eliminate the added stress of having a person physically presented that is observing you. The participants was told that the robots would be listening, and the facilitator would only be monitoring through video for the purpose of keeping the systems working. After the experiment the participants was debriefed, and told that the facilitator had in fact been listening and been the one generating the head nodding, but nothing was recorded, and the participant is asked if it was okay. The reason of not informing the participants before hand is that the participants would then just see the robot as the media through which they would be talking to the facilitator.

In figure 4.3, page 21, a top-down view of the experiment setup at Aalborg University with pictures of the setting underneath can be seen. The same can be see for the experiment at SOSU Nord in figure 4.4, page 22. The main difference from the two settings is the two-way mirror at Aalborg University for the facilitator to observe the experiment. At SOSU Nord it was done using the camera mounted in the room in the corner marked with 4 in figure 4.4.

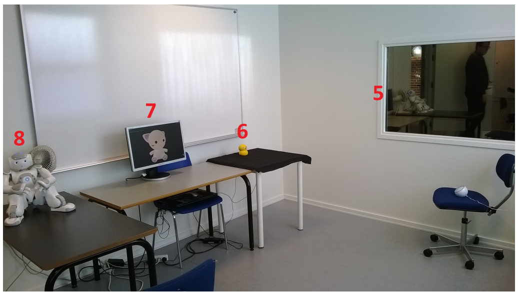
The first part of the study will be the replication of the study by Koda et al. [11] in which it was concluded that culturally appropriate head nodding from an listener agent is important seeing that the interaction sessions were prolonged when using culturally appropriate head nodding. Also the subjective perception of the robot changed depending on the head nodding condition (Japanese, US, no nodding). In the study by Koda et al. [11] physiological measurement (heart rate) was measured during the interaction session with the agent, but this part is removed in this study. The reasoning for not having this measurement is the difficulty of getting accurate measurement relating to the actual interaction with the agent. A particpants heart rate will most likely be effected by just participating in the experiment, and this would therefore require the participants to have the equipment for measuring the heart rate on during a *normal* conversation with a human, and then use this to compare with the interaction with the robot. Seeing that this is very time consuming for the participant it will not be included.

So in this part of the study only the Cat will be used, so the hypotheses based on the



(a)

(b)



(c)

Figure 4.3: **(a)** Top-down view from Aalborg University. 1) Facilitator for systems, 2) facilitator for the participants, 3) participant, 4) camera, 5) two-way mirror, 6 to 8) Robots **(b)** The view from the room with the facilitator controlling the systems. **(c)** The room for the experiment

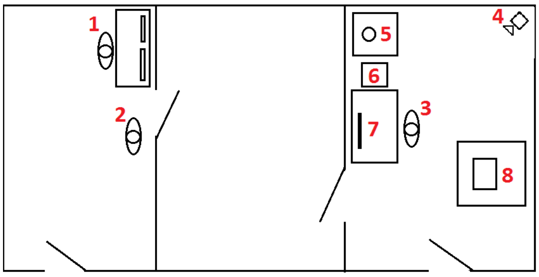
above mentioned study will be the following

> ***H1:*** Culturally appropriate head nodding behaviour from a virtual agent will elicit longer speech duration compared to a virtual agent showing culturally inappropriate head nodding or no head nodding.

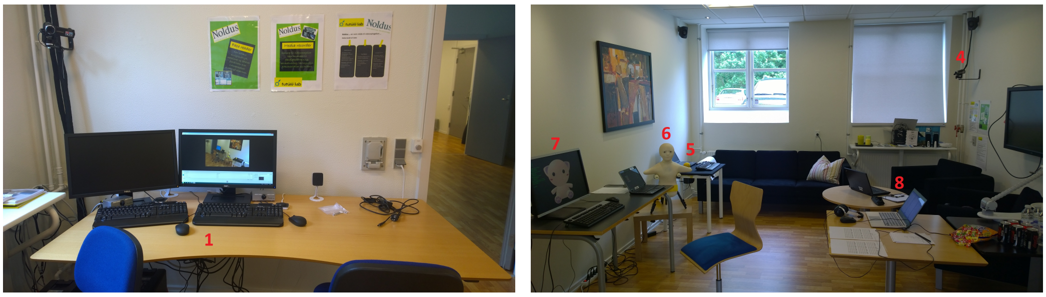
> ***H2:*** A virtual agent using culturally appropriate head nodding will be perceived as more intelligent compared to a virtual agent showing culturally inappropriate head nodding or no head nodding.

4.2.1 Design

The experimental design of the experiment varies between the one at Aalborg University and SOSU Nord.



(a)



(b) (c)

Figure 4.4: (a) Top-down view from SOSU Nord. 1) Facilitator for systems, 2) facilitator for the participants, 3) participant, 4) camera, 5 to 7) Robots, 8) Laptop with questionnaires **(b)** The view from the room with the facilitator controlling the systems. **(c)** The room for the experiment

Aalborg University In this experiment a between-subject design was used with the culturally appropriate head nodding as the independent variable; DK, US, and No nodding(NN), and the speech duration measured in seconds and the perceived intelligence rated on a 5-point Likert scale as the dependent variables.

SOSU Nord In this experiment a within-subject design with the same independent and dependent variables as at Aalborg University.

The reason for the difference in experiment design comes down to the experience from the first experiment at Aalborg University where the participant would only try one type of head nodding types from the Cat since the experiment design was design for the participant to talk to three agent (Cat, My Keepon, and Nao) and the robots would be using the same head nodding condition (DK, US, and no nodding). This will make the comparison of perception in intelligence based on culture less reliable seeing that the participant would not encounter all types of head nodding conditions.

4.2.2 Participants

All participants in the two experiment were naive, and was not informed of the real purpose of the experiment before the debriefing.

Aalborg University A total of 18 participants participated in this study (nine male and nine female) with ages from 20 to 25 ($M = 22$, $SD = 1.5$). All participants were students from Aalborg University.

SOSU Nord A total of six participants participated in this study (one male and five female) with ages from 21 to 52 ($M = 37.5$, $SD = 10.1$). The participants were recruited at SOSU Nord.

4.2.3 Apparatus

A virtual agent developed by Koda et al. [11] was used as the listener agent performing the head nodding. The head nod used for the two experiments was determined through the pilot studies (see section 4.1). The Cat is also blinking during the experiment which can't be turned off.

For some added information about the participants personality an online survey is filled out at the start of the experiment. For this the EPQR-A questionnaire [7] which is a reliable abbreviated version of the EPQ (Eysenck Personality Questionnaire) [5] with focus on extroversion versus introversion. The purpose of including this questionnaire is based on the assumption that an extrovert might talk longer and vice versa. It can therefore be helpful in explaining any potential outlier in the collected data.

Lastly the Godspeed questionnaire [2] is used the perceived intelligence of the agent after each talk session. It consists of five questions rated on a 5-point Likert scale.

All of the used questionnaires can be found in Danish in the appendix B.

From the research on head nodding from different cultures (see section 2.2) the Danish frequency is determined to 5.8 seconds and 8.5 seconds between each head nod, so an average at around 7 seconds will be the target (eight to nine head nods per minute). For the US head nodding the frequency is 22.5 seconds (two to three head nods per minute).

4.2.4 Procedure

The facilitator for the participants greets the participant upon arrival and starts the experiment by explaining what is going to happen throughout the entire experiment. The participant is explained that they will be talking to an artificial listener that might responded to what the participant is talking about. The participant is explained that the agent will at no point take the role as speaker, it will only act as the listener. The participant is told that the study is about social robotics, but no more information is given at the start of the experiment. After explaining to the participant that they will be talking to the agent they are asked if they are still interested in participating. If the participant agree they start the experiment by filling out an online survey regarding general information (age and gender) and the personality questionnaire (EPQR-A). After that the participant is presented with a list of topics to talk about as inspiration. This is all done together with the facilitator. For the study at Aalborg University a consent form is signed allowing video recording. This is left out at SOSU Nord as the video camera will not be recording. When the participant has chosen a topic to talk about they are seated in front of the Cat, and told that they should speak for as long as possible about the topic and only the one topic. The facilitator then leaves the room, and the participant starts talking to the agent. Once the participant feels that they are done talking about the topic, they put a hand in the air to signal to the facilitator that they are done and the facilitator enters the room again. The participant and the facilitator talks out loud about the general experience, and the facilitator behind the scene, controlling the system takes notes. The participant is then asked to fill out the Godspeed questionnaire about the perceived intelligence of the agent based on the just ended talk session.

For the study at Aalborg University the participant is randomly assigned a head nodding type (DK, US, or No Nodding), and will only interact with this case. At SOSU Nord the participants will go through all three types of head nodding with three different topics. In order to avoid any carry-over bias from the three conditions they are counterbalance between the six participants.

After the experiment is done the facilitator for the participants will debrief the participant, and explain that the Cat was controlled by a human facilitator (they never meant this facilitator) and explained that the purpose of the study is to investigate culturally appropriate head nodding from an listener agent. The participant is offered coffee, water and/or soda as a thanks for participating.

4.2.5 Results

The dependent variable speech duration is always shown in seconds, and it will therefore not be written after each time the variable is mentioned.

Aalborg University A one-way independent ANOVA was performed on the data from Aalborg University (DK, US, and no nodding). No significant difference was found regarding longer speech duration between the nodding types DK (N = 7, M = 81, SD = 36), US (N = 5, M = 183, SD = 129), and no nodding (N = 6, M = 93, SD = 93). After removing potential outliers (N = 15) a significant differences was found using a one-way independent ANOVA when comparing the speech duration between the nodding types DK (N = 6, M = 57, SD = 34), US (N = 4, M = 133, SD = 71), and no nodding (N = 5, M = 57, SD = 34) at $F(2, 14) = 4.013$, $p < 0.05$. The participants talk significantly longer under the US condition determined by a Fisher's LSD post-hoc test that showed a significant difference between US and the other two conditions (DK and NN) at $p < 0.05$.

The extroversion/introversion score can not give any interesting results as most of the participant scored high on the survey making a comparison pointless.

For the perceived intelligence for the study at Aalborg University the results will not be used as the participants did not go through all conditions regarding the cultural appropriate head nodding for the Cat.

SOSU Nord A one-way repeated-measures ANOVA was performed on the data from SOSU Nord (DK, US, and no nodding). No significant difference was found regarding longer speech duration between the nodding types DK (N = 6, M = 67, SD = 39), US (N = 6, M = 83, SD = 27), and no noddding (N = 6, M = 81, SD = 43). Removing potential outliers made no difference. The one-way repeated-measures ANOVA resulted in $F(2, 10) = 0.408$, $p = 0.68$.

As with the participants of the experiment at Aalborg Universty the participants at SOSU Nord scores high in the survey regarding extroversion/introversion, so no further investigation is performed on this data.

For the perceived intelligence score a summary is shown in figure 4.5 which shows that

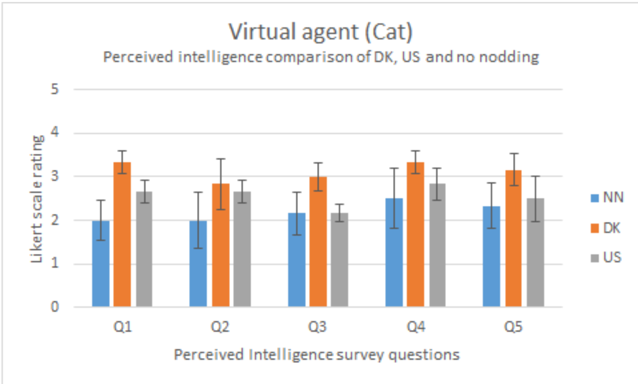


Figure 4.5: A summary of the score from the survey regarding perceived intelligence.

there might be a difference in perceived intelligence between the difference types of head nodding when look at question #1 (Q1) and question #3 (Q3). A Friedmann's ANOVA is calculated for Q1 and Q3. A significant difference was found for Q1, $X^2(2) = 7.1$, $p < 0.05$. A post-hoc analysis shows that the only significant difference is between DK and no nodding conditions, $p < 0.05$. No significant differences was found in the Q3 condition.

4.2.6 Discussion

Hypothesis H1 can be rejected as it was found that culturally appropriate head nodding did not elicit longer speech duration in either of the two studies. This finding contradicts the assumptions made by Koda et al. [11] where the culturally appropriate head nodding did elicit longer speech duration. But the result is the same as Krosager et al. [23] found when using the Nao, a physical humanoid robot. The interesting part is that the participant in the study from Aalborg University talked longer with the Cat under the US condition. This indicates that having feedback might elicit longer speech duration, but having too many head nods per minute might negatively influence the results as previous studies indicate [17]. This trend was also visible in the results from SOSU Nord with the highest average talking time was found under the US condition, but without being significant.

The second hypothesis H2 stating that culturally appropriate head nodding will influence the perception of the Cat is also rejected as only one part of the Godspeed questionnaire was found to be significant. Question #1 covering how competent the participants rated the Cat showed that the Danish condition was rated highest. This shows that even though the participants did not have longer talking sessions with the cat the higher frequency of head nodding added to the feeling of talking to a competent listener. If looking at the data in graph 4.5 there is a tendency that the Danish condition scores higher regarding perceived intelligence of the Cat. This could indicate that using time as the dependent variable for measuring culturally appropriate use of head nodding might not be the best solution. How can we be sure that time will reflect the perception of the agent in question?

Seeing that the results contradicts the findings by Koda et al. [11] it would be interesting to investigate whether or not the embodiment of the listener agent will influence the speech duration. This is speculated as the Cat agent might have a higher appeal toward Japanese participants with the cartoon- (anime) like appearance. It might be more relatable for Japanese participants to talk this kind of agent than for Danish participants. This speculation leads to the next part of the study - will the embodiment of the agent influence the speech duration?

4.3 Main Study - Perception of listener agent differs based embodiments

For this second part of the project the embodiment of the listeners agents will be evaluated. As mentioned in section 3 four different embodiments for a listener agent was included in the experimental design. As mentioned in the previous sections the experiment was carried out twice at different locations; Aalborg University and SOSU Nord. There are a few differences between the two experiment, which will be mentioned as it is brought up in the following sections just like the previous sections. The main differences between the two studies is that the Nao was included at Aalborg University and the Telenoid at SOSU Nord.

Based on the findings, from the previous sections, that culturally appropriate head nodding will not elicit longer speech duration compared to un-specific and no head nodding, the focus will be om whether the embodiment might have an influence when working with Danish participants. Also previous findings by Krosager et al. [23] indicates that the mere physical present of the agent will influence the speech duration from the speaker compared to a virtually present agent. The hypotheses will therefore be

> ***H3:*** The embodiment of the listener agent will influence the speech duration from the speaker.

H4: The embodiment of the listener agent will influence the perceived intelligence rated by the speaker.

H5: A virtual listener agent will elicit longer speech duration compared to a physically present agent.

4.3.1 Design

The experiment varies from each other between the two locations.

Aalborg University In this experiment a within-subject design was used for the robot embodiment as the independent variable (My Keepon, Nao, and Cat), and again having speech duration measured in seconds and the perceived intelligence measured on a 5-point Likert scale as the dependent variables.

SOSU Nord In this experiment a between-subject design was used for the robot embodiment as the independent variable (My Keepon, Telenoid, and Cat), and the dependent variables again the same as at Aalborg University.

The same issue regarding the design choice for the experiment at Aalborg University will influence the results gathered from the perception of the robot as it will maybe be based on a combination of embodiment and head nodding type, and this is not being controlled efficiently enough in the experiment at Aalborg University.

4.3.2 Participants

All participants in the two experiment were naive, and was not informed of the real purpose of the experiment before the debriefing

Aalborg University A total of 18 participants participated in this study (nine male and nine female) with ages from 20 to 25 ($M = 22$, $SD = 1.5$). All participants were students from Aalborg University.

SOSU Nord A total of 18 participants participated in this study (3 male and 15 female) with ages from 18 to 52 ($M = 37$, $SD = 10.1$). The participants were recruited at SOSU

Nord.

4.3.3 Apparatus

The questionnaires, virtual agent, and research is the same as in the previous section (see section 4.2.3). For the experiment at Aalborg University the Nao robot with software for controlling head nodding and the face tracking system active is included. For the experiment at SOSU Nord the Telenoid robot with software for controlling head nodding is included.

4.3.4 Procedure

The procedure is also the same explained in the previous section (see section 4.2.4) with the main difference highlight for the two locations:

Aalborg University In this case the participant will pick three topics, and talk one time with each agent (My Keepon, Nao, and Cat). The three agents will be nodding with the same head nodding type (DK, US, No nodding). The participant will fill out the Godspeed questionnaire after each talk session. The order the participants meet the agent is randomize as well as the head nodding type.

SOSO Nord Each participant will again pick three topics, but only talk to one type of agent (My Keepon, Telenoid, and Cat). The agent will be head nodding according to the three conditions (DK, US, No nodding). The head nodding condition is counterbalanced for each robot between participants to eliminate any carry-out effects in the end results. The participants will fill out the Godspeed questionnaire after each talk session.

4.3.5 Results

The dependent variable speech duration is always shown in seconds, and it will therefore not be written after each time the variable is mentioned.

Aalborg University A one-way repeated measures ANOVA was performed on the data from Aalborg University (My Keepon, Nao, Cat). No significant difference was found regarding longer speech duration based on embodiment; conditions My Keepon ($N = 18$,

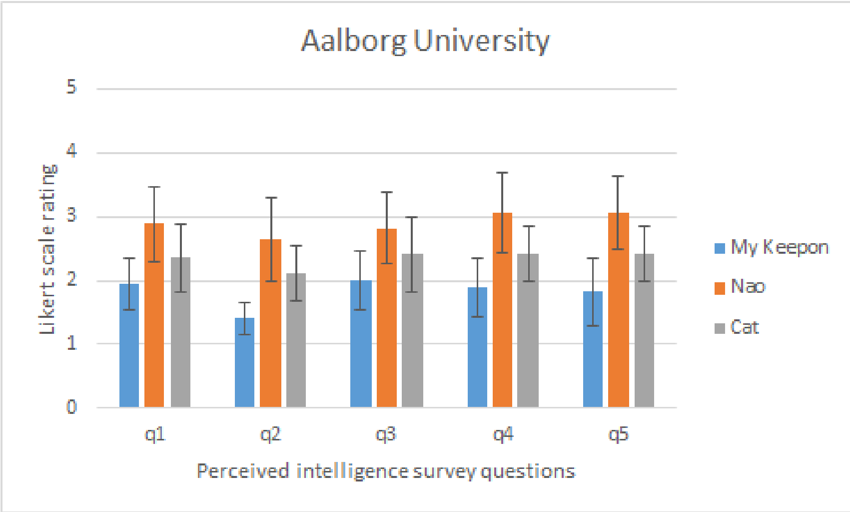


Figure 4.6: A summary of the score from the survey regarding perceived intelligence from Aalborg University.

M = 124, SD = 105), Nao (N = 18, M = 97, SD = 60), Cat (N = 18, M = 113, SD = 94) at $F(2, 34) = 2.2$, $p = 0.13$. Removing potential outliers also showed no significant result.

For the perceived intelligence score a summary is shown in figure 4.6 which shows a tendency for the Nao to score highest and My Keepon scoring lowest. A Friedman's ANOVA is used for each question from the survey to see if any significant differences are found. Significant differences is found for all five questions; Q1: $X^2(2) = 7.3$, $p < 0.05$, Q2: $X^2(2) = 13.5$, $p < 0.05$, Q3: $X^2(2) = 9.8$, $p < 0.05$, Q4: $X^2(2) = 10.9$, $p < 0.05$, and Q5: $X^2(2) = 14.4$, $p < 0.05$. The Friedman's ANOVA is used as the post-hoc test to determine between which embodiments the significant is found. The Nao scores significantly higher than My Keepon in all quesitons, $p < 0.05$, and the Cat scores significantly higher than My Keepon in Q2 and Q5, $p < 0.05$. The Nao scores significantly higher than the Cat in Q5, $p < 0.05$.

A dependent t-test us used to determine if a virtual agent (N = 18, M = 116, SD = 96.9) elicit longer speech duration compared to a physical agent (N = 18, M = 108, SD = 77.5). There is no significant differences between the two groups, $t(17) = 1.02$, $p = 0.32$.

SOSU Nord A one-way independent ANOVA was performed on the data from SOSU Nord (My Keepon, Telenoid, Cat). No significant difference was found regarding longer

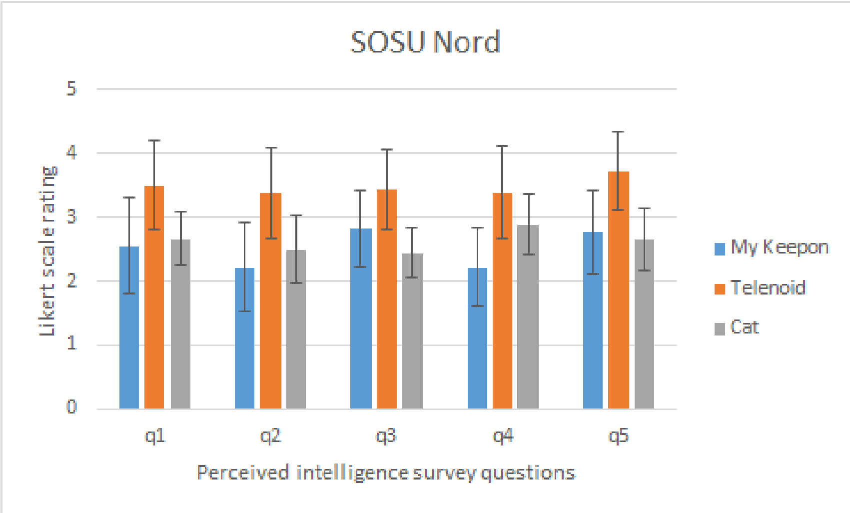


Figure 4.7: A summary of the score from the survey regarding perceived intelligence from SOSU Nord.

speech duration based on the embodiment; conditions My Keepon (N = 18, M = 108, SD = 45), Telenoid (N = 18, M = 95, SD = 32), Cat (N = 18, M = 78, SD = 35), F(2, 53) = 2.89, p = 0.07. Removing potential outliers also showed no significant result.

For the perceived intelligence score a summary is shown in figure 4.7 which shows a tendency for the Telenoid to score highest and My Keepon or the Cat scoring lowest. A Kruskal-Wallis test is used for each question from the survey to see if any significant differences are found. There is a significant differences in all questions, except for Q1: Q2 $X^2(2) = 6.5$, $p < 0.05$, Q3: $X^2(2) = 7.3$, $p < 0.05$, Q4: $X^2(2) = 7.1$, $p < 0.05$, and Q5: $X^2(2) = 7.9$, $p < 0.05$. A Mann-Whitney U test is used for post-hoc testing to determine which embodiments significantly differs. The Telenoid scores significantly higher in Q2, Q4, and Q5 than My Keepon (Q2: U = 90.5, $p < 0.05$, Q4: U = 86.5, $p < 0.05$, Q5: U = 97.0, $p < 0.05$), and significantly higher in Q3 and Q5 compared to the Cat (Q3: U = 79.0, $p < 0.05$ and Q5: U = 81.0, $p < 0.05$).

An independent t-test is used to determine if a virtual agent (N = 6, M = 77.7, SE = 11.0) elicit longer speech duration compared to a physical agent (N = 12, M = 101.5, SE = 9.8). There is no significant differences between the two groups, t(16) = 1.5, p = 0.15.

4.3.6 Discussion

Hypothesis H3 is rejected as no significant differences in speech duration was found in the experiments at Aalborg University and SOSU Nord using different robots as the listener agent, but a significant difference in embodiment is found when evaluating the different embodiment using a subjective approach so hypothesis H4 is retained. At both locations the humanoid robot is perceived as more intelligent which might be related to the expectations of a humanoid robot. The human features might add to the perception that the robot is intelligence even though it shows no other behavior compared to My Keepon and the Cat, but as the behavior is controlled by a human based on human behavior it will probably be better suited for humanoid robots. This might be interesting to investigate using the research from the robot seal (PARO) which is created to show animal behavior, but still interact with humans. Will this kind of behavior be more appropriate for My Keepon and the Cat agent? Also the demographics might influence the perception again based on the difference in culture and the expectations from robots. In Denmark social robotics for the general public is still not something that is widely used yet, and the prior understanding of different robot embodiments might be effected by this.

In the paper by Krosager et al. [12] it is found that the people interacting with a virtual version of the listener agent will talk significantly longer compared to the people talking to the physically present agent. This was the basis for hypothesis H5; a virtual agent will elicit longer speech duration compared to a physically present agent. This hypothesis is reject as no significant difference was found when comparing the virtual agent (Cat) with the physically present agents (My Keepon, Nao, and Telenoid). The two groups (virtual versus physical) was insignificantly different at both locations; Aalborg University and SOSU Nord. One speculation is that the mere presence of a human facilitator in the room might influence the speech duration, and this confounding factor was remove in this project.

Since it was shown that the perception of the agent will change based on the embodiment evaluated with subjective measures and that having feedback signals (the head nodding) will elicit longer speech duration compared to having no feedback when talking to a virtual agent. Therefore a look at the results from the two experiments with both independent variables (head nodding and robots) will be the focus for the last evaluation of the two experiments.

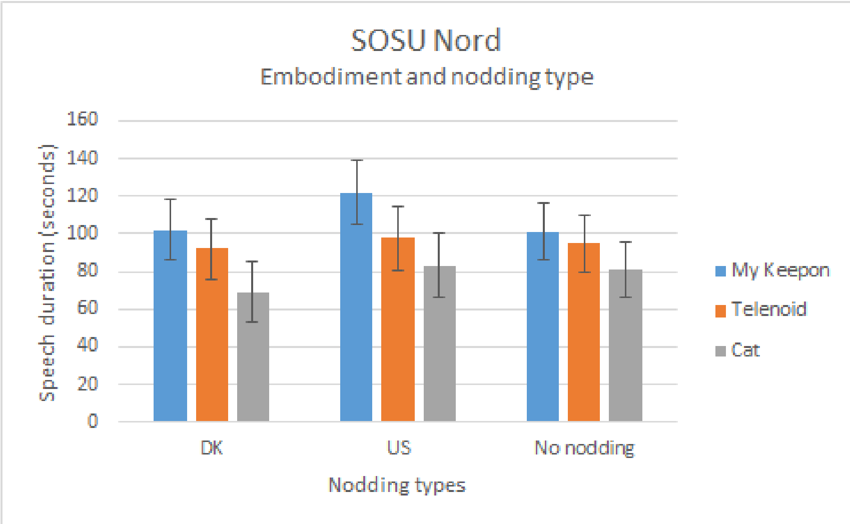


Figure 4.8: A summary of the speech duration from the different agents divided by nodding type.

4.4 Evaluation - The effect of head nodding from different embodiments during conversation

In this sections the data from SOSU Nord is evaluated with focus on the effect of having head nodding (DK, US, and no nodding) from the different agents still using the measurement time and subjective rating. Based on the previous findings the assumption for this section is that the humanoid robot will be perceived as the most intelligent under the head nodding condition US compared to the other two embodiments and head nodding conditions.

In figure 4.8 a summary of data from the experiment from SOSU Nord with all agents divided by the nodding conditions can be seen. A mixed factorial ANOVA was used on the data from SOSU Nord to see if there where any significant differences when combining embodiment and head nodding types, and no significance was found $F(2, 13) = 0.125$, $p > 0.05$.

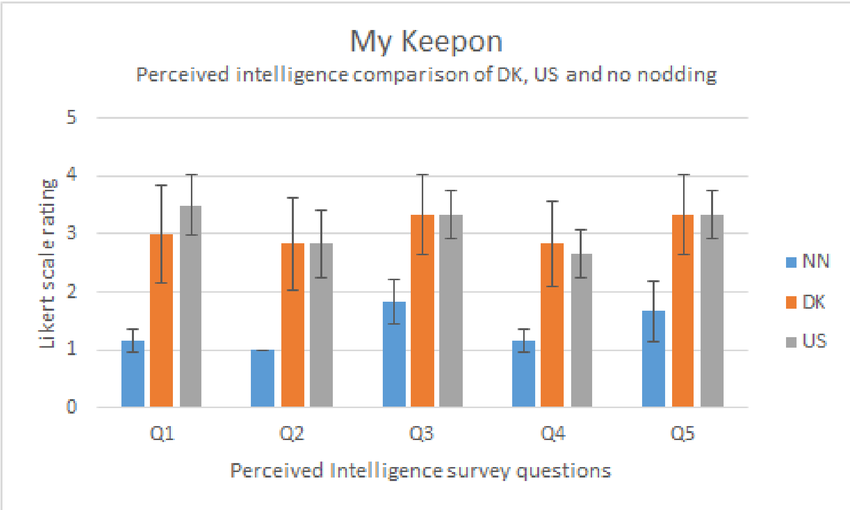
The perceived intelligence of My Keepon and the Telenoid is summarized in figure 4.9, an a clear tendency of the no nodding condition scoring lowest can be seen. A Friedman's ANOVA is used to determine if there are any significant differences in each question. For My Keepon Q1, Q2 and Q4 showed significant differences, (Q1: $X^2(2) = 6.9$, $p = 0.03$;

Q2: $X^2(2) = 7.4$, $p = 0.03$; Q3: $X^2(2) = 7$, $p = 0.03$), so a Wilcoxon Signed-Rank test was used for post-hoc evaulation which showed that for Q1, Q2, and Q3 the US condition significantly differs from the no nodding condition, $p < 0.05$. A Friedman's ANOVA is also used for the Telenoid case and in this case there is significant differences in all questions, $p < 0.05$. The Wilcoxon Signed-Rank test used for the post-hoc evaluation showed that the DK and US conditions significantly differs from the no nodding condition, $p < 0.05$.

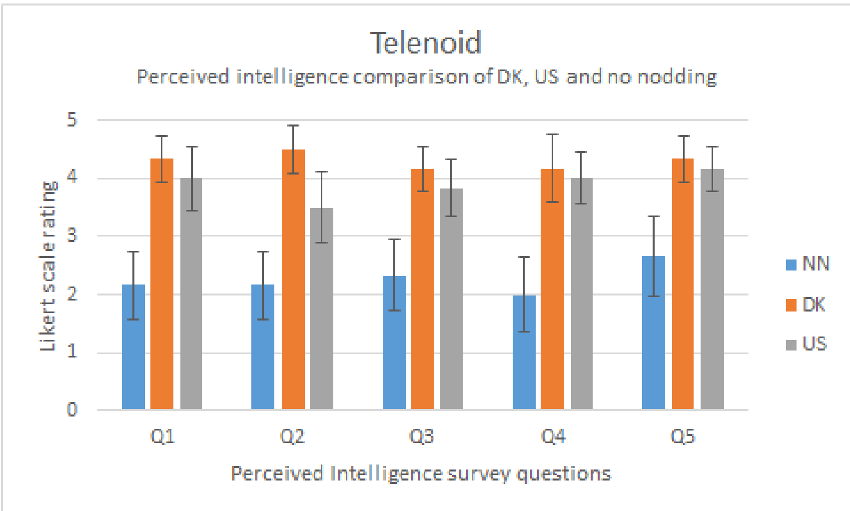
General observation and comments As mentioned in section 4.2.4 about procedure the facilitator and participants would talk out loud during the questionnaires and after the experiment, and from this some general observations has been made:

- **DK condition:**
  - A general comment is that the listener agent is nodding too much.
  - The agent comes of as too understanding, because of the higher frequency in head nodding.
  - The higher frequency in head nodding adds to the risk of the timing being wrong.
- **US condition:**
  - This was mentioned as the preferred case from more of the participants.
- **No nodding condition**
  - Referred to as *talking to a wall*.

My Keepon was called a *toy/plaything* and some the participants found it odd speaking to a cat. The Cat was also getting complains that the blinking was weird and did not seem natural.



(a)



(b)

Figure 4.9: (a) A summary of the score from the survey regarding perceived intelligence at SOSU Nord for My Keepon **(b)** A summary of the score from the survey regarding perceived intelligence at SOSU Nord for the Telenoid

EXPERIMENT(S)

Chapter 5

Discussion

A replication of the study [11] was carried out with the change to the original study being the use of Danish participants instead of Japanese participants. In the original study it was concluded that the use of culturally appropriate head nodding would elicit longer speech duration from the human speaker talking to a virtual agent. The results from the replication contradicts these findings seeing that culturally appropriate head nodding did not elicit longer speech duration compared to unspecific head nodding or no head nodding. The culturally appropriate head nodding did not significantly get higher score on Likert scale rating regarding the perceived intelligence which also contradicts the findings from the original study, but there were a tendency for the culturally appropriate head nodding being rated highest.

Different embodiments was included to investigate whether the embodiment of the agent would elicit longer speech duration and this was also rejected as no significant difference was found, but through the subjective measure of perceived intelligence it was found that humanoid robots scores significantly higher compared to other robots when using human behavior for the robots. The effect of having the robot physically present in the room during the interaction was compared to the virtual agent, and no significance in speech duration was found.

Lastly the combination of embodiment and cultural head nodding was investigate, and again no significant difference was found when measuring speech duration in any of the cases, but the perceived intelligence scored significantly higher for both My Keepon and the Telenoid in all most all of the questions from the Godspeed questionnaire when head nodding was included compared to no head nodding. This indicates that backchannel head nodding from a listener agent is important, but the use of culturally appropriate head nodding is not necessarily the way to go. This is also seen from the general observation where the DK condition is referred to as been *too much*, and the US condition mentioned as the preferred condition, but the no nodding condition is as expected not appreciated as it for most of the participants is described as talking to a wall. The comments regarding My Keepon being called a toy seems to fit with the humanoid robots scoring higher in the perceived intelligence rating.

These findings indicates that adding cultural background to an agent might not be suitable as the participant might not see the agent as coming from the same culture, and the cultural background can therefore be interpret as being wrong.

After analyzing the data the two experiments at Aalborg University and SOSU Nord it seems clear that time is not an appropriate way of measuring the influence of backchannel head nodding. There are so many things that can effect the time (e.g. the topics), and it is therefore hard to know what triggers a longer speech duration. Based on the results from the Godspeed questionnaire it would seem more appropriate to use more subjective measurement methods (e.g. interviews) to get a better understanding of which behavior we as humans expects from robot interaction in the early state of investigating robot behavior. After getting more control and knowledge over what constitute appropriate robot behavior it might be interesting to look at measurements such as time again.

Another issue with this approach is the use of only one predetermined head nod. This will quickly give the agents a mechanical feel as they are just repeating the same motion every time. It would therefore be interesting to have some variations in the head nod generated by the agents, and see if that will effect the perception of them. It would also be interesting to add verbal backchannels (Umm, okay, etc.) as this is very common in human conversation.

The timing during both experiments was one of the bigger issues with backchannel head nodding as this is often done subconsciously, and the timing would therefore be a few seconds off, and this would in some cases interrupt the speakers concentration. It would therefore be interesting to use the Telenoids head-mounted setup, that will move the Telenoids head just like the facilitators and see if the timing would be more accurate in this way.

One of the bigger challenges with implementing human behavior in robot interaction comes down to the hope of measuring the effect. In order to measure the effect, the effect needs to be isolated in the experiment. In this case having only non-verbal backchannel head nodding in a experimental setup will create a *fake* scenario, which might effect the participant. The participants often tries to push themselves to talk more because they are part of an experiment, and often want to help the people conducting the experiment by giving them *better* data. There is also the novelty effect associated with robots within HRI as most of the participants has never worked with robots before, and this can also have an influence on the data collected. Also the general interest for this kind of technology might influence the willingness during the experiment. At SOSU Nord the participants was older, and generally interested in social robotics, because they could see the use for it in their field of work. The same interest was not seen with the participants at Aalborg University, and it might therefore be important to have participants from areas with interest in social robotics.

5.1 Guidelines for an autonomous system

With the knowledge from the two experiments some general guidelines for an autonomous system able of generating appropriate head nodding can be made but answer the last part from the introduction; is it possible to develop an autonomous system capable of generating real-time feedback signals during a face-to-face conversation that will help to maintain it in longer time? The part about maintaining the conversation for longer time should be change to be a rating of the perception of the system using e.g. the Godspeed questionnaire instead of time. In order to developed a system capable of performing as the facilitator in the experiment the use of sound would only be needed, as the placement of most head nods was when the speaker would have a short break in the speech-flow, and combine this with the frequency condition US.

DISCUSSION

Chapter 6

Conclusion

In this project a series of experiments regarding human-robot interaction has been report. It was found that, contrary to previous research in other cultures, culturally appropriate head nodding does not have an effect on prolonging speech duration from a human speaker, but the use of backchannel head nodding will in some cases add to prolonged interaction time. It was also found that humanoid robots are perceived as more intelligent based on a conversation with the robot as listener, compared to other types of agents. It was concluded that using time to measure a participants perception of an agent is not appropriate, but subjective measures such as surveys might be more appropriate.

CONCLUSION

Appendix A

Hack My Keepon

The following guide to hack My Keepon is developed using the online resource *BeatBots/My Keepon* [4], and it will go through the equipment and software needed combined with a systematic guide of the setup. For more details on the hardware inside My Keepon see [20]

A.1 Equipment

In this section the equipment used for hacking My Keepon will be listed. The numbering in the list corresponds to the numbering in figure A.1.

1. My Keepon[1]
2. Power supply for My Keepon (12V and 1.5A) or 8 AA batteries
3. Arduino board (In this guide a Arduino Uno is used)
4. USB-A / USB-B (Printer cable)
5. Philips screwdriver
6. Wire-cutter
7. Four wires or a jumper wire 4-pin
8. Connectors for soldering to the PCB board inside My Keepon
9. Soldering kit

[1] http://tinyurl.com/oo7yfd8

Figure A.1: An overview of the equipment needed in order to hack My Keepon

A.2 Software

For communication with My Keepon using the Arduino board the Arduino IDE[2] has to be installed on the PC that is connected to the Arduino board. Based on what the goal with hacking My Keepon is it might be interesting to develop some UI for interfacing My Keepon from the PC, and this can be done using a variety of software tools. In this guide Processing[3] is used for this purpose. Below is the two libraries used for the Arduino and Processing:

- **Arduino:** My Keepon Arduino Controller[4]
- **Processing:** Serial Library

A.3 Guide

In the following sections it will be explained how My Keepon is hacked. After finishing hacking My Keepon there will be a quick guide for controlling My Keepon through a Processing Sketch. Lastly the software used for this project will be presented.

[2]Version 1.0.5-r2
[3]Version 2.1.1
[4]https://github.com/BeatBots/My Keepon/blob/master/My Keepon.ino

A.3.1 Open My Keepon

Take out the four screws at the back of My Keepon in order to open it up. Unplug the bigger black and blue cable from top part of My Keepon from the bottom part as shown in picture A.2.

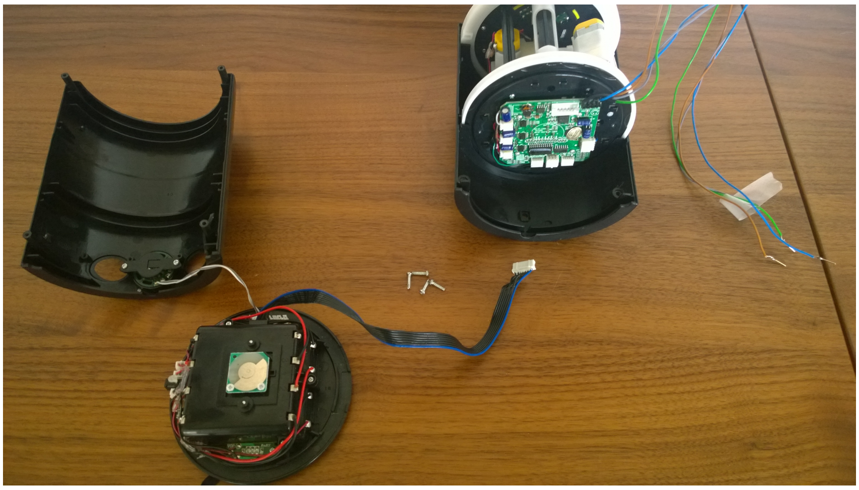


Figure A.2: My Keepon opened and the black/blue cable unplugged

A.3.2 Unscrew and unplug the PCB board

Now unplug all the cables from the PCB board. I will recommend take you use your mobile phone's camera to take pictures of where you have to reconnect the different cables. After that unscrew the PCB board so it looks like the picture A.3. The interesting part of the board is on the lower left corner. Below the smiley on the board there are four holes marked **V**, **CI**, **DA**, and **G**.

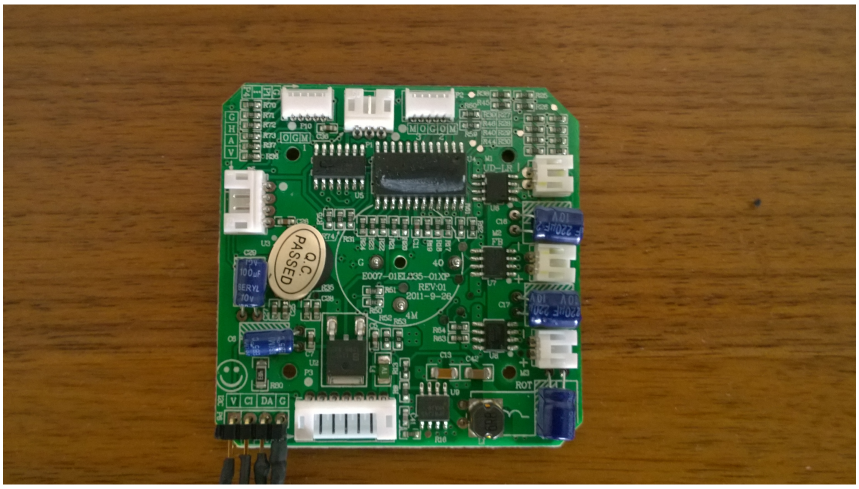


Figure A.3: The PCB board from My Keepon

A.3.3 Start soldering

Now that the PCB board is out of My Keepon it is time to solder on the connector to the board.

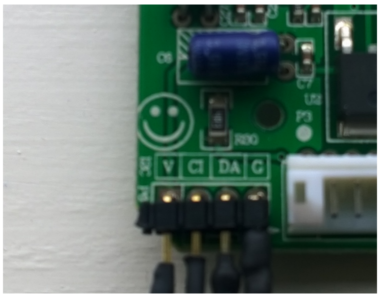


Figure A.4: Close up of the PCB board with the connector solder on

As mentioned in section A.1 it can be done with either four wires or a jumper wire (4-pin), but both solutions requires a connector on the PCB board. The connector solder on to the board can be seen in figure A.4.

As shown in the figure a connector with four wires soldered onto it was used. To make it easier later on the four wires are in different colors, and the chosen has a 90° bend, so it follows the direction as the other wires plugged into the PCB board.

After soldering on the wires My Keepon is ready to be closed up again. First start by reconnecting the unplugged cables and screw the PCB board back onto My Keepon.

Last part of closing up My Keepon is to choose where the wires will exit My Keepon.

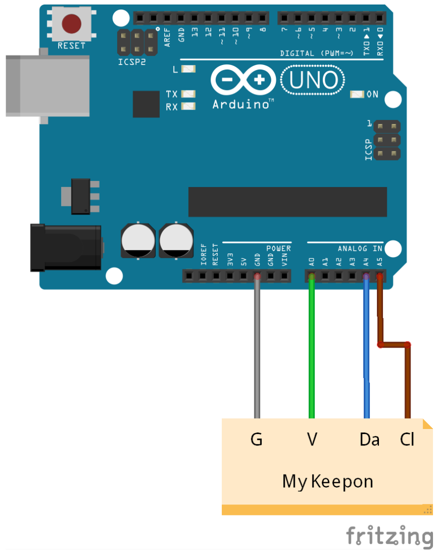


Figure A.5: Illustration of the connection between My Keepon and the Arduino Uno

A.3.4 Connecting My Keepon to the Arduino

As just mentioned you have to decide where the wires you just soldered onto the PCB board should exit My Keepon whether you are using four wires or the jumper wire. I have remove the buttons found on the front of My Keepon since I am not going to use them for my setup. This gives room for the wires to exit through the holes left by the buttons. If this is not the solution, you prefer you could opt for drill a hole in the case of My Keepon and make a neater finish.

With My Keepon fully assembled again and the wires connected to the PCB board, you are now ready to connect My Keepon to the Arduino.

In figure A.5, you can see the connection between My Keepon and the Arduino Uno. The colors of the wires corresponds to the color of the wires I used, but you can use whatever, color combination that, floats your boat.

With My Keepon still power off connect the Arduino to your pc and upload the code *My Keepon Arduino Controller*. After the upload, open the built-in serial monitor in the Arduino IDE and set the baud rate to 115200, and then power on My Keepon. You should now receive the message "“My Keepon detected"”.

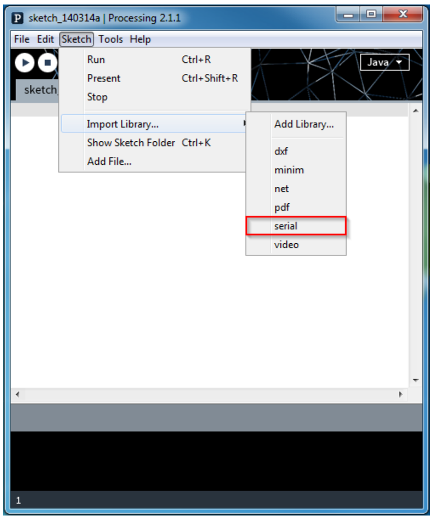


Figure A.6: Import the Serial Library

You can now communicate with My Keepon through the serial monitor with the pre-defined commands, shown below in table 1, e.g. "“SOUND PLAY 20;"” **It is important to remember the closing semicolon.**

A.3.5 Setup Processing

After all the hardware is setup it is time to get the software up and running. I personally prefer to use Processing because of the ease of use and large online community, where help is never far away. However, if you prefer any other IDE it should be just as easy to set it up, since all you need is a serial connection to the Arduino Board.

As I mentioned under *software* the main parts you need is the Processing IDE and the Serial library for communicating with the Arduino. You could of course add a Bluetooth or Ethernet Shield to the Arduino if you need it to be wireless, but that beyond this guide.

The first thing you need to do is download the Processing IDE found at https://processing.org/download/, and for now you need the 64-bit version in order for the Serial library to work. (This might be different based on when you read this guide.)

The first step is to import the Serial library and this is where Processing is easy to use. All you have to do is *Sketch -> Import Library -> serial*, and after that, the Processing IDE takes care of the rest. It can be seen in figure A.6.

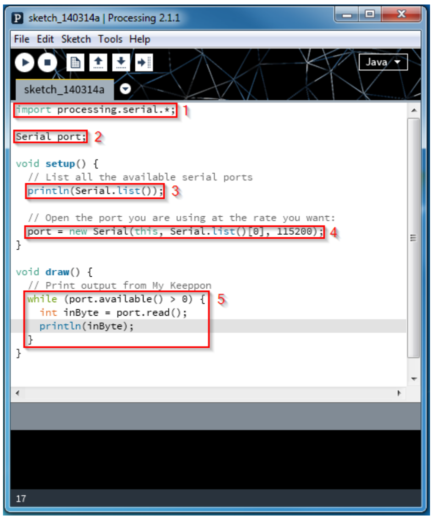


Figure A.7: All the code needed for Processing to read the output from the Arduino

In figure A.7 all the code you need in order to setup the serial connection to the Arduino can be seen. I have marked the code with numbers for easy referencing below.

1. This line of code is added as you import the serial library as mentioned above.
2. Create a variable of type Serial and call it whatever you wish. In my case it is called *port*.
3. This line prints all available ports into the console of the Processing IDE.
4. This part of the code sets up the connection for the port variable, and it just picks the first one on the list with *Serial.list()[0]*, and this might been to be change if you got more if connected to your PC. Also, remember to set the correct baud-rate.
5. This code chunk checks if the number of ports available is above 0, and if it is it reads from the serial connection and writes it to the variable *inByte*, which is displayed in the console in the text line.

The code will not run if you do not unless you got the Arduino connected. Now that you got My Keepon hacked, the Arduino running, and Processing connected you are ready to play around with My Keepon as you wish to. On the DVD you can find the code used for the experiments.

Appendix B

Questionnaires

The Danish version of the questionnaires EPQR-A, Godspeed and general information (age and gender) is found in the following sections. It will therefore be in Danish the following section.

For all questionnaire the information given by the participants is anonymous

B.1 Demografi

Køn *

◯ Mand

◯ Kvinde

Alder *

[*Skriv alder*]

B.2 Om dig *(EPQR-A)*

Er du en snakkesalig person? *

◯ Ja

◯ Nej

Er du ret livlig? *

◯ Ja
◯ Nej

Kan du let få liv i en ellers kedelig fest? *

◯ Ja
◯ Nej

Har du det med at holde dig i baggrunden ved sociale lejligheder? *

◯ Ja
◯ Nej

Er du for det meste stille når du er sammen med andre? *

◯ Ja
◯ Nej

Ser andre folk dig som værende meget livlig? *

◯ Ja
◯ Nej

B.3 Om robotten *(Godspeed)*

Hvilken robot snakkede du med? *

[*Cat, My Keepon, Nao, Telenoid*]

Angiv venligst på skalaen hvordan dit indtryk af robotten var *

Sæt kun et kryds for hvert spørgsmål

	1	2	3	4	5	
Inkompetent	◯	◯	◯	◯	◯	Kompetent *

	1	2	3	4	5	
Uvidende	◯	◯	◯	◯	◯	Vidende *

	1	2	3	4	5	
Uansvarlig	◯	◯	◯	◯	◯	Ansvarlig *

	1	2	3	4	5	
Uintelligent	◯	◯	◯	◯	◯	Intelligent *

	1	2	3	4	5	
Tåbelig	◯	◯	◯	◯	◯	Fornuftig *

Appendix C

Topics

The list of topics that the participants can choose from is the following (and it is in Danish).

C.1 Emner

Du skal vælge forskellige emner at snakke om.

Her er en liste over mulige samtale emner, men hvis du selv har et andet emne er du også velkommen til at vælge det.

- Sport
- Kæledyr
- Mad
- Bøger
- Film
- Musik
- Rejser
- Arbejde
- Skole
- Spil
- Biler
- Ferie
- Hobbyer
- Ambitioner
- Mode

Bibliography

[1] JENS ALLWOOD and LOREDANA CERRATO. “A study of gestural feedback expressions”. In: *First nordic symposium on multimodal communication*. Copenhagen. 2003, pp. 7–22 (see p. 7)

[2] CHRISTOPH BARTNECK et al. “Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots”. In: *International journal of social robotics* 1.1 (2009), pp. 71–81 (see p. 23)

[3] J B BAVELAS, L COATES, and T JOHNSON. “Listeners as co-narrators”. In: *Journal of personality and social psychology* 79.6 (Dec. 2000), pp. 941–952 (see p. 6)

[4] *Beatbots/MyKeepon*. `https://github.com/beatbots/mykeepon` (see p. 43)

[5] PETER ELSASS. *Assessmentmetoder: haÌŁndbog for psykologer og psykiatere*. [Virum]; [KÃÿge]: Dansk Psykologisk Forlag ; [eksp. DBK], 2006 (see p. 23)

[6] ANDY P. FIELD. *How to design and report experiments*. London ; Thousand Oaks, Calif: Sage publications Ltd, 2003 (see p. 15)

[7] LESLIE J FRANCIS, LAURENCE B BROWN, and RONALD PHILIPCHALK. “The development of an abbreviated form of the Revised Eysenck Personality Questionnaire (EPQR-A): Its use among students in England, Canada, the USA and Australia”. In: *Personality and individual differences* 13.4 (1992), pp. 443–449 (see p. 23)

[8] M. HEERINK et al. “The influence of social presence on acceptance of a companion robot by older people”. In: *Journal of Physical Agents* 2.2 (2008), pp. 33–40 (see p. 7)

[9] DIRK HEYLEN. “Challenges ahead: head movements and other social acts during conversations”. In: ed. by LYNN HALLE et al. Hatfield, UK: The Society for the Study of Artificial Intelligence and the Simulation of Behaviour, 2005, pp. 45–52 (see pp. 6, 7)

[10] DIRK HEYLEN et al. “Generating listening behaviour”. In: *Emotion-Oriented Systems*. Springer, 2011, pp. 321–347 (see pp. 6, 7)

[11] TOMOKO KODA et al. “Cultural Study on Speech Duration and Perception of Virtual Agent’s Nodding”. In: *Intelligent Virtual Agents*. Ed. by YUKIKO NAKANO et al. Vol. 7502. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2012, pp. 404–411 (see pp. 5, 7–12, 20, 23, 26, 27, 37)

[12] ANDERS KROSAGER et al. *Influence of Backchannel Head Nod by a Physical Agent in First Conversational Meeting with Danish Participants*. Tech. rep. Aalborg University, 2013 (see pp. 12, 32)

[13] SENKO K MAYNARD. “Interactional functions of a nonverbal sign Head movement in japanese dyadic casual conversation”. In: *Journal of Pragmatics* 11.5 (1987), pp. 589–606 (see p. 7)

[14] EVELYN Z MCCLAVE. “Linguistic functions of head movements in the context of speech”. In: *Journal of Pragmatics* 32.7 (2000), pp. 855–878 (see p. 7)

[15] PATRIZIA PAGGIO and COSTANZA NAVARRETTA. “Feedback and gestural behaviour in a conversational corpus of Danish”. In: *Proceedings of the 3rd Nordic symposium on multimodal communication, NEALT*. 2011, pp. 33–39 (see p. 8)

[16] PATRIZIA PAGGIO and COSTANZA NAVARRETTA. “Head movements, facial expressions and feedback in Danish first encounters interactions: a culture-specific analysis”. In: *Universal Access in Human-Computer Interaction. Users Diversity*. Springer, 2011, pp. 583–590 (see pp. 7, 8)

[17] RONALD POPPE, KHIET P TRUONG, and DIRK HEYLEN. “Backchannels: quantity, type and timing matters”. In: *Intelligent Virtual Agents*. Springer. 2011, pp. 228–239 (see pp. 6, 7, 26)

[18] RONALD POPPE et al. “Backchannel Strategies for Artificial Listeners”. In: *Intelligent Virtual Agents*. Ed. by JAN ALLBECK et al. Lecture Notes in Computer Science 6356. Springer Berlin Heidelberg, Jan. 2010, pp. 146–158 (see p. 6)

[19] MATTHIAS REHM et al. “From observation to simulation: generating culture-specific behavior for interactive systems”. In: *AI & society* 24.3 (2009), pp. 267–280 (see p. 8)

[20] *Reverse Engineering My Keepon*. https://github.com/qdot/keepoff/blob/master/doc/keepon_reverse_engineering.asciidoc (see p. 43)

[21] LAUREL D. RIEK, PHILIP C. PAUL, and PETER ROBINSON. "When my robot smiles at me: Enabling human-robot rapport via real-time head gesture mimicry". In: *Journal on Multimodal User Interfaces* 3.1-2 (Mar. 2010), pp. 99–108 (see p. 7)

[22] M. SCHRODER et al. "Building Autonomous Sensitive Artificial Listeners". In: *IEEE Transactions on Affective Computing* 3.2 (2012), pp. 165–183 (see p. 7)

[23] NICOLAJ SEGATO, ANDERS KROGSAGER, and MATTHIAS REHM. "Backchannel Head Nods In Danish First Meeting Encounters With A Humanoid Robot: The Role Of Physical Embodiment". 2014 (see pp. 5, 9, 20, 26, 27)

[24] V. H. YNGVE. "On getting a word in edgewise". In: *Chicago Linguistics Society, 6th Meeting* (1970), pp. 567–578 (see p. 6)