Människa och robot som berättare: hur gester påverkar förståelse och minne

Human and Robot Narrator: How Gestures Affect Comprehension and Recollection

Johannes Rosenfrost

handledare/supervisor(s)

Birger Johansson

KOGM20

June 4, 2023

Masteruppsats i kognitionsvetenskap
Avdelningen för kognitionsvetenskap
Filosofiska institutionen
Lunds universitet

Master's thesis (2 years) in Cognitive Science
Lund University Cognitive Science
Department of Philosophy
Lund University

Human and Robot Narrator:
How Gestures Affect Comprehension and Recollection

Johannes Rosenfrost
jo7118he-s@student.lu.se

This thesis investigated if a human or a robot narrator affected listeners' comprehension and memory recollection similarly in a narrative story. 103 participants ($Mean_{Age}$ = 35) were recruited to an online experiment to investigate if gestures affected the participants' narrative comprehension and recollection more than a narrator that did not produce gestures. Participants were presented with the same narrative setting in four different conditions; either with a human or a robot narrator, with or without gestures. Participants answered questions in a questionnaire, and in total, there were 21 questions regarding the narrative, of these, 13 questions were directly related to a gesture, and 8 questions were related to the narrative without any accompanying gesture. The gestures produced were congruent redundant iconic and deictic gestures. Participants were graded and analyses were conducted on three dependable variables; total sum, the sum of the gesture-related questions, and the sum of no-gesture-related questions. The total sum was used to test the hypotheses. The sum of gesture-related and no-gesture-related questions were analyzed to see if gestures affect comprehension and recollection of specific and/or general information. A second experiment was also conducted on the same participants as in the first experiment to see how participants perceived the human and robot gestures. Most participants perceived the gestures produced by the narrator as similar to each other. The results found in the experiment of this thesis showed a small indication that participants who viewed the robot narrator, both in the condition with gestures and the condition without gestures, had a stronger benefit in comprehension and recollection compared to the human narrator conditions. Similar indications were found in the gesture-related questions and a strong indication were found in the no-gesture related questions. The credible intervals overlapped between the human and robot narrator conditions, meaning that the conditions with gestures scored similar to each other, regardless of the narrator, and the same applies to the conditions without gestures. The gesture conditions scored significantly higher compared to the conditions without gestures, showing that human and robot narrators producing gestures affected the listener's compression and recollection in a similar way.

1 Introduction

Communication is far wider than just verbal communication. Facial expressions, proximity, gazes and gestures are a natural part of human interaction. When we speak and want to convey new information or emphasize words, we often use gestures. Gestures are performed in various types of communication, including narration, instruction, and conversation (Cassell, McNeill, & McCullough, 1999; Cook & Fenn, 2017; Dargue, Sweller, & Jones, 2019; Goldin-Meadow, 2003; McNeill, 1992).

Speech and gestures form a unified system in speech production and speech comprehension; they are an integrated part of human multimodal communication. Multimodal communication is communication that uses more than one modality; one modality is, for example, purely verbal communication, and another modality could be gestures. Multimodality is for example, gestures accompanying speech (Clark, 1996; Cook & Fenn, 2017; Goldin-Meadow, 2003; Kelly, Barr, Church, & Lynch, 1999; Kendon, 2004; McNeill, 2005).

Verbal communication is often used to describe something, and gestures draw attention to something in the communication, meaning that gestures act as visual-spatial support to the accompanying verbal communication (Cook & Fenn, 2017; Kendon, 2004; McNeill, 2005; Valenzeno, Alibali, & Klatzky, 2003). Gestures can infer more or new information than just verbal communication, for instance, a gesture of throwing a basketball, accompanied by saying "They played a game". Gestures that accompany verbal communication are called co-speech gestures. These gestures are by default multimodal communication (Cassell et al., 1999; Clark, 1996; Cook & Fenn, 2017; Goldin-Meadow, 2003; Kelly et al., 1999; Kendon, 2004; McNeill, 1992, 2005; Valenzeno et al., 2003).

There are several categories of co-speech gestures, a widely used definition comes from McNeill (1992); beat, iconic, deictic, and metaphoric. These gestures refer to specific information (for example, spatial information), and they are created by different movements (Cassell et al., 1999; Kendon, 1980; McNeill, 1992). Here are some examples:

- *Beat* gestures are simple movements that have no associated semantic meaning, for example, flicking with your hand. Beat gestures emphasize verbal information or reflect the tempo of speech
- *Iconic* gestures are movements that refer to physical phenomena, for example, creating a circle with your hands, and saying "ball"
- *Deictic* gestures are movements that refer to spatial items or direction, for example, pointing with a finger towards a direction, and saying "left"
- *Metaphoric* gestures are movements that represent an abstract concept, for example, moving your hands apart, and saying "space"

Experiments have long been conducted within the field of gestures, and there are several advantages for the listener within gesture research, some examples are listener's comprehension, memory recollection, second language learning, mathematics, problem-solving (Alibali, Heath, & Myers, 2001; Alibali, Spencer, Knox, & Kita, 2011; Cook, 2018; Dargue et al., 2019; Goldin-Meadow, Nusbaum, Kelly, & Wagner, 2001; Macedonia & Knösche, 2011; Valenzeno et al., 2003).

This thesis will describe how different gestures can improve listeners' comprehension and recollection skills, how we perceive and observe gestures, what kind of advantages there are with gestures, and how or why the agent producing these gestures can affect listeners' comprehension and recollection. The primary objective of this thesis is to investigate whether the comprehension and recollection of a narrative story can be influenced by the use of gestures during narration, and whether the narrator performing these gestures plays a significant role or not. In other words; can listeners be equally or similarly affected by a robot as a human narrator?

There are many forms of robots, for instance, industrial robots, drones, and social robots. Social robots are often humanoid robots, such as Epi (Johansson, Tjøstheim, & Balkenius, 2020), NAO (Aldebaran, n.d.), or BERTI (Bremner, Pipe, Melhuish, Fraser, & Subramanian, 2009). The focus of this thesis will be on humanoid social robots. These robots are designed to be similar to humans, with functions and appearance resembling a human. These robots are perceived as social interactively agents, with similar social functions as humans, among these functions are gestures (de Wit, Vogt, & Krahmer, 2023; Fong, Nourbakhsh, & Dautenhahn, 2003; Saunderson & Nejat, 2019; Van den Berghe, Verhagen, Oudgenoeg-Paz, Van der Ven, & Leseman, 2019).

Since humans and social robots share social and communicative functions, it is of interest to study social robots and their effects on humans. This thesis will utilize Epi (see Figure 1), a humanoid robot created by Lund University Cognitive Robotics Lab (Johansson et al., 2020).

The following section will describe the theoretical background of gestures and the benefits gestures have for listeners. This section will also describe how we perceive and interpret gestures.

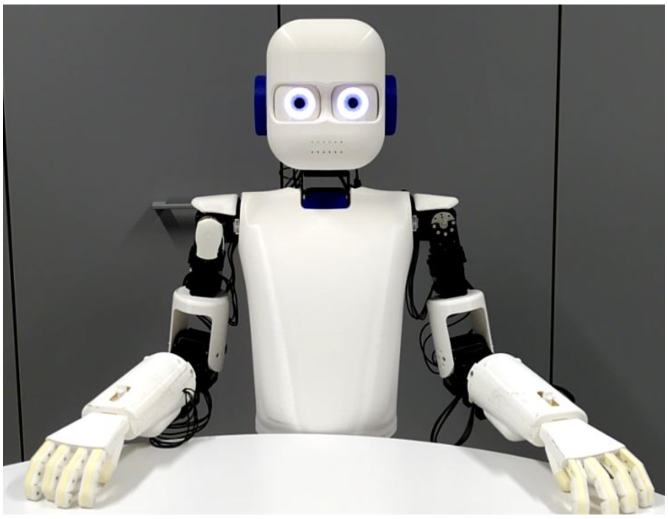

Figure 1. Epi, a humanoid robot

2 Theoretical background

Effect of gestures

There are several cognitive advantages with gestures, for example, they have been shown to decrease the cognitive load, because co-speech gestures are visual-spatial and the information in the produced co-speech sentence is processed over both visual and verbal memory (Cook & Fenn, 2017; Cook, Yip, & Goldin-Meadow, 2012; Goldin-Meadow, 1999; Goldin-Meadow et al., 2001; Wagner, Nusbaum, & Goldin-Meadow, 2004).

Due to this dual processing, gestures display additional or beneficial information with the accompanied speech. Co-speech gestures also form a unified representation, both in the production of information but also in comprehension and recollection of information. Gestures aid both listeners and speakers since they can both benefit from the additional modality (Alibali, Kita, & Young, 2000; Austin & Sweller, 2014; Cassell et al., 1999; Church, Garber, & Rogalski, 2007; Goldin-Meadow & Alibali, 2013; Kelly, Özyürek, & Maris, 2010).

In the field of gestures, there is no surprise that producing co-speech gestures influence comprehension, interpretation, and recollection of verbal information (Dargue et al., 2019; Hostetter, 2011; Kelly et al., 2010; McNeill, 1992). Gestures captures attention, and provide redundancy and clarification, as Valenzeno et al. (2003) points out. In their experiment children learned more from lessons which included both speech and gestures compared to lessons without gestures. This shows that gestures can improve comprehension and recollection since gestures reinforce the verbal information (Cassell et al., 1999; Church et al., 2007; Cook & Fenn, 2017; Kelly et al., 2010; McNeill, 2005; Overoye & Storm, 2019; Valenzeno et al., 2003).

There are several studies that show the benefits of gestures in both produced and observed gestures (Dargue et al., 2019; Hostetter, 2011). For instance, Kartalkanat and Göksun (2020) study, where a story was narrated with either iconic, beat, or without gestures to both children and adult participants. Participants recalled more information when they observed gestures compared to when they did not observe gestures. Iconic gestures had more impact on recollection than beat gestures (Kartalkanat & Göksun, 2020).

Similar results have been found in a study by Cutica and Bucciarelli (2008), where participants viewed video recordings of a narrator telling a story. In their study, gestures were produced in one condition and no gestures were produced in the other condition. Participants in the gesture condition recollected more information than the other condition (Cutica & Bucciarelli, 2008).

Furthermore, there are some studies that have looked into if co-speech gestures only benefit comprehension and recollection of specific information in a narrative, that is, details, or if they are beneficial for all information in a narrative, that is, general information. For instance, Bharadwaj, Dargue, and Sweller (2022) pointed out that some studies have looked into recalling general information, where the participants were asked questions like "Tell me everything you remember". Bharadwaj et al. (2022) discussed that there are inconclusive results regarding the recollection of general information. The studies conducted by Bharadwaj et al. (2022); Dargue and Sweller (2020); Dargue et al. (2019); Macoun and Sweller (2016) however, concluded that gestures mostly improve specific information with accompanying gestures (the questions about specific information, were open-ended, for example, "What did they do?"), but there can be a potentiality to recall general information. Bharadwaj et al. (2022) study looked into if producing gestures at encoding (the listener produced gestures while listening) could benefit both specific and general recall. Bharadwaj et al. (2022) concluded that the participants were better at recalling specific information rather than general information. Dargue et al. (2019) meta-analysis studied if observing gestures had similar results, that is, if observing gestures benefits specific or general information. In their meta-analysis, they concluded that observing gestures benefit comprehension and recollection regardless if the information is specific or general. However, they discussed that gestures only benefit general questions in some cases (Bharadwaj et al., 2022; Dargue & Sweller, 2020; Dargue et al., 2019; Macoun & Sweller, 2016).

In short, gestures affect comprehension and recollection, and when gestures are produced or observed people can recall more of the verbal information than they would without gestures (Bharadwaj et al., 2022; Cameron & Xu, 2011; Dargue et al., 2019; Hostetter, 2011; Overoye & Storm, 2019).

Social robots

As mentioned previously, humanoid robots are designed for interaction (Baxter, Kennedy, Senft, Lemaignan, & Belpaeme, 2016; de Wit et al., 2023; Fong et al., 2003; Saunderson & Nejat, 2019). Robots have an embodied presence, similar to humans, and since social robots have an embodied physical presence that have an affordance[1] for social interaction. Robots have similar non-verbal cues as humans since robot also have an embodied presence, like humans have (Baxter et al., 2016; de Wit et al., 2023; Michaelis & Di Canio, 2022; Saunderson & Nejat, 2019). Robots' gestures, gaze, and proximity can affect how humans interact with and perceive them, in a wide range of situations and domains. Studies of embodied cognition have shown how important non-verbal communication can be (Bartneck et al., 2020; Baxter et al., 2016; Broadbent, 2017; de Wit et al., 2023; Fong et al., 2003; Michaelis & Di Canio, 2022; Robins, Dautenhahn, te Boekhorst, & Billard, 2003; Saunderson & Nejat, 2019; Van den Berghe et al., 2019), therefore robots' communicative effects are of interest to study.

Furthermore, humans anthropomorphize robots, which means that we perceive them as having human behavior, such as social and communicative behavior. For example, movements that are similar to gestures produced by robots will be perceived as the same gestures produced by a human (de Wit et al., 2023; Fink, 2012; Huang & Mutlu, 2013; Saunderson & Nejat, 2019; Stiefelhagen et al., 2004).

Robots' embodiment and their co-speech gestures have an impact on comprehension and memory recollection. There have been several studies that have shown that robots improve children's vocabulary and language comprehension, some of these studies have used gestures, but not all (Belpaeme, Kennedy, Ramachandran, Scassellati, & Tanaka, 2018; de Wit et al., 2023; Saunderson & Nejat, 2019; Vogt et al., 2019). Similar experiments have been conducted on adults, who improved the same skills as the children. An embodied agent (robot) increased listeners' learning and comprehension better than a computer screen or similar device, in a wide range of domains, such as language learning, second-language learning, mathematics, and social skills. In short, robots have an impact on human communication similar to human-human-based communication (Baxter et al., 2016; Belpaeme et al., 2018; de Wit et al., 2023; Kennedy, Baxter, & Belpaeme, 2015a, 2015b; Kennedy, Baxter, Senft, & Belpaeme, 2016; Robins et al., 2003; Saunderson & Nejat, 2019; Van den Berghe et al., 2019; Vogt et al., 2019).

Most studies on robots producing gestures have focused on simple sentences and not a complex story, unlike much of gesture-based studies have done. For complex stories see for example Bharadwaj et al. (2022); Dargue and Sweller (2018a, 2018b, 2020); Kartalkanat and Göksun (2020), and for robot gestures studies see for example Bremner and Leonards (2016); Cabibihan, So, and Pramanik (2012); Van Dijk, Torta, and Cuijpers (2013). In Bremner and Leonards (2016) study participants saw iconic gestures produced by either a human or a robot. These gestures accompanied verbal communication in the form of simple sentences (for example "I played chess"). Gesture comprehension was tested, which means, that participants was tested if they could comprehend the gesture produced. In the study, they concluded that participants could identify the robot's gestures. However, participants were better at identifying human gestures than robot gestures (Bremner & Leonards, 2016).

Similar results were found in Cabibihan et al. (2012) study. The study used a humanoid robot that produced gestures and studied if participants could comprehend these gestures. Cabibihan et al. (2012) used iconic gestures produced by either a human or a robot and then asked the participant if they could comprehend the gestures. Half (eight) of the gestures produced by a human were comprehended by the participants, and six of the comprehended gestures were also comprehended in the robot gesture condition (Cabibihan et al., 2012). Humans perceive and comprehend robot gestures in a similar way as human gestures. However, as previously mentioned, studies on robot gestures have mostly used short sentences. The only study found with a more complex story did not yield the same results as previous human-human gesture studies (Bremner & Leonards, 2016; Bremner et al., 2009; Bremner, Pipe, Melhuish, Fraser, & Subramanian, 2011; Cabibihan et al., 2012; Stolzenwald & Bremner, 2017; Van Dijk et al., 2013), which will be discussed in a later section.

So it seems that humans perceive human and robot gestures in a similar way and that humans and robots have similar gestural behavior. There are however questions regarding the perception of gestures; what is the difference between a gesture and other hand movements? Are the beneficial effect of gestures dependent on how we recognize and perceive them? In the next section, these questions will be discussed.

Different types of gestures

It is important to note that gestures are intentional, goal-directed movements, but not all movements are gestural (Kendon, 1994; Novack, Wakefield, & Goldin-Meadow, 2016). Novack et al. (2016) conducted an experiment with object-oriented movements and gestures. The participants watched a silent video recording of a person either manipulating an object with their hand (touching it), doing the same movement but without an object, or doing the movement above the object. The study concluded that the participants could distinguish what movements were object-orientated and what movements were gestural (Novack et al., 2016).

Novack et al. (2016) study showed that the gestures we observe form our way to think and communicate with each other. This can be seen in a study by Holler and Wilkin (2011). The study concluded that a speaker's gestures are observed and sometimes reproduced by the listener. An example of a

[1] "The term *affordance* refers to the relationship between a psychical object and a person (or for that matter, any interacting agent, whether animal or human or even machines and robots). An affordance is a relationship between the properties of an object and the capabilities of the agent that determine just how the object could possibly be used." (Norman, 2013, p.11)

gesture from their study is a speaker flapping their arms to refer to a figure similar to a bird. These gestures were reproduced by the listener when they communicated about this figure back to the listener. These mimicry movements enhance our communication since the speaker and listener establish a mutually shared understanding (Holler & Wilkin, 2011).

Gestures, as mentioned before, aid our comprehension because of the additional information. For instance, deictic gestures aid our spatial communication due to the spatial information produced by the gestures (Austin & Sweller, 2014; Austin, Sweller, & Van Bergen, 2018). Different gestures have different results on comprehension. For example, beat gestures have shown mixed results on comprehension, some experiments with beat gestures have shown improvement in comprehension, and some have not. Deictic, iconic, and metaphoric gestures seem to have an effect on comprehension and memory recollection (Aussems & Kita, 2019; Austin et al., 2018; Dargue et al., 2019; Hostetter, 2011). However, these gestures are produced spontaneously and often mixed with each other in the same sentence. These gestures also carry specific semantic information, that is, meaningful information for the speaker and the listener (Aussems & Kita, 2019; Austin & Sweller, 2014; Austin et al., 2018; Dargue et al., 2019; Hostetter, 2011; McNeill, 1992).

As mentioned before, deictic gestures mostly convey spatial information, like pointing to an object, however, they also require spatial information. Metaphoric gestures are gestures that represent abstract concepts, but these risk losing their meaning since they depend on the speaker's intention and the listener's interpretation, for example moving your hands apart to describe "space", which can refer to personal space, outer space, or as a distance. Iconic gestures are gestures that depict psychical phenomena such as pulling your hands over your head to show putting on a hat. Iconic gestures are easy to comprehend and encode, due to their clear semantic meaning. Iconic gestures do not require any spatial information and they are less dependent on interpretation compared to metaphoric gestures, that is, iconic gestures are less dependent on context (Aussems & Kita, 2019; Austin & Sweller, 2014; Austin et al., 2018; Cassell et al., 1999; Dargue et al., 2019; Hostetter, 2011; Kendon, 1980; McNeill, 1992).

How deictic, iconic, and metaphoric gestures contain semantic meaning and how we define and perceive the information depends on if the gestures are redundant or non-redundant. An example of a non-redundant gesture is indirect request as studied in Kelly et al. (1999) where pointing at a window and saying "I'm getting hot". In this example, listeners are unable to interpret the whole meaning without gestures and speech. If the gestures contain new semantic meaning not produced in the accompanying verbal information this is a non-redundant gesture. If the gesture contains the same information as the verbal information then it is redundant (Cassell et al., 1999; Dargue et al., 2019; Hostetter, 2011; Kelly et al., 1999).

As previously mentioned, non-redundant gestures contain semantic information that is not present in the verbal information. Non-redundant gestures give the listener contextual meaning of verbal communication. However, there are some instances where gestures could be confusing for the listener. The gesture and the verbal communication can be mismatched or incongruent. A mismatched/incongruent gesture has either no semantic relation to verbal communication or is simply incorrect, like saying "putting on a hat" but gesturing "putting on pants". Atypical gestures are gestures that are produced less frequently, or gestures that are produced too early or too late compared to the accompanying verbal communication. Matched gestures (also called congruent) are gestures produced with a clear semantic relation to verbal communication. Typical gestures are produced more frequently, for instance, saying "They played a game" and gesturing to throw a basketball. Non-redundant gestures can be both congruent and incongruent, while a redundant gesture is often congruent (except if a listener interprets it faulty). As previously mention, redundant gestures contain the same semantic meaning in both the verbal communication and the accompanying gesture, which means, redundant information (Cassell et al., 1999; Dargue & Sweller, 2018b; Hostetter, 2011; Hostetter & Alibali, 2011; Kelly, Kravitz, & Hopkins, 2004; Kelly, McDevitt, & Esch, 2009; McKern, Dargue, Sweller, Sekine, & Austin, 2021; McNeil, Alibali, & Evans, 2000; McNeill, 2005).

Several studies have looked into the different types of gestures; typical, atypical, congruent, in-congruent, redundant, and non-redundant gestures (Dargue et al., 2019; Hostetter, 2011). Macoun and Sweller (2016) conducted a study where both redundant and non-redundant gestures were produced by a narrator telling a story. They found that children observing non-redundant and redundant gestures both improved comprehension of the narration compared to no gestures and beat gestures with the same narration. However, the experiment had no separate condition for non-redundant and redundant gestures (Macoun & Sweller, 2016).

In Dargue and Sweller (2018a) study they used iconic gestures that they classified as reinforcing or contradictory. A video of a narrated story with either iconic gestures or beat gestures was presented to children. Dargue and Sweller (2018a) concluded that iconic reinforcing gestures improved children's comprehension, and that contradictory gestures had no significant effect on comprehension (Dargue & Sweller, 2018a).

Hostetter (2011) conducted a meta-analysis that studied if either non-redundant or redundant gestures benefited the listener more. The meta-analysis drew the conclusion that non-redundant gestures have a bigger effect on listeners' comprehension. Hostetter (2011) argued that this benefit exists due to non-redundant gestures provides new information. However, the author pointed out that the difference between non-redundant and redundant gestures is not entirely clear (Hostetter, 2011).

In a later meta-analysis conducted by Dargue et al. (2019), they concluded that no matter what type of gestures (redundant or non-redundant, congruent or incongruent, typical or atypical), the gestures still improve comprehension and recollection. Furthermore, they concluded that deictic, metaphoric, and iconic gestures improved comprehension and recollection. Beat gestures were shown to have less impact on comprehension and recollection. They didn't make a full meta-analysis of beat gestures but pointed out that later published articles seemed to conclude that even beat gestures improve comprehension (Dargue et al., 2019).

Overall, no matter the gesture type, gestures have been shown to improve comprehension and recollection of the information given. Congruent gestures have been shown to affect comprehension and recollection more compared to incongruent gestures (Cassell et al., 1999; Dargue et al., 2019; Hostetter, 2011). Furthermore, non-redundant gestures seem to have a higher risk of being misinterpreted compared to redundant gestures. In the next section the research question, experimental conditions, and hypothesis will be discussed.

This thesis

The purpose of the thesis is to investigate whether gestures have an impact on the comprehension and recollection of a narrative story and whether the agent performing these gestures plays a significant role on the impact of comprehension and recollection. In previous studies with robot gestures, the focus has been mostly on iconic, deictic, and beat gestures in simple sentences (Bremner & Leonards, 2016; Bremner et al., 2009; Cabibihan et al., 2012; Stolzenwald & Bremner, 2017; Van Dijk et al., 2013). In Bremner et al. (2011) a similar setting to this thesis can be found. In Bremner et al. (2011) experiment, participants viewed a humanoid robot telling a story that either produced gestures or did not produce gestures. The story regarded information about their laboratory and the robot produced both beat and what they called representational gestures. However, the results indicated that there were no significant differences in recollection between the gesture and no-gesture conditions (Bremner et al., 2011).

This thesis aimed to conduct similar experimental conditions as described in for instance Bremner et al. (2011); Dargue and Sweller (2018b, 2020) but also from further similar experiments studied in Dargue et al. (2019); Hostetter (2011) where the narration is more complex than simple sentences (Bremner et al., 2011; Dargue & Sweller, 2018a, 2018b; Dargue et al., 2019; Hostetter, 2011).

Since humans are used to seeing human speakers producing gestures or not, and these observed gestures improve comprehension and recollection, is this also true for robots? To answer this question, a defined setting was created with the use of congruent redundant, iconic gestures, and two deictic gestures (up and down). Another effect that this thesis studied was if observing gestures affected comprehension and recollection of specific or general information. This effect was studied, separately from the hypotheses. In Dargue et al. (2019) meta-analyses they concluded that observing gestures had an effect on recalling both specific information and general information in a narrative, but that gestures benefited comprehension and recollection of specific information more than general information. This thesis, therefore, used open-ended questions both for information with accompanied gestures (gesture-related questions), and without accompanied gestures (no-gesture-related questions).

The research question this thesis aimed to answer were:

- Do co-speech gestures produced by a human or a robot affect memory recollection and comprehension of a narrative story in a similar way to each other?

In this context, "a similar way" means that the listener should have similar benefits in comprehension and recollection. This was investigated by scoring participants' answers regarding the narrative. Participants should score similarly in the conditions with gestures, regardless of the narrator, and the same for the condition without gestures.

To investigate the research question and the hypothesis four different experimental conditions were created (see Table 1).

The research question entailed three hypotheses:

Table 1. Experimental conditions

Condition 1	Robot narrator without gestures
Condition 2	Robot narrator with gestures
Condition 3	Human narrator without gestures
Condition 4	Human narrator with gestures

H1 The listener should score similarly in the no-gestures conditions regardless of the narrator

H2 The listener should score similarly in the gestures conditions regardless of the narrator

H3 The participants in the gestures conditions should score higher compared to the no-gesture conditions regardless if the narrator is a human or a robot

This experiment used video recordings to ensure that all participants received the same information, which ensured that all conditions could be equally compared to each other.

3 Method

The experiment utilized a between-subject design, with each participant being assigned to watch only one of the four different videos. Each video contained the same story. The different videos correlate to the conditions seen in Table 1 (Video 1 correlates to Condition 1, and so on, see Figure 2). The participants were told to pay attention to the narrator but were not informed of the purpose of the questionnaire until they were debriefed after answering the questionnaire.

The narrative story is about Sara who is late for work. The story contains key plot moments that have accompanied gestures or not, depending on the condition. These plot moments were used as a base for the questions the participants answered. The story was roughly 1 minute and 50 seconds long and it is in Swedish. Open-ended semantic questions were asked, such as "What was the name of the main character in the story?", or "How did the main character get on the bus?". In total there were 21 questions, of these questions, 13 were directly related to a semantic congruent redundant gesture (gesture-related question), and 8 questions were unrelated to a co-speech gesture (no-gesture-related question). The gesture-related questions were of the type: "What did the main character do first before she sat down in her room?", "How did the main character greet the receptionist?". The no-gesture-related questions were of the type: "What was the name of the main character in the story?", "What day was it?".

The same voice is utilized in all videos. Videos 1 and 3 contain identical voice recordings, and Video 2 and 4 contain identical voice recordings. The videos were recorded in the Cognitive Robotics Lab at Lund University. All videos contained the narrator (human or robot) behind a white round table and with a neutral grey background, which ensured that there were no visual distractions. In the gestures conditions, the narrators' hands were placed on the table, so that the hands could easily be seen. For the no-gesture conditions, the arms of both narrators were kept at the side and did not move, but the table was still there (see Figure 2).

The gestures produced in the experiment were selected to be appropriate for the story and to align with the physical capabilities of the robot. These gestures were executed at specific points in the plot, as mentioned before, and similar

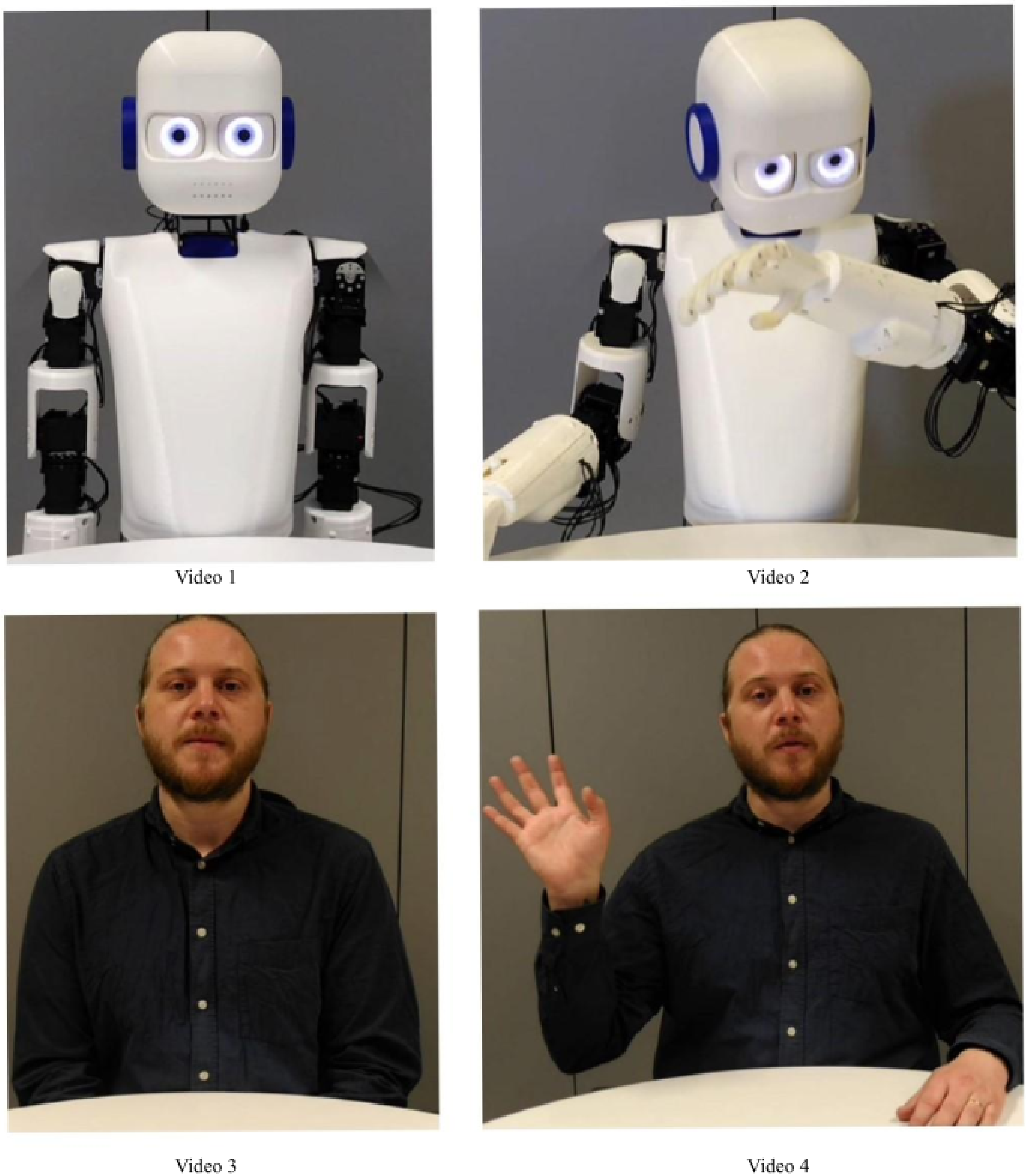

Figure 2. The videos that correlate to the experimental condition

gestures were used by both the robot and the human narrator. The primary constraint of the robot's movement is the restricted degrees of freedom. Epi's arms are equipped with five degrees of freedom, with three in the shoulder, one in the elbow, and one in the wrist. Epi has one degree of freedom in the torso. The torso can move along its central axis, meaning it can rotate sideways. The head has two degrees of freedom and can tilt and pan (Johansson et al., 2020). These limitations had to be accounted for in the human narrator settings so the conditions could be similar.

The movements that Epi produced were created via puppeteering, meaning that the researcher moved the robot and its joints to produce the gestures. These movements were recorded via a digital platform running the Ikaros system that controls the robot, the digital platform communicates back to Epi (Balkenius, Morén, & Johansson, 2007; Balkenius, Morén, Johansson, & Johnsson, 2010; Johansson et al., 2020). Epi also has LED lights. To create similarity between conditions, the mouth lights up at random intervals to simulate speaking. The intervals were also created via puppeteering.

Table 2. Participants in the experimental conditions.

Video	Participants	$Mean_{Age}$	SD_{Age}	Male	Female	Non-binary
Video 1: Robot narrator without gestures	25	35 (range 19-69)	14.6	11	13	1
Video 2: Robot narrator with gestures	25	33 (range 20-68)	9.0	13	11	1
Video 3: Human narrator without gestures	26	38 (range 18-73)	15.4	17	8	1
Video 4: Human narrator with gestures	27	33 (range 19-64)	11.8	11	16	0

Pilot

A pilot study was conducted for two reasons; first to make sure the robot's behavior and gesturing were perceived as human, and second to see if the story or gestures needed changes. The pilot ensured that the experimental design was sound. The pilot was conducted on eight people, two participants in each of the four conditions. The questionnaire was created in Google Forms (Google , n.d.). The robot narrator's gestures were perceived as similar to the human narrator's gestures by the participants. No statistical analysis was conducted on the data from the pilot. The story and the gestures were rewritten since some of the gestures produced did not correspond to the story or were non-redundant or similar.

Experiment

The experiment was conducted online via a questionnaire created in Psytoolkit (Stoet, 2010, 2017). The questionnaire was programmed so that the videos could only be seen once in the questionnaire. The videos were uploaded on YouTube (Youtube, n.d.). After the video, the questions were presented to the participants one at a time. Participants answered via written form in the questionnaire instead of multiple choices answers so the participants could not answer correctly by accident. The answers were later graded with scores. The max score was 42, and the questionnaire contained 21 questions regarding the story (see Appendix for URLs, script, and questionnaire)

Participants

Swedish-speaking participants were recruited via email correspondence to teachers, doctorates, professors, et cetera at the university level, from several universities in Sweden and over several subject fields, for instance, linguistics, information technology, psychology, and cognitive science. Further participants were recruited via friends, family, Facebook (Meta Platforms, n.d.), posting the questionnaire on SurveyCircle (SurveyCircle, n.d.), posting the questionnaire on Accindi (Accindi, n.d.), and finally, flyers were handed out over several weeks to people on the campus of Lund University. In total in the experiment, there were 126 participants, 23 participants were excluded since they did not answer all questions. In total, 103 participants completed the questionnaire ($Mean_{Age}$ = 35, age range 18-73, SD_{Age} = 14), 52 identified as male, 48 identified as female, and 3 identified as non-binary (see Table 2). Participants did not receive any compensation except for a few points on SurveyCircle.com (SurveyCircle, n.d.).

Grading

The grading was conducted by sorting answers alphabetically without the researcher knowing which of the videos the participants had viewed. Every sequential question was also sorted alphabetically to further increase a fair grading and to further differentiate participants. Participants could receive 2,1, or 0 points. 2 points were given to correct and comparable answers. An example of a question used was "How did the main character get on the bus?". The correct answer was "forcing the doors open" which would be 2 points, but the participants also got 2 points for "forces" but also for "pulls" or similar synonyms. Categorical answers on the same question such as "violently" got 1 point, and 0 points for faulty answers or "I don't remember". 10% of the data were graded by another party to ensure grading was fair and conducted properly. Grading by both parties was identical.

Second experiment

The second experiment was conducted directly after the participant had answered the narrative questions, in the same questionnaire created in Psytoolkit (Stoet, 2010, 2017). This experiment was conducted to see if participants perceived the gestures used by the human and robot narrator as natural or not. The video contained both the human and robot narrator, in a split-screen video, so participants viewed both narrators at the same time producing the same gestures. After the video, participants graded how natural or unnatural the human and robot narrator was. A 1-5 scale was used where 1 was "Very unnatural" and 5 was "Very natural". The same scale was used for both narrators. Participants were then asked to comment on what kind of movements seemed unnatural/natural or if it was noticeable that the robot did not move in the same way as the human. In total, there were seven questions regarding the comparison video (See Appendix).

4 Results

Statistical analyses

All analyses were conducted in RStudio (RStudio Team, 2020) running R 4.2.3 (R Core Team, 2021) and all analyses were done using Bayesian linear regression models.

The experiment contained three dependent variables:

- Total Sum (ST): the total amount of points participants received (SG and SNG)
- Sum of gesture (SG): the total amount of points participants received in the gesture-related questions
- Sum of no-gestures (SNG): the total amount of points participants received in the questions that were not related to gestures

The conditions and demographic information (age and gender) were the independent variables.

Analysing the hypotheses

The results from the questionnaire were summed (ST) and analyzed against the condition. The results from the Bayesian model are displayed in Table 3 and plotted in Figure 3. The score for participants that viewed Video 2 was 6.75 (CI 95% [3.02, 10.50]) points higher than for Video 1. The score for Video 3 was 2.56 (CI 95% [-6.23, 1.26]) points lower than for Video 1. Finally, Video 4 scored 5.82 (CI 95% [2.13, 9.59]) points higher than Video 1. The posterior distribution was 100% positive indicating high confidence in the results.

Table 3. Fitted values for summed total (ST), credible intervals (CI), and standard deviation (SD) for the summed total score per video

Video	ST	CI 95%	SD
Video 1	27.65	25.03, 30.27	6.47
Video 2	34.38	31.80, 37.01	6.47
Video 3	25.04	22.51, 27.60	6.49
Video 4	33.51	30.91, 35.94	6.93

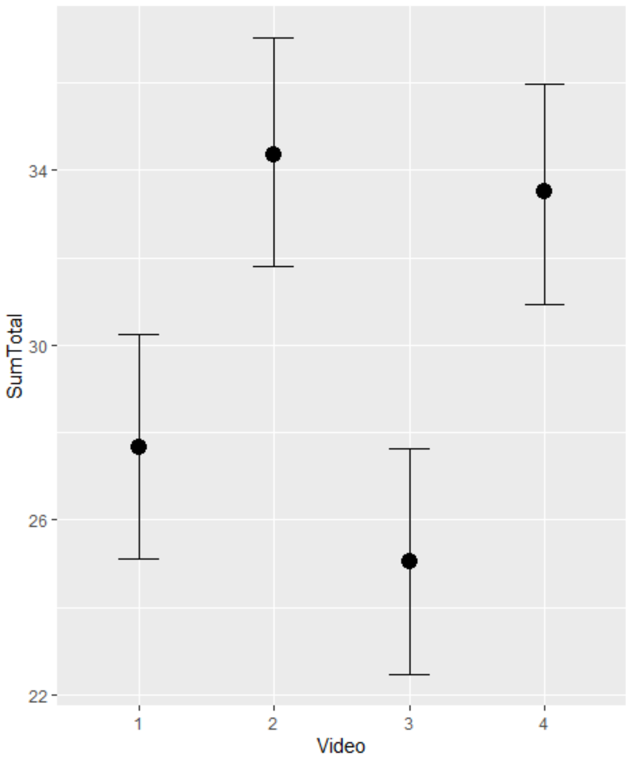

Figure 3. Difference in the scoring sum total (ST) between videos

As seen in Table 3 and Figure 3, listeners scored similarly in the no-gesture conditions meaning hypothesis 1 (which assumed listeners should score similarly in the no-gesture condition regardless of the narrator) found support. Furthermore, hypothesis 2 (which assumed listeners should score similarly in the gesture conditions regardless of the narrator) found support. As seen in Table 3 and Figure 3 the gesture conditions scored significantly higher than the no-gesture conditions, which means that hypotheses 3 (which assumed that participants in the gesture conditions should score higher compared to no-gesture conditions regardless of the narrator) also found support. There is an overlap in the credible intervals for the gesture conditions between the narrators and the same for no-gesture conditions. However there is less credible interval overlap between human and robot narrators in the no-gesture conditions, meaning that there is an uncertain indication for a possible greater effect of the robot narrator, these indications for a possible effect will be discussed later.

Further analyses

Further analyses were conducted to see if the score for gesture-related questions (SG) and score for no-gesture-related questions (SNG) improved depending on which conditions the participant viewed. This analysis was conducted to see if gestures could increase comprehension and recollection of specific information or general information. The results from gesture-related questions were summed (SG) and analyzed against the condition.

Table 4. Fitted values for summed gesture-related score (SG), credible intervals (CI), and standard deviation (SD) for the summed gesture-related score per video

Video	SG	CI 95%	SD
Video 1	13.96	11.88, 16.03	4.27
Video 2	19.84	17.88, 21.91	5.26
Video 3	11.89	9.90, 13.93	5.29
Video 4	19.19	17.18, 21.09	5.47

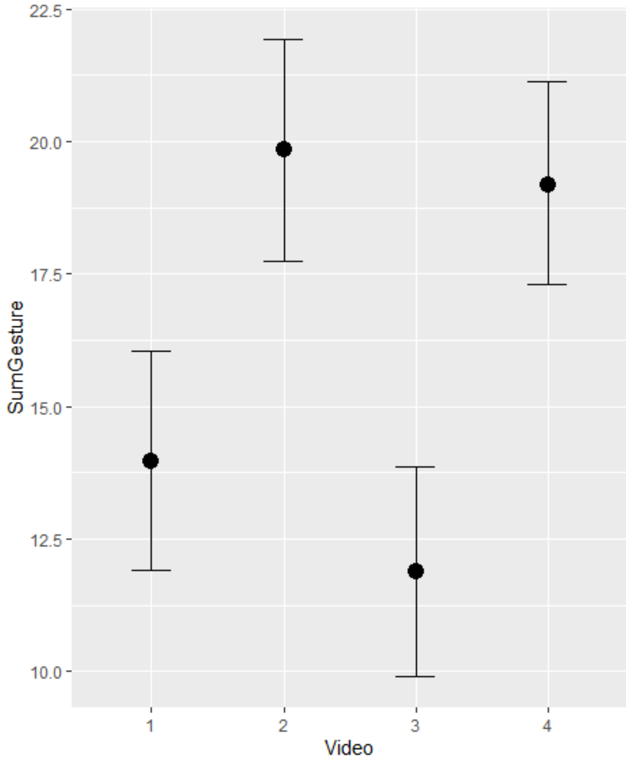

Figure 4. Difference in the scoring (SG) between videos

The results are displayed in Table 4 and plotted in Figure 4. The score indicated a significant difference between gesture conditions and no-gesture conditions. The score for the gesture-related questions for participants that viewed Video 2 was 5.94 (CI 95% [3.09, 8.73]) points higher than for Video 1. The score for Video 3 was 2.04 (CI 95% [-4.89, 0.89]) points lower than for Video 1. Finally, Video 4 scored 5.22 (CI 95% [2.34, 8.05]) points higher than Video 1. The posterior distribution was 100% positive indicating high confidence in the results. In this analysis, there is a similar overlap in credible intervals for the gesture and no-gesture conditions as in the sum total. There is also less overlap in the credible intervals for no-gesture conditions, meaning that participants did not score too differently from each other within the conditions with or without gestures. But there are indications of an uncertain effect for the lower score for participants in the human narrator conditions, here as well.

The results from no-gesture-related questions were summed (SNG) and analyzed against the condition. The results are displayed in Table 5 and plotted in Figure 5. The score for participants that viewed Video 2 was 0.27 (CI 95% [0.50, -0.73]) points higher than for Video 1. The score for Video 3 was 1.08 (CI 95% [-2.06, -0.14]) points lower than for Video 1, and finally, video 4 scored 0.05 (CI 95% [-0.92, 1.02]) points higher than Video 1. The posterior distribution was 100% positive indicating high confidence in the results.

Table 5. Fitted values for summed no-gesture-related score (SNG), credible intervals (CI), and standard deviation (SD) for the summed no-gesture-related score per video

Video	SNG	CI 95%	SD
Video 1	14.23	13.57, 14.92	3.31
Video 2	14.52	13.82, 15.20	1.53
Video 3	13.15	12.48, 13.83	1.82
Video 4	14.29	13.62, 14.96	1.83

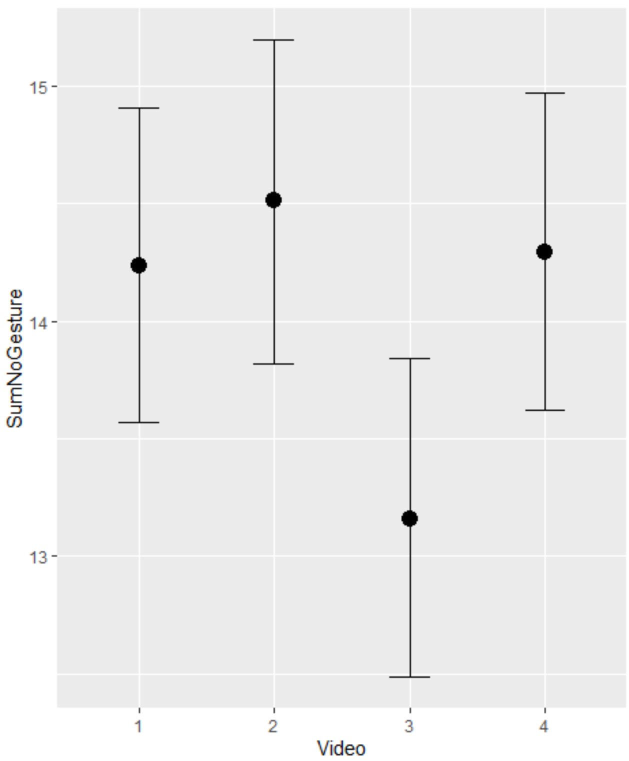

Figure 5. Difference in the scoring of sum no-gesture (SNG) between videos

As can be seen in Table 5 and in Figure 5 the score and credible intervals overlap heavily between all conditions (meaning no significance difference) except for Video 3, which had less overlap, compared to the other conditions. There are therefore strong indications that participants in Video 3 (human without gesture condition) could have had an effect on why they scored lower on the no-gesture-related questions compared to the other conditions. However, since the credible interval overlaps the other conditions, there is an uncertainty of this effect.

Demographic effects

Further Bayesian linear models were conducted to analyze demographic effects. Gender and age had no effect on any of the dependable variables (ST, SG, SNG).

Second experiment: Analysing perception of movements

The analysis of the second experiment was conducted to see if the movements of the robot and human narrator were perceived as natural or not. Participants that viewed Video 1 graded the robot narrator 3.08 (CI 95% [2.72, 3.44]) out of 5. Participants that viewed Video 2 graded the robot 0.60 (CI 95% [0.09, 1.11]) higher compared to Video 1. Participants that viewed Video 3 graded the robot 0.57 (CI 95% [0.06, 1.09]) higher compared to Video 1. Participants that viewed Video 4 graded 0.59 (CI 95% [0.08, 1.09]) higher compared to Video 1. There is a heavy overlap in the credible intervals between conditions, meaning that there was no significance between conditions (see Table 6). Posterior distribution was 100% positive.

Table 6. Fitted values for perception of robot narrator (RN), credible intervals (CI), and standard deviation (SD) for the participants grading of robot narrator per video

Video	RN	CI 95%	SD
Video 1	3.08	2.73, 3.44	0.85
Video 2	3.69	3.31, 4.05	0.78
Video 3	3.65	3.30, 4.02	0.79
Video 4	3.67	3.30, 4.01	0.76

Participants that viewed Video 1 graded the human narrator 3.76 (CI 95% [3.39, 4.12]) out of 5. Participants that viewed Video 2 graded the human 0.24 (CI 95% [-0.26, 0.75]) higher compared to Video 1. Participants that viewed Video 3 graded the human 0.02 (CI 95% [-0.51, 0.48]) lower compared to Video 1. Participants that viewed Video 4 graded 0.09 (CI 95% [-0.42, 0.60]) higher compared to Video 1. There is a heavy overlap in the credible intervals between conditions (see Table 7). Posterior distribution was 100% positive, indicating high confidence in the results, meaning that there was no significance between conditions.

Table 7. Fitted values for the perception of the human narrator (HN), credible intervals (CI), and standard deviation (SD) for the participants grading of human narrator per video

Video	HN	CI 95%	SD
Video 1	3.76	3.37, 4.11	0.83
Video 2	4.00	3.66, 4.35	0.76
Video 3	3.74	3.49, 4.19	0.87
Video 4	3.84	3.39, 4.09	0.75

Gender and age had no effect on the perception of the movements of the human or robot narrators.

The participants also answered questions in written form in the questionnaire , that regarded if and which gestures seemed unnatural. On the question if the robot's gestures seemed unnatural, 42 people answered that they did not perceive the robot's movements as unnatural, the other most common answers were the "mouth", "circle", "ball", "running", "little things", and "details" and 8 people answered that they perceived the robot's movements as unnatural. Regarding the human narrator and the same question, 46 people answered that they did not perceive the human's movements as unnatural, the other most common answers were "circle", "ball", "running" and that "the gestures were exaggerated", and 2 people answered that they perceived the human movements as unnatural. Another question that the participant answered regarded if the robot and human narrator seemed to conduct the same movements. Here all participants except for 4 answered some degree of yes. 58 people simply answered "yes", and the other participants answered "yes, mirrored sometimes", "yes, but sometimes stuttered" and so on. In the question regarding if it was noticeable that Epi could not move in the same way as the human, ("Was it noticeable that the robot could not move in the same way as the human"). 70 participants answered variants of "no", for instance, "no, but sometimes slower", "no, but sometimes it was the wrong hand", and so on. 19 participants commented that it was noticeable that the robot could not move in the same way as the human, of these 9 participants answered simply "yes" and the others "yes, it was the wrong hand, sometimes", "yes, the robot did not use their fingers", and so on, the other participants either wrote "the robot was stiffer, but no", "the robot seemed to make the same movements" or "I don't think so", "the head movements" and so on. Overall more participants perceived that the movements between narrators were more similar than not.

5 Discussion

The results found in this thesis supported all hypotheses, participants scored similarly in the conditions with gestures regardless of the narrator, the same applies to the conditions without gestures. The participants also scored higher in the gesture conditions compared to participants in the no-gesture conditions.

There were also similar results in this thesis as with previous studies with robot gestures, see, for example, Bremner and Leonards (2016); Bremner et al. (2009); Stolzenwald and Bremner (2017) and Cabibihan et al. (2012). However, in this thesis, compared to Bremner et al. (2011), the gesturing robot improved the listener's comprehension and recollection more compared to the robot without gestures, thus finding different results compared to Bremner et al. (2011). Possible explanations for the results in this thesis, compared to Bremner et al. (2011), can be the difference in the narrative content, online instead of on-site, how the robot was perceived, and how the gestures were performed.

Furthermore, there was no statistical significance between participants regarding the human and robot narrator conditions, which means that participants scored similarly within the human and robot narrator conditions. However, there is a small indication that participants in the robot narrator conditions scored higher, especially in the conditions without gestures, compared to the human narrator. There was a strong indication of an effect in the scoring of the participants in the human narrator without gesture condition, regarding the no-gesture related questions (SNG). Participants scored lower in the human without gesture condition in the questions regarding no-gestures (SNG) compared to the other conditions. There can be several reasons for the difference in scoring and the small indication that participants scored higher in the robot condition. Possible explanations will be discussed in the next sections, focusing on discussing the hypotheses first and later possible explanations for differences between narrators.

Hypothesis 1

Hypothesis 1 (which assumed that the listener should score similarly in the no-gesture conditions regardless of the narrator) found support since the participant scored similarly in the no-gesture conditions regardless of the narrator. The analysis indicated no significant difference in scoring between the robot and the human narrator in the no-gesture conditions. As mentioned in the results, there is an overlap in credible intervals between the conditions without gestures (see Table 3 and Figure 3). Participants still scored higher in the robot narrator condition, indicating a possible uncertain beneficial effect of the robot narrator compared to the human narrator in the conditions without gestures.

Hypothesis 2

Hypothesis 2 (which assumed that the listener should score similarly in the gesture conditions regardless of the narrator) found support since the participant scored similarly in the gesture conditions regardless of the narrator. The analysis indicated no significant difference in scoring between the robot and the human narrator in the gesture conditions. However, the robot narrator conditions scored slightly higher compared to the human narrator conditions. The differences between human and robot narrator conditions scores are therefore of interest to discuss, since the participants in the robot narrator conditions scored higher than the participants in the human narrator conditions. However, this effect is highly uncertain since there is a heavy overlap it still indicated the possible or potential beneficial effects of the robot narrator, compared to the human narrator.

Hypothesis 3

Hypothesis 3 (which assumed that the participants in the gestures conditions give a higher score compared to the no-gesture conditions regardless of the narrator) was tested as a replication from several similar studies, see for instance Bharadwaj et al. (2022); Dargue and Sweller (2018a, 2018b); Dargue et al. (2019); Hostetter (2011). This hypothesis was tested to see if the experimental settings were sound. The hypothesis found support, since the gesture conditions improved the listeners' comprehension and memory recollection more, no matter if the listener watched a human or a robot narrator. Gestures did improve comprehension and recollection more compared to conditions without gestures (see Table 3, and Figure 3). This thesis, therefore, had similar results compared to previous gesture studies.

Further analysis

The no-gesture and gesture-related questions were differentiated to see if gestures improved the comprehension and recollection of all narrative questions (general information) or if co-speech gestures only improved comprehension and recollection of verbal information accompanied by co-speech gestures (specific information). General and specific information has been studied previously, for example, Bharadwaj et al. (2022); Dargue and Sweller (2020); Dargue et al. (2019); Macoun and Sweller (2016). Instead of using free recall (questions like "Tell me everything you remember"), this thesis used different open-ended questions (see Appendix) to differentiate between information accompanied by gestures and information not accompanied by gestures. In the gesture-related questions, participants in the gesture conditions scored significantly higher than participants in the conditions without gestures (see Table 4, and Figure 4). Participants in the gesture conditions improved comprehension and recollection of specific information more than participants in the conditions without gestures.

The no-gesture-related questions had a higher score in the gesture condition compared to the no-gesture condition, which indicated that gestures can improve recollection of both general and specific information. The credible intervals overlap, which means that this effect was uncertain. Even heavier overlap is seen in the score of the participants in the robot without gestures condition, since they scored so similarly to the gesture conditions (see Table 5 and Figure 5). However, participants in the human narrator without gestures condition had lower scores in the sum of no-gesture, compared to the other conditions. This difference is a strong indication that there was an effect for the participants in the human narrator condition without gestures, meaning there could be an effect of why they scored lower regarding these questions. In the next section, possible explanations for similarities and differences between human and robot narrators will be discussed.

Possible explanations

The results in the human and robot narrator conditions are interesting, especially in the no-gesture-related questions. It is unclear why the robot narrator conditions indicate an improved effect on comprehension and recollection, and why participants scored lower in the human narrator conditions. A potential or possible explanation for the increased score in the robot conditions could be a novelty effect. Baxter et al. (2016); Kennedy et al. (2015a, 2015b, 2016); Van den Berghe et al. (2019) all discuss that there is a novelty effect in human-robot interaction and that robots can affect participants more compared to things participants are used to, since interaction with robots is a novelty.

Participants might be more engaged with the robot than the human, especially if participants have never interacted with a robot. This novelty effect can be a factor in why the participants scored higher in the robot conditions, especially for the no-gesture condition, since participants viewed one of the four videos only once, thus never getting too accustomed to the narrator. However, the novelty of a robot can potentially have a negative effect on participants' recollection, since they might be more focused on the robot than the task at hand, as discussed by Kennedy et al. (2015b). This negative effect was not shown in the experiment in this thesis. Nonetheless, the novelty effect could potentially be an influence on the participants' higher scores in the robot conditions.

Further possible or potential explanations for the difference in scoring between human and robot narrator conditions could be that there is an "uncanny" effect in the human narrator condition without gestures. Since humans use gestures as an integrated part of communication (Clark, 1996; Goldin-Meadow, 2003; Kendon, 2004; McNeill, 1992, 2005) it can be uncanny, or weird, for the listeners to not see gestures, or the listeners can be more skeptical towards a speaker without gestures. However, previous research has made no remarks about this, see for example, Dargue et al. (2019); Hostetter (2011). Compare this behavior (not using gestures) with a robot, where this behavior could be seen as more natural. The listener could at least be more forgiving of the robot, especially since the listeners are most likely unaware of the robot's degree of movement capabilities. Another possible or potential explanation can be that Epi might look more human or be perceived as more human than previous robots, for example, BERTI was utilized in Bremner et al. (2009, 2011). BERTI was designed for gesturing but has no face. This could be a factor in the communicative interaction since gazes and facial expressions are an important part of human communication (Cook & Fenn, 2017; Goldin-Meadow, 2003; McNeill, 1992). Furthermore, NAO that was used in Bremner and Leonards (2016); Stolzenwald and Bremner (2017) is only 58cm tall Aldebaran (n.d.) and can be argued to be perceived as more toylike than humanlike, due to its size. EPI was 101 cm in 2020 (Johansson et al., 2020) (no newer measurement reported), so Epi's size is more comparable to the human narrator than NAO, meaning that Epi is most likely perceived as more similar to humans since it both has a face and is bigger than NAO. It is important to note that both NAO and BERTI have been shown to affect humans interacting with them, (Baxter et al., 2016; Belpaeme et al., 2018; Bremner & Leonards, 2016; Bremner et al., 2009, 2011; Saunderson & Nejat, 2019; Stolzenwald & Bremner, 2017). However, since previous gesture studies have not reported an uncanny effect of a speaker that do not produce gestures, and that robot has shown results within gesture studies no matter size, or face, these explanations are speculative at best.

Another possible or potential explanation for the similarity in the conditions with gestures, and in the conditions without gestures, could be that the gestures were produced in regards to Epis movement capabilities. However, the gestures themselves should not affect the difference in the conditions too much, since they were similar, and it is unlikely that this explanation answers to the difference in scoring within conditions, or for the stronger indication of improvements for participants in the robot narrator condition without gestures.

In conclusion, of the possible and potential explanations mentioned in this section, the most reasonable explanation for the differences in scoring by the participants in the human and robot conditions, is most likely the novelty effect. The novelty effect has previously been shown to affect humans interacting with robots (Baxter et al., 2016; Kennedy et al., 2015a, 2015b, 2016; Van den Berghe et al., 2019), and it is a likely effect in the experiment in this thesis as well. Other possible explanations can be the gestures themselves, perceiving Epi as more human, or other effects, such as participants could have perceived Epi as having facial expressions or gaze behavior.Nevertheless, the results found in this thesis support that robots and their gesturing have a similar effect on listeners' comprehension and recollection as previous human gesturing studies have established, see for instance Aussems and Kita (2019); Austin and Sweller (2014); Church et al. (2007); Dargue and Sweller (2018a, 2018b, 2020); Dargue et al. (2019); Hostetter (2011).

Limitations

The first limitation in this thesis regards the experiment, specifically the gestures themselves. As previously mentioned, the gestures were created with regard to Epi's limitations, which could answer to why the score was so similar between the human and robot narrator. How participants perceived the gestures could possibly have an effect on the results. Although the analysis of the perception of gestures indicates that people perceived the movements as more natural than unnatural. The comments also indicated that participants did not differentiate between narrators too much. In short, the narrators were viewed by most participants as conducting similar movements. Regardless of these results, there still existed a limitation in the gestures due to the limitations of the robot, since the robot could not move in the same way as the human narrator. However, the gestures produced in the experiment in this thesis can still be regarded as typical, congruent, and redundant gestures, according to previous research, see for example Cassell et al. (1999); Dargue and Sweller (2018a, 2018b); Dargue et al. (2019); Hostetter (2011); Kelly et al. (2009); McKern et al. (2021); McNeil et al. (2000); McNeill (2005).

Another limitation was that the experiment was online. Participants were asked if they were alone or not, but other than this no further questions regarding the location or setting were asked. Several times in the questionnaire the participants were prompted to be alone, but no control could be done in this regard. Similarly, another limitation is that the participants could open the videos in a new tab on their browser and watch the video as many times as they wanted on YouTube (Youtube, n.d.). However, the results did not indicate that participants did so (only 3 participants got all questions right, these were in the gesture conditions, 2 in the human narrator condition, and 1 in the robot narrator condition).

Furthermore, the questionnaire was programmed so the participants were forced into the narrative questionnaire after being on the questionnaire page that contained the video after 2 minutes and 20 seconds. This ensured participants could not view the narration several times, at least in the questionnaire. Despite the loss of control of the participants' setting, conducting this experiment online ensured that it reached as many participants as possible.

Future research

Similar experiments as conducted in this thesis can be conducted on-site. Similar gestures produced in this thesis, can also be created by motion-capturing and could be played on Epi or similar/different robots. Further testing on Epi or similar robots regarding gesture comprehension, memory recollection, or other skills and subjects can be conducted. For instance, testing gestures' effect with a robot on even longer narratives, within mathematics, problem-solving, or other fields. Testing can also be conducted with non-redundant gestures, incongruent and/or atypical gestures, and further testing of iconic, metaphoric, deictic, and/or beat gestures can be conducted.

The novelty effect and other possible explanations for differences in score should also be studied in future research. Future research can include longer sessions with the robot or some time to get familiar with the robot before the experiments. Once the novelty effect wears off the results might differ and human narrators and robot narrators could potentially be more closely viewed. Nevertheless, future research is needed since the results in this thesis support that humans and robots have similar gestural effects on comprehension and recollection in complex narrations.

6 Conclusion

This thesis has investigated if gestures and verbal communication produced by a human or a robot affect memory recollection and comprehension of a complex narrative story in a similar way. Since robots are an embodied social agent they have a psychical affordance for social interaction. Studies within human-robot interaction have previously shown that robots used in social interactions affect participants similarly to human interaction. This thesis found that gestures, regardless of whether a human or a robot narrator, affect the listeners' comprehension and recollection similarly. Gestures have been shown to improve comprehension and recollection and this thesis and its results are consistent with these previous findings. Co-speech gestures, even if produced by a robot instead of a human, have similar improvements on the listeners' comprehension and recollection. There was a small indication that the robot narrator affected the listener's comprehension and recollection more than the human narrator, and there was a stronger indication regarding the no-gesture-related questions between the human and robot narrator. The indications of participants scoring higher in the robot narrator conditions could be because of a novelty effect or other factors. More studies are needed to further study this effect. In conclusion, this thesis found support that human and robot narrators affected humans' comprehension and recollection in a similar way.

7 Acknowledgement

First and foremost I would like to express my deepest appreciation and gratitude to my supervisor Birger Johansson, for being of great help during this thesis. I am grateful for Birger Johansson's invaluable feedback, guidance, and the several readings of this thesis. Thank you sincerely for all the help! I am also grateful to Annika Wallin for these two excellent years in this program, her always being so helpful, and just supporting us all. I would also like to extend my sincerest thanks to my classmates of HAKOG -21 for two fulfilling years, and their feedback on this thesis, special thanks to Mia Huovilainen for her feedback and comments on this thesis. I am also grateful for my classmate's endless moral support for these two years, and all collaborations during these two years. Thanks to Lund University Cognitive Robotic Lab for letting me use Epi, and thanks to Epi for being such a good sport. Thanks to all my friends, and my family for all their moral support, feedback, sharing the questionnaire, and just being there for me, special thanks to Per Eborn who helped me with the pilot study. Last but not least I would like to extend my heartfelt gratitude to my wife Malin Rosenfrost for all her endless moral support, her belief in me, being there for me during all the mental breakdowns, and for motivating me during this thesis. I would also like to thank my cats for their emotional support and their endless cuteness.

References

Accindi. (n.d.). Retreived in: April, 2023, https://www.accindi.se/.

Aldebaran. (n.d.). Retreived in: May, 2023, https://www.aldebaran.com/sites/default/files/press-kit/NAO-press-kit-EN.pdf.

Alibali, M. W., Heath, D. C., & Myers, H. J. (2001). Effects of visibility between speaker and listener on gesture production: Some gestures are meant to be seen. *Journal of Memory and Language*, *44*(2), 169-188. Retrieved from https://www.sciencedirect.com/science/article/pii/S0749596X00927529 doi: https://doi.org/10.1006/jmla.2000.2752

Alibali, M. W., Kita, S., & Young, A. J. (2000). Gesture and the process of speech production: We think, therefore we gesture. *Language and Cognitive Processes*, *15*(6), 593-613. Retrieved from https://doi.org/10.1080/016909600750040571 doi: 10.1080/016909600750040571

Alibali, M. W., Spencer, R. C., Knox, L., & Kita, S. (2011). Spontaneous gestures influence strategy choices in problem solving. *Psychological science*, *22*(9), 1138–1144.

Aussems, S., & Kita, S. (2019). Seeing iconic gestures while encoding events facilitates children's memory of these events. *Child development*, *90*(4), 1123–1137.

Austin, E. E., & Sweller, N. (2014). Presentation and production: The role of gesture in spatial communication. *Journal of experimental child psychology*, *122*, 92–103.

Austin, E. E., Sweller, N., & Van Bergen, P. (2018). Pointing the way forward: Gesture and adults' recall of route direction information. *Journal of Experimental Psychology: Applied*, *24*(4), 490-508.

Balkenius, C., Morén, J., & Johansson, B. (2007). System-level cognitive modeling with ikaros. *LUCS: Lund University Cognitive Studies*, *133*.

Balkenius, C., Morén, J., Johansson, B., & Johnsson, M. (2010). Ikaros: Building cognitive models for robots. *Advanced Engineering Informatics*, *24*(1), 40–48.

Bartneck, C., Belpaeme, T., Eyssel, F., Kanda, T., Keijsers, M., & Šabanović, S. (2020). *Human-robot interaction: An introduction*. Cambridge University Press.

Baxter, P., Kennedy, J., Senft, E., Lemaignan, S., & Belpaeme, T. (2016). From characterising three years of hri to methodology and reporting recommendations. In *2016 11th acm/ieee international conference on human-robot interaction (hri)* (pp. 391–398).

Belpaeme, T., Kennedy, J., Ramachandran, A., Scassellati, B., & Tanaka, F. (2018). Social robots for education: A review. *Science robotics*, *3*(21), eaat5954.

Bharadwaj, A., Dargue, N., & Sweller, N. (2022). A hands-on approach to learning: Gesture production during encoding and its effect on narrative recall. *Cognitive Science*, *46*(12), e13214.

Bremner, P., & Leonards, U. (2016). Iconic gestures for robot avatars, recognition and integration with speech. *Frontiers in psychology*, *7*, 183.

Bremner, P., Pipe, A., Melhuish, C., Fraser, M., & Subramanian, S. (2009). Conversational gestures in human-robot interaction. In *2009 ieee international conference on systems, man and cybernetics* (pp. 1645–1649).

Bremner, P., Pipe, A. G., Melhuish, C., Fraser, M., & Subramanian, S. (2011). The effects of robot-performed co-verbal gesture on listener behaviour. In *2011 11th ieee-ras international conference on humanoid robots* (pp. 458–465).

Broadbent, E. (2017). Interactions with robots: The truths we reveal about ourselves. *Annual review of psychology*, *68*, 627–652.

Cabibihan, J.-J., So, W.-C., & Pramanik, S. (2012). Human-recognizable robotic gestures. *IEEE Transactions on Autonomous Mental Development*, *4*(4), 305–314.

Cameron, H., & Xu, X. (2011). Representational gesture, pointing gesture, and memory recall of preschool children. *Journal of Nonverbal Behavior*, *35*, 155–171.

Cassell, J., McNeill, D., & McCullough, K.-E. (1999). Speech-gesture mismatches: Evidence for one underlying representation of linguistic and nonlinguistic information. *Pragmatics & cognition*, *7*(1), 1-34.

Church, R. B., Garber, P., & Rogalski, K. (2007). The role of gesture in memory and social communication. *Gesture*, *7*(2), 137–158.

Clark, H. H. (1996). *Using language*. Cambridge university press.

Cook, S. W. (2018). Enhancing learning with hand gestures: Potential mechanisms. In *Psychology of learning and motivation* (Vol. 69, pp. 107–133). Elsevier.

Cook, S. W., & Fenn, K. M. (2017). The function of gesture in learning and memory. , 129-153.

Cook, S. W., Yip, T. K., & Goldin-Meadow, S. (2012). Gestures, but not meaningless movements, lighten working memory load when explaining math. *Language and cognitive processes*, *27*(4), 594–610.

Cutica, I., & Bucciarelli, M. (2008, 07). The deep versus the shallow: Effects of co-speech gestures in learning from discourse. *Cognitive science*, *32*, 921-35. doi: 10.1080/03640210802222039

Dargue, N., & Sweller, N. (2018a). Donald duck's garden: the effects of observing iconic reinforcing and contradictory gestures on narrative comprehension. *Journal of experimental child psychology*, *175*, 96–107.

Dargue, N., & Sweller, N. (2018b). Not all gestures are created equal: the effects of typical and atypical iconic gestures on narrative comprehension. *Journal of Nonverbal Behavior*, *42*, 327–345.

Dargue, N., & Sweller, N. (2020). Two hands and a tale: When gestures benefit adult narrative comprehension. *Learning and Instruction*, *68*, 101331.

Dargue, N., Sweller, N., & Jones, M. P. (2019). When our hands help us understand: A meta-analysis into the effects of gesture on comprehension. *Psychological Bulletin*, *145*(8), 765.

de Wit, J., Vogt, P., & Krahmer, E. (2023). The design and observed effects of robot-performed manual gestures: A systematic review. *ACM Transactions on Human-Robot Interaction*, *12*(1), 1–62.

Fink, J. (2012). Anthropomorphism and human likeness in the design of robots and human-robot interaction. In *Social robotics: 4th international conference, icsr 2012, chengdu, china, october 29-31, 2012. proceedings 4* (pp. 199–208).

Fong, T., Nourbakhsh, I., & Dautenhahn, K. (2003). A survey of socially interactive robots. *Robotics and autonomous systems*, *42*(3-4), 143–166.

Goldin-Meadow, S. (2003). *Hearing gesture: How our hands* help us thinking gesture: How our hands help us think*. Oxford University Press.

Goldin-Meadow, S. (1999). The role of gesture in communication and thinking. *Trends in cognitive sciences*, *3*(11), 419–429.

Goldin-Meadow, S., & Alibali, M. W. (2013). Gesture's role in speaking, learning, and creating language. *Annual review of psychology*, *64*, 257–283.

Goldin-Meadow, S., Nusbaum, H., Kelly, S. D., & Wagner, S. (2001). Explaining math: Gesturing lightens the load. *Psychological science*, *12*(6), 516–522.

Google . (n.d.). Retreived in: Mars, 2023, https://docs.google.com/forms//.

Holler, J., & Wilkin, K. (2011, 06). Co-speech gesture mimicry in the process of collaborative referring during face-to-face dialogue. *Journal of Nonverbal Behavior*, *35*, 133-153. doi: 10.1007/s10919-011-0105-6

Hostetter, A. B. (2011). When do gestures communicate? a meta-analysis. *Psychological bulletin*, *137*(2), 297.

Hostetter, A. B., & Alibali, M. W. (2011). Cognitive skills and gesture–speech redundancy: Formulation difficulty or communicative strategy? *Gesture*, *11*(1), 40–60.

Huang, C.-M., & Mutlu, B. (2013). Modeling and evaluating narrative gestures for humanlike robots. In *Robotics: Science and systems* (Vol. 2).

Johansson, B., Tjøstheim, T. A., & Balkenius, C. (2020). Epi: An open humanoid platform for developmental robotics. *International Journal of Advanced Robotic Systems*, *17*(2), 1729881420911498.

Kartalkanat, H., & Göksun, T. (2020). The effects of observing different gestures during storytelling on the recall of path and event information in 5-year-olds and adults. *Journal of experimental child psychology*, *189*, 104725.

Kelly, S. D., Barr, D. J., Church, R. B., & Lynch, K. (1999). Offering a hand to pragmatic understanding: The role of speech and gesture in comprehension and memory. *Journal of memory and Language*, *40*(4), 577–592.

Kelly, S. D., Kravitz, C., & Hopkins, M. (2004). Neural correlates of bimodal speech and gesture comprehension. *Brain and language*, *89*(1), 253–260.

Kelly, S. D., McDevitt, T., & Esch, M. (2009). Brief training with co-speech gesture lends a hand to word learning in a foreign language. *Language and cognitive processes*, *24*(2), 313–334.

Kelly, S. D., Özyürek, A., & Maris, E. (2010). Two sides of the same coin: Speech and gesture mutually interact to enhance comprehension. *Psychological science*, *21*(2), 260–267.

Kendon, A. (1980). Gesticulation and speech: Two aspects of the process of utterance. *The relationship of verbal and nonverbal communication*, *25*(1980), 207–227.

Kendon, A. (1994). Do gestures communicate? a review. *Research on Language and Social Interaction*, *27*(3), 175-200. Retrieved from https://doi.org/10.1207/s15327973rlsi2703_2 doi: 10.1207/s15327973rlsi2703_2

Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge University Press.

Kennedy, J., Baxter, P., & Belpaeme, T. (2015a). Can less be more? the impact of robot social behaviour on human learning. In *Proceedings of the 4th international symposium on new frontiers in hri at aisb*.

Kennedy, J., Baxter, P., & Belpaeme, T. (2015b). Comparing robot embodiments in a guided discovery learning interaction with children. *International Journal of Social Robotics*, *7*, 293–308.

Kennedy, J., Baxter, P., Senft, E., & Belpaeme, T. (2016). Social robot tutoring for child second language learning. In *Proceedings of the 11th acm/ieee international conference on human-robot interaction (hri)* (pp. 231–238).

Macedonia, M., & Knösche, T. R. (2011). Body in mind: How gestures empower foreign language learning. *Mind, Brain, and Education*, *5*(4), 196–211.

Macoun, A., & Sweller, N. (2016). Listening and watching: The effects of observing gesture on preschoolers' narrative comprehension. *Cognitive Development*, *40*, 68–81.

McKern, N., Dargue, N., Sweller, N., Sekine, K., & Austin, E. (2021). Lending a hand to storytelling: Gesture's effects on narrative comprehension moderated by task difficulty and cognitive ability. *Quarterly Journal of Experimental Psychology*, *74*(10), 1791–1805.

McNeil, N. M., Alibali, M. W., & Evans, J. L. (2000). The role of gesture in children's comprehension of spoken language: Now they need it, now they don't. *Journal of Nonverbal Behavior*, *24*, 131–150.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. University of Chicago Press, Chicago.

McNeill, D. (2005). *Gesture and thought*. University of Chicago Press. http://dx. doi. org/10.7208/chicago.

Meta Platforms. (n.d.). Retreived in: Mars, 2023, https://www.Facebook.com/.

Michaelis, J. E., & Di Canio, D. (2022). Embodied geometric reasoning with a robot: The impact of robot gestures on student reasoning about geometrical conjectures. In *Proceedings of the 2022 chi conference on human factors in computing systems* (pp. 1–14).

Norman, D. (2013). *The design of everyday things: Revised and expanded edition*. Basic books.

Novack, M. A., Wakefield, E. M., & Goldin-Meadow, S. (2016). What makes a movement a gesture? *Cognition*, *146*, 339-348. Retrieved from https://www.sciencedirect.com/science/article/pii/S0010027715300925 doi: https://doi.org/10.1016/j.cognition.2015.10.014

Overoye, A. L., & Storm, B. C. (2019). Remembering what was said and done: The activation and facilitation of memory for gesture as a consequence of retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(3), 526.

R Core Team. (2021). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from https://www.R-project.org/

Robins, B., Dautenhahn, K., te Boekhorst, R., & Billard, A. (2003, 01). *Effect of repeated exposure of a humanoid robot on children with autism-can we encourage basic social interaction skills* (Tech. Rep.).

RStudio Team. (2020). Rstudio: Integrated development environment for r [Computer software manual]. Boston, MA. Retrieved from http://www.rstudio.com/

Saunderson, S., & Nejat, G. (2019). How robots influence humans: A survey of nonverbal communication in social human–robot interaction. *International Journal of Social Robotics*, *11*, 575–608.

Stiefelhagen, R., Fugen, C., Gieselmann, R., Holzapfel, H., Nickel, K., & Waibel, A. (2004). Natural human-robot interaction using speech, head pose and gestures. In *2004 ieee/rsj international conference on intelligent robots and systems (iros)(ieee cat. no. 04ch37566)* (Vol. 3, pp. 2422–2427).

Stoet, G. (2010). Psytoolkit: A software package for programming psychological experiments using linux. *Behavior research methods*, *42*, 1096–1104.

Stoet, G. (2017). Psytoolkit: A novel web-based method for running online questionnaires and reaction-time experiments. *Teaching of Psychology*, *44*(1), 24–31.

Stolzenwald, J., & Bremner, P. (2017). Gesture mimicry in social human-robot interaction. In *2017 26th ieee international symposium on robot and human interactive communication (ro-man)* (pp. 430–436).

SurveyCircle. (n.d.). Retreived in: Mars, 2023, `https://www.surveycircle.com/en/`.

Valenzeno, L., Alibali, M. W., & Klatzky, R. (2003). Teachers' gestures facilitate students' learning: A lesson in symmetry. *Contemporary Educational Psychology*, *28*(2), 187–204.

Van den Berghe, R., Verhagen, J., Oudgenoeg-Paz, O., Van der Ven, S., & Leseman, P. (2019). Social robots for language learning: A review. *Review of Educational Research*, *89*(2), 259–295.

Van Dijk, E. T., Torta, E., & Cuijpers, R. H. (2013). Effects of eye contact and iconic gestures on message retention in human-robot interaction. *International Journal of Social Robotics*, *5*, 491–501.

Vogt, P., van den Berghe, R., De Haas, M., Hoffman, L., Kanero, J., Mamus, E., … others (2019). Second language tutoring using social robots: a large-scale study. In *2019 14th acm/ieee international conference on human-robot interaction (hri)* (pp. 497–505).

Wagner, S. M., Nusbaum, H., & Goldin-Meadow, S. (2004). Probing the mental representation of gesture: Is handwaving spatial? *Journal of memory and language*, *50*(4), 395–407.

Youtube. (n.d.). Retreived in: May, 2023, `https://www.youtube.com/`.

8 Appendix

Video URL

Video	URL
Video 1	`https://youtu.be/cNxqyicxYrM`
Video 2	`https://youtu.be/g9xIHVXxUF4`
Video 3	`https://youtu.be/32kLo-Hoo8M`
Video 4	`https://youtu.be/JjPg4YCaUQc`
Comparision	`https://youtu.be/AIr6jgp8oQ4`

Table 8. Video URLs

8.1 Script (Swedish)

(bold) indicate gestures.

Hej! **(Vinka)** Jag ska berätta en historia. Det är viktigt att du uppmärksammar vad som sägs.

Sara klär på sig och tittar på sin klocka **(titta på klockan)** och inser att hon är sen till jobbet idag. Hon gör iordning frukosten och kastar i sig flingorna **(kasta i sig mat)**. Hon skyndar sig sen genom hallen, hon smäller upp dörren **(smälla upp dörr)**, och låser den efter sig. Hon springer **(springa)** efter bussen som precis ska åka iväg. Hon forcerar **(forcerar)** upp dörrarna och lyckas gå på bussen. Busschauffören skakar på huvudet **(nej)**, men låter henne ändå åka med bussen. Sara sätter sig ner och tittar ut genom fönstret, där hon ser några barn leka med en boll **(boll)**. Tillslut närmar sig bussen Saras kontorsbyggnad. Så fort bussen stannar så börjar hon springa mot sitt jobb. När hon kommer fram så vinkar **(vinka)** hon mot receptionisten utan att märka att disken är tom. Efter att hon har gått in i hissen så klickar Sara frenetiskt på hissknappen, men tappar sin mobil, hon plockar upp den **(plocka upp)** och rycker till när hon ser vad klockan är. Så fort hissdörrarna öppnas så springer **(springa)** Sara igenom korridoren, det tar flera minuter innan hon inser att hon är på fel våning. Sara vänder sig om så snabbt att hon nästan trillar **(trilla)**. Hon tänker att det går snabbare att ta trapphuset så hon skyndar sig dit. På vägen hoppar hon över **(hoppa över)** en dammsugare som står på golvet. Hon öppnar dörren till trapphuset och tar två trappsteg i taget uppåt **(upp)**. När hon når rätt våning märker hon att ingen annan är där. Hon hinner inte tänka på det, utan fortsätter till sitt rum. När hon kommit fram så tänder hon lampan, och går runt **(gå runt)** skrivbordet. Med en duns sätter hon sig ner **(ner)** i sin stol. Hon torkar svetten ur pannan **(torka svett ur panna)** och vänder uppmärksamheten mot sin dator. Så fort den startat öppnar hon upp alla program. Hon har varken fått mail eller meddelanden. Hon tycker att det är konstigt och tar upp sin telefon **(ta upp telefon)**. Till sin stora förvåning ser Sara att hon tagit fel på dagen. Det är ju lördag.

8.2 Questionnaire (Swedish)

8.2.1 Narrative-questionnaire

(NG) = No-gesture related questions
(G)= Gestures related questions

1. Vad hette huvudpersonen i berättelsen? **(NG)**
2. Vad åt huvudpersonen i berättelsen? **(G)**
3. Hur öppnade huvudpersonen sin dörr? **(G)**
4. Låste huvudpersonen sin dörr? **(NG)**
5. Hur åkte huvudpersonen till jobbet? **(NG)**
6. Hur tog sig huvudpersonen in i bussen? **(G)**
7. Hur reagerade busschauffören när huvudpersonen gick på bussen? **(G)**
8. Vad såg huvudpersonen i bussens fönster? **(NG)**
9. Vad lekte barnen med? **(G)**
10. Hur hälsade huvudpersonen på receptionisten? **(G)**
11. Var det någon i receptionen? **(NG)**
12. Vad tappade huvudpersonen i hissen? **(NG)**
13. Vad gjorde huvudpersonen så fort hon kommer ut ur hissen? **(G)**
14. Vad var det som fick huvudpersonen att nästan trilla? **(G)**
15. Vad hoppade huvudpersonen över i korridoren? **(G)**
16. I vilken riktning rörde sig huvudpersonen i trapphuset? **(G)**
17. Vad gjorde huvudpersonen innan hon satte sig ner i sitt rum? ((G))
18. Hur satte sig huvudpersonen ner? **(G)**
19. Vad gjorde huvudpersonen först efter att hon satt sig ner? **(G)**
20. Vad märkte huvudpersonen när hon tittade på datorn? **(NG)**
21. Vilken dag var det? **(NG)**

8.2.2 Perception-questionnaire

- Hur uppfattade du rörelserna?

  Jag uppfattade rörelsena som roboten utförde som (skala 1-5):

  1. Väldigt onaturliga
  2. Ganska onaturliga
  3. Varken eller
  4. Ganska naturliga
  5. Väldigt naturliga

- Hur uppfattade du rörelserna?

  Jag uppfattade rörelsena som människan utförde som (skala 1-5):

  1. Väldigt onaturliga
  2. Ganska onaturliga
  3. Varken eller
  4. Ganska naturliga
  5. Väldigt naturliga

- Tyckte du att någon av rörelserna som roboten gjorde verkade onaturliga? I såna fall vilka?

- Tyckte du att någon av rörelserna som människan gjorde verkade onaturliga? I såna fall vilka?
- Verkade både roboten och människan utföra samma rörelser? Skriv ner dina tankar oavsett vad du tycker.
- Märktes det att roboten inte kunde röra på sig på samma sätt som människan?

8.3 Script (English, translated from the Swedish story)

Words in **(bold)** indicate gestures.

Hello! **(Waving)** I will tell you a story. It is important that you pay attention to what is being said.

Sara gets dressed and looks at her watch **(look at her watch)** and realizes that she is late for work today. She prepares her breakfast and eats her cereal fast **(eating cereal)**. She then hurries through the hall, she slams open the door **(slam open door)**, and locks it behind her. She runs **(run)** after the bus that is about to leave. She forces **(forces)** the doors open and manages to get on the bus. The bus driver shakes his head **(no)**, but still let her ride the bus. Sara sits down and looks out the window, where she sees some children playing with a ball **(ball)**. Finally, the bus approaches Sara's office building. As soon as the bus stops, she starts running towards her job. When she arrives, she waves **(wave)** at the receptionist without noticing that the counter is empty. After she enters the elevator, Sara frantically clicks the elevator button, but drops her cell phone, she picks it up **(picks up)** and flinches when she sees what time it is. As soon as the elevator doors open, **(run)** Sara runs through the corridor, it takes several minutes before she realizes that she is on the wrong floor. Sara turns around so quickly that she almost falls **(falls)**. She thinks it's faster to take the stairwell, so she rushes there. On the way, she jumps over **(jump over)** a vacuum cleaner that is on the floor. She opens the door to the stairwell and takes two steps at a time upwards **(up)**. When she reaches the right floor, she notices that no one else is there. She doesn't have time to think about it, but continues to her room. When she arrives, she turns on the lamp, and walks around **(walk around)** the desk. With a thud, she sits down **(down)** in her chair. She wipes the sweat from her forehead **(wipe sweat from her forehead)** and turns her attention to her computer. As soon as it starts, she opens all programs. She hasn't received any emails or messages. She finds it strange and picks up her phone **(pick up phone)**. Much to her surprise, Sara sees that she got the day wrong. It's Saturday.

8.4 Questionnaire (English)

8.4.1 Narrative-questionnaire

(NG) = No-gesture related questions
(G)= Gestures related questions

1. What was the name of the main character in the story **(NG)**
2. What did the main character in the story eat? **(G)**
3. How did the main character open their door? **(G)**
4. Did the main character lock their door? **(NG)**
5. How did the main character get to work? **(NG)**
6. How did the main character get on the bus? **(G)**
7. How did the bus driver react when the main character got on the bus? **(G)**
8. What did the main character see in the window of the bus? **(NG)**
9. What did the children play with? **(G)**
10. How did the main character greet the receptionist? **(G)**
11. Was there anyone at the reception? **(NG)**
12. What did the main character drop in the elevator? **(NG)**
13. What did the main character do as soon as she got out of the elevator? **(G)**
14. What made the main character almost fall ? **(G)**
15. What did the main character jump over in the hallway? **(G)**
16. In which direction did the main character move in the stairwell? **(G)**
17. What did the main character do before she sat down in her room? ((G))
18. How did the main character sit down? **(G)**
19. What did the main character do first after she sat down? **(G)**
20. What did the main character notice when she looked at the computer? **(NG)**
21. What day was it? **(NG)**

8.4.2 Perception-questionnaire

- How did you perceive the movements?

  I perceived the movements performed by the robot as (scale 1-5):

  1. Very unnatural
  2. Quite unnatural
  3. Neither natural or unnatural
  4. Quite natural
  5. Very natural
- How did you perceive the movements?

  I perceived the movements that the human performed as (scale 1-5:

  1. Very unnatural
  2. Quite unnatural
  3. Neither natural or unnatural
  4. Quite natural
  5. Very natural
- Did you think any of the movements the robot made seemed unnatural? If so, which ones?
- Did you think any of the movements the human made seemed unnatural? If so, which ones?

- Did both the robot and the human seem to perform the same movements? Write down your thoughts regardless of what you think.
- Was it noticeable that the robot could not move in the same way as the human?