Evaluating the Effect of Language Fluency and Task Competency on the Perception of a Social Robot

A

Thesis

Presented to

The faculty of the School of Engineering and Applied Science

University of Virginia

in partial fulfillment
of the requirements for the degree

Master of Science
in
Systems Engineering

Department of Engineering Systems and Environment
University of Virginia

by

Shahira Ali

May 2023

APPROVAL SHEET

This

Thesis

is submitted in partial fulfillment of requirements

for the degree of

Master of Science

Author: Shahira Ali

This Thesis has been read and approved by the examining committee:

Advisor: Tariq Iqbal

Committee Member: Tariq Iqbal (Advisor), Assistant Professor, Systems and Information Engineering

Committee Member: Afsaneh Doryab (Chair), Assistant Professor, Systems and Information Engineering

Committee Member: Reneé Cummings, Assistant Professor of the Practice, Data Science

Accepted for the School of Engineering and Applied Science:

Jennifer L. West, School of Engineering and Applied Science

May 2023

ABSTRACT

Recent advancements in robot capabilities have enabled them to interact with people in various human-social environments (HSEs). In many of these environments, the perception of the robot often depends on its capabilities, e.g., task competency, language fluency, etc. To enable fluent human-robot interaction (HRI) in HSEs, it is crucial to understand the impact of these capabilities on the perception of the robot. Although many works have investigated the effects of various robot capabilities on the human's perception of the robot separately, in my M.S. research, I present three large-scale HRI study (total n = 120) to investigate the combined impact of both language fluency and task competency on the perception of a robot.

For the first study, we enlisted monolingual participants who only spoke English (n = 60) in order to control for the participants' linguistic background. Then, to establish a social interaction while also incorporating a task to complete between the human and the robot, we designed a simple guessing game modeled after the children's game called 'What Am I?' for the participant and robot to play together where the robot assumed the identity of an animal and gave the participant hints to guess the animal. To accommodate the differences in language proficiency and task competence of the robot, we equipped it with the ability to be both fluent and disfluent in English and correctly match or fail to identify the animal, respectively. Finally, we developed four distinct combinations of language fluency and task competency (fluent-competent, fluent-incompetent, disfluent-competent, and disfluent-incompetent) which made up the varying robot conditions in the study. Participants were randomly placed into one of these four groups, and to understand how the robot's condition impacts human perception of the robot, we collected and analyzed participants' perceptions of the robot's verbal competence, intelligence, and reliability along with the robot meeting expectations, being a good teammate, and if the participant is willing to work with the robot again.

From the monolingual perspective in Study I, the fluent-competent robot was rated higher than the disfluent-incompetent robot in every perception category except willingness to work again. The results indicate that participants found both the fluent robots (fluent-competent and fluent-

incompetent) to be significantly more verbally competent than the two disfluent robots (disfluent-competent and disfluent-incompetent). Participants also perceived the fluent-competent, fluent-incompetent, and disfluent-competent robots to be more intelligent than the disfluent-incompetent robot. Participants in the fluent-competent, fluent-incompetent, and disfluent-competent conditions all perceived the robot to be significantly more reliable than how the participants in the disfluent-incompetent condition perceived the robots to be. Regarding perceptions of expectations being met, participants in the fluent-competent and disfluent-competent conditions perceived the robot to have met their expectations more than the participants in the disfluent-incompetent condition did. Lastly, participants in the fluent-competent and disfluent-competent conditions rated the robot more highly for being a good teammate than participants in the disfluent-incompetent condition did. There was no statistical significance of the effects of varying robot conditions on participants' willingness to work with the robot again.

In the second study, we aimed to maintain the same experimental setup and design as the first study, but instead of requiring participants to be monolingual in English, we sought out participants ($n = 60$) who were multilingual, with English as one of their fluently spoken languages. The results of this study suggest that participants found the fluent-competent robot to be significantly more verbally competent than the other three varying robot conditions (fluent-incompetent, disfluent-competent, and disfluent-incompetent). Participants perceived both the competent robots (fluent-competent and disfluent-competent) to be more intelligent than the incompetent robots (fluent-incompetent and disfluent-incompetent robots). The results also indicate that both the competent robots (fluent-competent and disfluent-competent) were perceived to be more reliable than the incompetent robots (fluent-incompetent and disfluent-incompetent robots). Additionally, participants rated the fluent-competent robot significantly higher in reliability than the disfluent-competent robot. The participants also perceived both the competent robots (fluent-competent and disfluent-competent) to have met their expectations more than the incompetent robots (fluent-incompetent and disfluent-incompetent robots). Participants in the fluent-competent and disfluent-competent conditions rated the robot significantly higher for being a good teammate than participants in the fluent-incompetent and disfluent-incompetent conditions did. Finally, the results suggest that the participants are more willing to work with both the competent robots (fluent-competent and disfluent-competent) than the incompetent robots (fluent-incompetent and disfluent-incompetent robots).

The third study entailed comparing the perception of robots from the monolingual perspective with the perception of robots from the multilingual perspective, and this was done by combining the data from both Study I and Study II (n = 120). The purpose of this third study was to investigate potential interactions between language groups and varying task conditions on perceptions of the robot. We found that there were statistically significant differences between the ratings of monolingual participants and that of multilingual participants in certain varying robot conditions. Multilingual participants perceived the verbal competence of the disfluent-competent and disfluent-incompetent robot to be significantly higher than monolingual participants did. Multilingual participants also rated the fluent-competent robot more highly as a good teammate than their monolingual counterparts. However, the monolingual participant group rated the fluent-incompetent robot significantly higher than the multilingual participants in intelligence, reliability, meeting expectations, and willingness to work with the robot again.

The results suggest that both language fluency and task competency may impact certain perceptions of the robot at different scales. For example, in Study I, while language fluency may play a more significant role than task competency in the monolingual perception of the verbal competence of a robot, both language fluency and task competency contribute to the perception of the intelligence and reliability of the robot. On the other hand, task competency played a more significant role than language fluency in the perception of meeting expectations and being a good teammate for monolingual individuals. In Study II, multilingual participants were more impacted by task competency in their perceptions of the robot's intelligence, ability to be a good teammate, meet expectations, and their willingness to work with the robot again. In Study III, monolingual participants prioritized fluency more than task competency in the perception categories of intelligence, reliability, meeting expectations, and willingness to work with the robot again as they rated the fluent-incompetent robot significantly higher in these conditions than multilingual participants did. Although the studies in our research did not investigate why monolingual or multilingual perceptions may have varied across certain conditions, these results may serve as a foundation for future research to explore additional relationships between monolingual and multilingual participants' perception of robots while also understanding the factors that may cause these perceptions to differ. Overall, the findings of these three studies highlight the relationship between language fluency and task competency through the lens of linguistic background in the context of social HRI and will enable the development of more intelligent robots in the future.

Table of Contents

Chapter 1

INTRODUCTION

The integration of robots to interact and work alongside people has become a growing area of research [1]–[7]. Advancements in robotics over the last few decades have enabled them to perform complex tasks and interact with people across various human-social environments (HSEs) [8]– [12]. Other works have explored how robots can understand and utilize various group dynamics to work with people in teams [99]-[103]. As a result, there has been an increase in research exploring the role of robots and understanding how robots are perceived in various scenarios based on these roles [13]– [17].

One particular capability of robots that has widely been explored in the context of social human-robot interaction (HRI) is their verbal communication skills [18]– [23]. Along this research direction, many works have focused on investigating the communication abilities of robots, and their effectiveness in various task settings [7], [15], [19], [24]. Another work evaluated interpersonal communication and relationship building between a robot and a human partner and found that verbal communication capabilities had a positive impact on the social perception of the robot as well as strengthened the bond between the collaborative pair [24]. Similarly, another study found that verbal communication positively impacts perceptions of friendliness and social presence in robots [25]. Given the critical role verbal communication plays in HRI, many works in HRI have focused on making communication as natural and fluent as possible [19], [20].

In human-human interactions, among various aspects of verbal communication, language fluency particularly plays a significant role in building a perception of a person. For example, some studies indicate that participants hold negative biases and stereotypes against non-native speakers [26]– [28]. Although the impact of language fluency on the perception of a person in human-human interaction has been considerably studied, there has not been any work exploring the impact of language fluency on the perception of the robot. As robots are now expected to interact with people in various social scenarios, it's crucial to explore how verbal communication, specifically language fluency, may play a role in HRI.

Along with verbal communication being a significant part of interactions, robots are also expected to interact with people and perform various tasks with them. In these scenarios, the task competency of the robot also affects how people perceive the robot. Several studies have investigated varying task competency and its influence on the perception of robots [29]– [35]. For example, robots that were interruptive during tasks were perceived as being less task competent than the robots that completed the tasks without interleaving [32], [35]. Additionally, results from some works indicate how human agents perceived an erroneous robot as less trustworthy and less reliable than a competent robot in a collaborative setting [36].

1.1 Thesis Statement

Verbal communication within human-robot interactions has been increasing as robots are being designed to better work with humans and integrate into society [18]– [20], [22], [50]-[52]. Similarly, there has been a significant exploration of human-robot interaction through task collaboration along with implications on perceptions of a competent robot [35], [36], [55]. Thus, social robots employing natural, fluent speech and being competent at tasks are established as separate, important components in the perception of robots by humans. However, given that a lot of task collaboration between humans and robots includes verbal communication, it is surprising that both components have not yet been jointly evaluated within the realm of HRI. Although task incompetency in robots has been investigated, there is no research surrounding disfluency within robots. Moreover, existing literature depicts how monolingual and multilingual individuals perceive non-native speech differently and how there are negative biases and perceptions present against disfluent speech in humans [26], [27], [41]–[45], [82], [83]. Therefore, it is crucial to investigate how varying language fluency and task competency in a social robot will affect people's perceptions of the robot through the lens of their linguistic background. Specifically, I aim to investigate the impact of the robot's language fluency and task competency (i.e., fluent-competent, fluent-incompetent, disfluent- competent, and disfluent-incompetent conditions) on the perceptions of the robot's verbal competence, intelligence, and reliability along with whether the robot meets expectations, is a good teammate, and if the participant is willing to work with the robot again. Understanding the perceptions that people with different language capabilities can have on a robot with varying language fluency and task competency will allow for robots to be designed more appropriately to integrate better into diverse societies.

1.2 Completed Work

To investigate the gap introduced in Section 1.1, we designed a human-robot interaction scenario modeled after the children's guessing game, "What Am I?," and conducted three large-scale HRI studies (n = 60, n=60, n=120). In this task, a NAO robot described various characteristics of an animal, and the human participant was prompted to guess what animal the robot was referring to. We varied the language fluency of the robot by allowing it to speak at a native or a non-native level and the task competency by allowing the robot to correctly match or fail to identify the animal. The first study explored the perceptions of the robot from a monolingual perspective, the second study focused on perceptions of the robot from a multilingual perspective, and the third study compared the perceptions of the two participant groups.

The results from the first study suggest that while language fluency may play a greater role than task competency in the perception of verbal competence of a robot, both language fluency and task competency contribute to the perception of intelligence and reliability of the robot. The results also indicate that task competency may play a greater role than language fluency in the perception of the robot meeting expectations and being a good teammate. In the second study, multilingual participants may have been more impacted by task competency in their perceptions of the robot's intelligence, ability to be a good teammate, meet expectations, and their willingness to work with the robot again while both language fluency and task competency may play a role in their perception of the robot's reliability. The findings in the third study suggest that there were significant interactions between the participant groups and their perceptions of the robot in certain conditions. The results indicate that multilingual individuals may perceive language disfluency more positively with respect to ratings of verbal competence than monolingual participants. Monolingual individuals, on the other hand, perceived fluent-incompetent robots more positively in ratings of intelligence, reliability, meeting expectations, and willingness to work with the robot again. This may suggest that monolingual and multilingual individuals' perceptions of robots are impacted differently by varying language fluency and task competency. These findings will allow us to have a deeper understanding of the relationship between language fluency and task competency on the perception of a robot and to develop ways to reduce non-native language biases and promote greater understanding and respect for linguistic diversity in the context of HRI.

1.3 Contributions

My thesis has three main components: perceptions of robots with varying language fluency and task competency from a monolingual perspective, multilingual perspective, and a comparison between the perceptions of the two participant groups. First, we developed a study design in which the robot presented different combinations of language fluency and task competency during a guessing game that it played with participants. In Study I, we found that there is a significant difference between perceptions of robots in varying robot conditions in certain perception categories from the perspective of monolingual individuals. For example, monolinguals may prioritize language fluency when perceiving verbal competence, both language fluency and task competency when perceiving intelligence and reliability, and task competency when determining if their expectations were met and if the robot was a good teammate. Next, we replicated the study design from Study I to create Study II, but we replaced monolingual participants with multilingual participants in this second study. We discovered that there is a significant difference between perceptions of robots in varying robot conditions in certain perception categories from the perspective of multilingual individuals. Specifically, multilinguals may be impacted by task competency more than language fluency in their perceptions of the robot's intelligence, ability to be a good teammate, meet expectations, and their willingness to work with the robot again while both language fluency and task competency may play a role in their perception of the robot's reliability. Third, in Study III, we analyzed the combined results of Study I and Study II. We observed that there were significant interactions between the participant groups and their perceptions of the robot in certain conditions. These findings implied that language fluency plays a greater role in the perception of a disfluent robot's verbal competence for monolingual participants than multilingual participants. Additionally, monolingual individuals rated fluent-incompetent robots more favorably in the perception categories of intelligence, reliability, meeting expectations, and willingness to work with the robot again. Based on these works, we hope for further research to expand on understanding why ratings between robot conditions and participant groups differed while also gaining valuable insight into what robot capabilities are necessary to maintain positive human-robot interactions.

1.4 Broader Impact

Our proposed studies and findings have a vast range of pertinence within the field of human-robot interaction. Although the impacts of verbal communication and task competency on perceptions of social robots have been investigated in isolation, our findings demonstrate a greater need for them to be explored together. In addition, given that social robots will have to work with diverse human populations who potentially identify with many different backgrounds, our research has opened the door to understanding the impact that a person's linguistic background can have on the perception of the robot. Moreover, these studies can inspire future work to investigate or further explore the potential impact that other participant backgrounds can have on perceptions of robots such as culture, personality, age, and gender.

Additionally, our results can be extended to make interactions between humans and social robots more natural and seamless since we are better able to understand the preferences in robot capabilities that different subset groups in human populations may have. Furthermore, our proposed implications of how varying language fluency and task competency in a social robot affects different participant groups' perceptions of the robot enable future robot design to prioritize which capabilities the robot needs to possess in order to best serve or work with the specific person or group of persons that it will be working with. For example, if multilingual individuals find task competency to play a greater role than language fluency in their perception of the robot partner's intelligence, then it would be beneficial to make sure that the robot has uncompromised task competency and not focus more on its language fluency when designing it to interact with multilingual humans. Finally, we believe that our research paves a way for the research community to evaluate perceived language fluency and task competency in robots through different linguistic backgrounds and make advancements on this research field in HRI.

Chapter 2

BACKGROUND

2.1 Language Fluency

The definition of fluency in the context of spoken language in human-human interactions has long been debated and studied. For example, Fillmore sees fluency as “the ability to fill time with talk” which implies that a person employs fewer pauses, stutters, and thinking time when speaking [63]. Brumfit perceives fluency “as the maximally effective operation of the language system so far acquired by the student” [64] Both of these approaches were combined to propose a potential broad measure of fluency to be “(1) the speed and flow of language production, (2) the degree of control of language items, and (3) the way language and content interact” [65]. However, Swain believes that fluency is when a message is conveyed “precisely, coherently, and appropriately.” Thus, in order to reach native-speaker fluency, there has to be an advancement from semantic to syntactic processing, so that the grammatical accuracy of the language is maintained in the delivery of the message [66].

Lennon differentiates between two approaches of fluency [37]. The first is a holistic, broad sense in which fluency refers to global oral proficiency. The second portrays fluency in a much narrower sense and relies on the efficient and effortless planning and production of speech, such as speaking speed or grammatical accuracy. Furthermore, language fluency has been linked to higher levels of perceived proficiency and is “best conceived of as fast, smooth, and accurate performance” [38], [39]. However, there is a greater risk of confounding measures when multiple measures are used to define and examine fluency [40]. Given the limited literature on definitive descriptions and classifications of language fluency within the human-human realm, it is not surprising that language fluency in HRI has yet to be adequately explored. In this study, we will be referring to language fluency in Lennon’s narrow sense and focus only on one measure of oral proficiency: grammatical accuracy.

2.1.2 Defining Monolingualism & Multilingualism

As technology and society advances, interactions between groups of people with different identities, backgrounds, and cultures are also bound to increase. With people being exposed to more languages and having varying competency levels in producing those languages, it has become more of a challenge to try and adapt a person's linguistic abilities to existing concepts of what is considered monolingual, bilingual, multilingual, and so on [68]. There are various definitions of monolingual, bilingual, and multilingual, but the *Longman Dictionary of Language Teaching and Applied Linguistics* and the *Merriam-Webster Dictionary* define them as follows:

> '*monolingual*'
>
> (n, adj.) "1. a person who knows and uses only one language 2. a person who has an active knowledge of only one language, though perhaps a passive knowledge of others." [69]
>
> (adj.) "having or using only one language." [70]
>
> *'bilingual'*
>
> (adj.) "a person who uses at least two languages with some degree of proficiency. In everyday use bilingual usually means a person who speaks, reads or understands two languages equally well (a balanced bilingual), but a bilingual person usually has a better knowledge of one language than another." [69]
>
> (adj.) "using or able to use two languages especially with equal fluency." [70]
>
> *'multilingual'*
>
> (n, adj.) "a person who knows and uses three or more languages. Usually, a multilingual does not know all the languages equally well." [69]
>
> (adj.) "using or able to use several languages especially with equal fluency." [70]

Pulling from other existing literature, a working definition of monolingualism can be "the condition or state of being able to speak only one language," or natively speaking one language, bilingualism to be "a native-like control of two languages," and multilingualism to be equal competence in more than one language or "the use of more than one language" [74]-[76]. For the sake of my research, the definition of monolingual that I adhere to is "native fluency in one language." Additionally, in my thesis, I interpret the terms 'multilingual' and 'multilingualism' to encompass 'bilingual' and 'bilingualism,' respectively. Therefore, the definition of multilingual that I employ in my research is "native fluency in more than one language." To control for the language capabilities of participants in this work, we required all participants to be a native, monolingual in English in Study I and multilingual, with one of the fluently spoken languages being English, in Study II.

2.2 Language Bias & Benefits

Monolingualism and multilingualism span beyond just linguistic capabilities and implications. There is literature supporting that multilingual individuals may have intellectual, emotional, and cultural benefits [68], [78]. Language study can be seen as an intellectual stimulus that can assist with cognitive processes and yield new ways of "thinking, learning, and organizing knowledge" [79]. Hawkins stated that the ability to learn one language makes an individual capable of learning other languages, as there are important linguistic skills and learning strategies that are developed through this language acquisition [80]. Knowing other languages also provides 'access to different bodies of knowledge which are unavailable to the monolingual speaker' [81]. It was also found that multilingualism had a correlation with greater cultural empathy and open-mindedness [71], [72]. Dewaele suggested that the implication of this was that knowing many languages "broadens the mind" and multilinguals are "more willing to accept that other people might have different values" [71]. There is also research supporting that individuals who know more languages express lower levels of communicative anxiety, and this finding was attributed "to the fact that multilinguals have more experience in communication with a wide range of interlocutors, which allows them to overcome unexpected communicative difficulties" [71], [72]. Relatively recent research has found that multilingualism, which is "operationalized as advanced levels of proficiency in several foreign languages and frequent use of these languages," is positively linked to cognitive empathy [77]. Furthermore, these greater levels of cognitive empathy can be perceived as an indication of multicompetence [77].

However, not all the associations with multilingualism are positive. The study of language bias and perception has become increasingly important as globalization has led to a more diverse and multicultural society. There are negative attitudes and biases towards people who speak many languages throughout society for various reasons. An area of interest is the perception of multilingual, non-native language speakers by monolingual, native language speakers. For example, several studies have been conducted to investigate the extent of bias, and the factors that influence perceptions of non-native speakers [26], [27], [41]– [45]. Other studies depict how students express a preference for native English-speaking teachers and negatively judge the teaching skills of non-native English teachers [28], [46], [47]. Similarly, monolingual children displayed a preference to affiliate and work with other monolingual children who spoke the same language as them than with bilingual children or foreign language speakers [84]. One study found non-native speakers to have lower perceived intelligence and trustworthiness by native speakers [48]. Volz et al. found that monolingual, native English speakers were less likely to believe nonnative speakers' messages, and this was in part attributed to the language disfluency in the messages [82]. Another study observed that non-native speakers in call centers were perceived as less verbally competent and trustworthy compared to their native English-speaking counterparts [83]. There is adequate literature surrounding native speaker bias and perceptions of non-native speakers in human-human interactions. However, little research has been done surrounding the impact of a person's linguistic abilities on their perceptions of a robot. Therefore, in this research, we take into account the languages that participants speak, and explore the impact that this linguistic background can have on perceptions of a robot with varying language fluency.

2.3 Language in Robots

Verbal communication has been a crucial area of research in robotics, as it enables robots to interact with humans in a more natural and intuitive way [18]– [20], [22]. Language plays an important role in human-robot interaction, as it allows robots to understand and respond to human commands, convey information, and establish social relationships with humans [50]– [52]. McGinn found that the robot's voice which with it executed verbal communication had an impact on the robot's perception [106]. The focus of several works has been to improve perceptions of trust in robots by employing deep learning speech and automatically generated explanations [53], [54]. With the importance of verbal communication and natural language established in HRI, it is important to understand how varying levels of language fluencies can impact perceptions of that robot.

2.4 Task Competency

Previous research in HRI has explored various aspects of robot behavior, including task performance, and their impact on the perception of the robot. For example, Salem et al. found that participants rated a robot as more reliable and trustworthy when it performed a task competently compared to when it made errors, suggesting that task competency influences perceptions in robots [36]. In a similar vein, Carter et al. found that participants rated robots who did not interrupt during a task as more competent than robots that did interrupt the collaborative task [35]. Furthermore, Walker et al. observed that participants perceived the competence of the robot more negatively when it deviated from tasks [104]. Research also showed that humans perceive robots as more competent and trustworthy when they perform analytical tasks compared to social tasks [105]. Clair et al. employed verbal feedback in a human-robot collaboration task and found that ratings of team performance and the robot as a teammate were improved [55]. Although some work explored the interaction between verbal messages and task competency, there is a gap in the exploration of the effect of language fluency and task competency in HRI. Therefore, in my research, we aim to examine the effects of varying task competency paired with varying language fluency on the perceptions of a social robot partner.

Chapter 3

HYPOTHESIS & RESEARCH QUESTIONS

To examine the impact of varying language fluency and task competency on the perception of a social robot, we have designed three studies in which each has four experimental conditions. In particular, we aim to investigate whether perceptions vary for a robot that is: (1) fluent and task competent; (2) fluent and task incompetent; (3) disfluent and task competent, or (4) disfluent and task incompetent. We refer to these four conditions as the varying robot conditions. Building on the literature, we want to investigate the impact of these varying robot conditions on the general perception of the robot (verbal competence, intelligence, reliability), the robot meeting expectations, the robot as a teammate, and participants' willingness to work with the robot again.

The hypotheses that I developed are based on existing correlations and perceptions of non-native English speakers from Section 2.2 and task competent robots from Section 2.4. Given that monolingual and native English speakers are less willing to work with and find non-native English speakers to be less verbally competent, intelligent, reliable, and trustworthy than native English speakers, I anticipate that this would likely also be the case regarding the perception of disfluent robots by monolingual English participants. In contrast, since multilingual individuals were found to exhibit higher cognitive empathy, open-mindedness, and willingness to accept that people may differ in values, I reason that they may be more understanding towards and accepting of the disfluent robots as well. Furthermore, multilinguals have lower levels of communicative anxiety, so I argue that they will be less deterred by the language errors made by the disfluent robots. As there is not sufficient research to date studying the combined impact of language fluency and task competency on robot perception, I base my hypotheses surrounding perception of varying task competency within robots on existing positive correlations established between robots that are task-competent and their perceptions of intelligence, reliability, and the robot as a teammate.

Therefore, the hypotheses for Studies I, II, and III are as follows:

3.1 Hypothesis

- H1: Both monolingual and multilingual participants will have a higher rating of intelligence, reliability, their expectations being met, the robot as a teammate, and the participants' willingness to work with the robot again in the fluent-competent condition than in the fluent-incompetent condition. Additionally, there will be no difference between the ratings of verbal competence in fluent-competent and fluent-incompetent conditions.
- H2: Both monolingual and multilingual participants will have a higher rating on general perception of the robot, their expectations being met, the robot as a teammate, and the participants' willingness to work with the robot again in the fluent-competent condition than in the disfluent-incompetent condition.
- H3: Monolingual participants will more positively perceive the fluent-incompetent robot than the disfluent-competent robot across all dependent measures while multilingual participants will more positively perceive the disfluent-competent robot than the fluent-incompetent robot across all dependent measures except for verbal competence.
- H4: Monolingual individuals will rate perceptions of the fluent-competent and fluent-incompetent robots higher than the disfluent-competent and disfluent-incompetent robots, respectively. Multilingual individuals will not differ in their ratings between the fluent-competent and fluent-incompetent robots and disfluent-competent and disfluent-incompetent robots, respectively.
- H5: Multilingual participants will rate the disfluent-competent and disfluent incompetent robots higher across the dependent measures compared to monolingual participants.
- H6: Monolingual participants will rate the fluent-incompetent robot higher across the dependent measures compared to the multilingual participants.

We aimed to address the hypotheses through the following research questions across three studies:

3.2 Study I Research Questions

- RQ1: How does the general perception (verbal competence, intelligence, and reliability) of the robot partner differ across the varying robot conditions from the perspective of monolingual individuals?
- RQ2: How do the perceptions of the participants' expectations being met differ across the varying robot conditions from the perspective of monolingual individuals?
- RQ3: How do the perceptions of the robot as a teammate differ across the varying robot conditions from the perspective of monolingual individuals?
- RQ4: How does the willingness to work with the robot differ across the varying robot conditions from the perspective of monolingual individuals?

3.3 Study II Research Questions

- RQ5: How does the general perception (verbal competence, intelligence, and reliability) of the robot partner differ across the varying robot conditions from the perspective of multilingual individuals?
- RQ6: How do the perceptions of the participants' expectations being met differ across the varying robot conditions from the perspective of multilingual individuals?
- RQ7: How do the perceptions of the robot as a teammate differ across the varying robot conditions from the perspective of multilingual individuals?
- RQ8: How does the willingness to work with the robot differ across the varying robot conditions from the perspective of multilingual individuals?

3.4 Study III Research Questions

- RQ9: How does the general perception (verbal competence, intelligence, and reliability) of the robot partner differ across the varying robot conditions between monolingual and multilingual individuals?

- RQ10: How do the perceptions of the participants' expectations being met differ across the varying robot conditions between monolingual and multilingual individuals?
- RQ11: How do the perceptions of the robot as a teammate differ across the varying robot conditions between monolingual and multilingual individuals?
- RQ12: How does the willingness to work with the robot differ across the varying robot conditions between monolingual and multilingual individuals?

Chapter 4

METHODS

4.1 What Am I? Study Design

To address these research questions, we designed a human-robot interactive scenario, which is loosely based on the guessing game, “What Am I?” This game consists of the interaction partner prompting the robot to guess the identity of an animal that the robot is assuming after hearing a few characteristics of that animal from the robot. The robot was positioned on a small table to face the human agent during the entire interaction. The participant was provided with a list of animals and their associated characteristics which they could reference throughout the game. The participant was either assigned to a fluent robot, which spoke English at a native fluency level, or a disfluent robot, which spoke English at a non-native fluency level for the entire duration of the game. We define fluency in detail in the following section.

To initiate the game, the robot greets the player and asks if they are ready to begin the round. After the partner affirms, the robot states the characteristics of the animal whose identity it assumed and asks the participant, “What am I?” (fluent robot) or “What I am?” (disfluent robot). Then, the participant identifies the animal, and the robot either correctly confirms the accuracy of the identity or incorrectly rejects the participant’s selection and states the wrong animal for the characteristics it assumed. The robot then continues to the next round and repeats this process with a different animal identity until a total of three rounds are completed. At the conclusion of the third round, the robot states the game has ended and thanks the human partner for playing. A sample script of the varying robot interactions is included in Table 1.

4.2 Language Fluency and Task Competency Classification

Table 1: Sample Script for The Four Varying Robot Conditions.

	Fluent	Disfluent
Task Competent	NAO: I like bananas and I climb trees. What am I? Human: Monkey NAO: Yes, that is correct. Let's go to the next round.	NAO: Me bananas and climb tree. What I am? Human: Monkey NAO: Yes, correct. We go next round.
Task Incompetent	NAO: I like bananas and I climb trees. What am I? Human: Monkey NAO: No, I am a rabbit. Let's go to the next round.	NAO: Me bananas and climb tree. What I am? Human: Monkey NAO: No, me rabbit. We go next round.

4.2.1 Language Fluency

There are many factors that can impact the spoken language fluency of a non-native speaker and the perception of their language fluency such as stutter, accent, hesitation, intonation, phonological processes, pronunciation rules, and ungrammatical words or sentence structures, etc. [37], [39], [56], [57]. For this study, we incorporated grammatical accuracy in a speech to represent the fluent robot partner. In order to accurately depict disfluency in the robot, we need to define speech disfluency. A possible taxonomy to understand speech disfluencies is by distinguishing speech disfluencies into two groups: disfluencies that stem from uncertainty and errors or error-type disfluencies (ETDs) [58]. This taxonomy lists the principal measures of uncertainty-related speech disfluencies to include factors such as hesitations and repetition while error-type disfluencies encompass measures such as grammatical errors and contamination. In this work, we interpreted the root of speech disfluency to be ETDs. Specifically, we incorporated grammatical errors in a speech to represent the disfluent robot partner. We also chose this factor because grammatical errors can be made easily decipherable to participants in robot speech, and we wanted to avoid disfluency measures that could be interpreted as inherent to the robot. For example, stutter or hesitation could have been incorrectly attributed to or perceived as malfunctions or natural behavior of the robot. The types of grammatical errors implemented in the disfluent robot's dialogue included inaccurate subject-verb agreement, incorrect singular/plural agreement, wrong pronoun usage (subjective vs objective), incorrect word form, and lack of article and preposition usage. These errors were based on the most common grammatical errors made by non-native English speakers [59]– [61]. A sample script of the disfluent robot interaction is included in Table 1.

4.2.2 Task Competency

Along with verbal language fluency, this study also looked at the perceptions of two types of task competencies by the robot: task competent (accurate task completion) and task incompetent (inaccurate task completion). The task competent condition was demonstrated in the game by the robot correctly affirming the animal that the human player guessed to be the robot's assumed identity. The task incompetent condition was represented in the game by the robot incorrectly rejecting the animal that the human player guessed to be the robot's assumed identity and stating a different incorrect animal as the answer. A sample script of the task competent and task incompetent robot interactions are included in Table I.

4.3 The Robot

In our study, the SoftBank Robotics' NAO humanoid robot was used. To capture the robot's nonverbal gestures, we utilized the expressive behavior modules in SoftBank's Choregraphe Suite and for the robot's verbal performance, we worked with Amazon Polly's text-to-speech platform [62]. The decision to use Amazon Polly over Choegraphe's text-to-speech option was due to Amazon Polly's robust annunciation and timing capabilities. To mitigate potential bias with the perceptions of the robot's gender, we use the gender-neutral "ivy" voice. Then, 18 sound clips of the fluent, disfluent, accurate task completion, and inaccurate task completion dialogue were added to the Choregraphe program and matched with some basic animation behaviors. In Choregraphe, we also utilized the speech recognition modules so that the robot could react accordingly to the participant's responses.

Chapter 5

STUDY I

5.1 Procedure

Participants reviewed a study information document for consent along with the task instructions and were asked to complete a brief demographic survey. Each participant was then randomly assigned to one of four participant groups where each group differed in its combination of language fluency and task competency. Next, the participants played three rounds of the "What Am I?" game with the NAO robot. In each round, the robot would assume the identity of an animal and give a description of that animal to the participant who would then have to guess the identity of the animal based on the provided description. After playing all three rounds, participants were asked to rate how much they agreed with the robot being verbally competent, intelligent, reliable, a good teammate, meeting their expectations, and willing to work with the robot again on a five-point Likert scale ranging from "strongly disagree" (1) to "strongly agree" (5). The entire study took approximately 15 minutes. At the conclusion of the study, participants were debriefed and compensated.

5.2 Participant Classification

To reduce the participants' linguistic bias in the perception of the language fluency and task competency of the robot, we only chose to select monolingual participants to participate in this study. The monolingual categorization was based on participants' responses to a pre-task survey that collected biographical, demographic, and linguistic information and was completed upon arrival. The linguistic portion comprised two questions where the first asked how many languages the participant had verbal native fluency in, and the second followed up by asking to list the language(s) from the previous question. Native fluency was defined to the participants on the survey as being a native language that does not contain unnatural grammatical errors, pauses, stutters, repetitions, or self-corrections when speaking in the language. The participants of this study identified themselves as monolingual native English speakers. In order to be classified as a monolingual native speaker, the participant had to self-report their linguistic abilities and answer the first question with the number “1” and the second question with the language “English”'.

5.3 Participants

A total of 60 adults participated in the study (70.0% female (n = 42), 30.0% male (n = 18)). The mean age of participants was 21 years (SD = 2.75). All participants were required to be English speakers, available to participate in-person, at least 18 years of age or older, and monolingual. The participants consisted of undergraduate and graduate students, and high school and college graduates. During the study, participants also reported their experience with robots in general (M = 1.42, SD = 0.740). Participants were compensated with a $5 gift card for participating in the study. The study was approved by the Institutional Review Board.

5.4 Measures

After the game, participants were asked to rate perceptions of the robot's verbal competence, intelligence, and reliability. We also had participants reflect on whether the NAO robot was a good teammate, rate their willingness to work with the robot again, and express whether the robot met their expectations on the Likert scale described in Section 5.1.

5.5 Results & Discussion

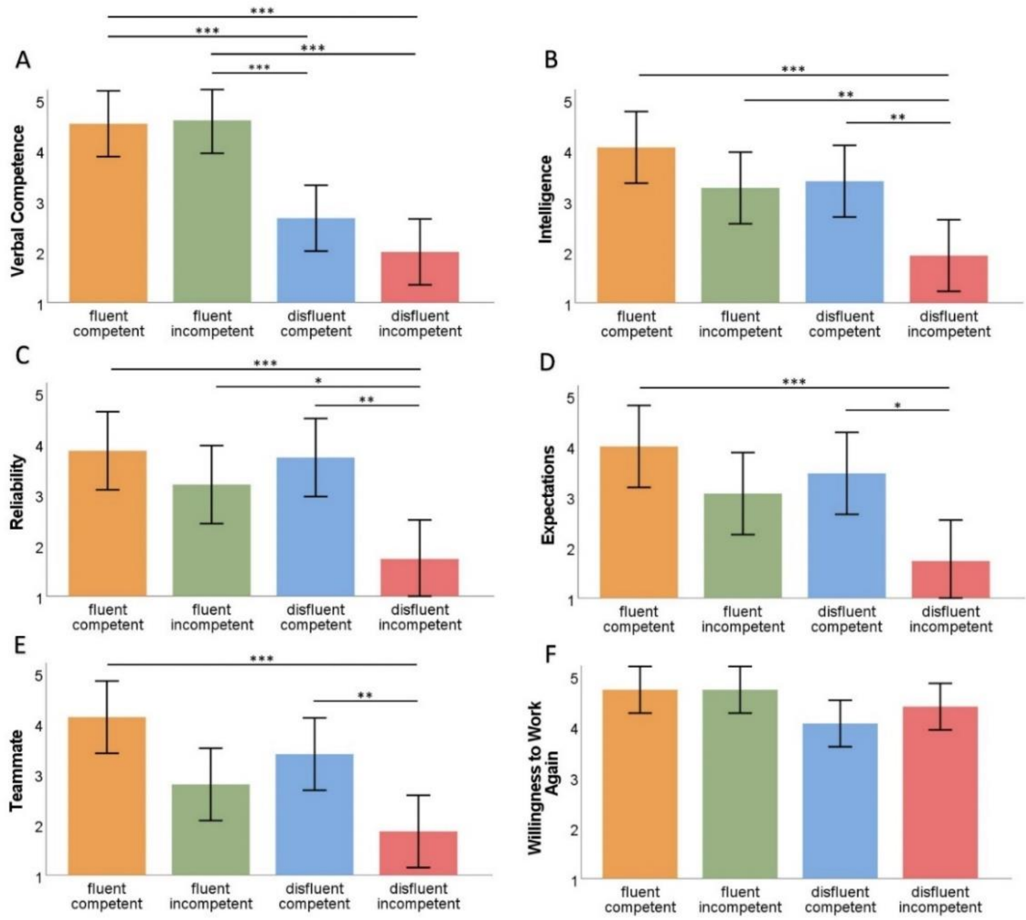

*Figure 1: Bar graphs showing monolingual participant perceptions of verbal competence, intelligence, and reliability as well as their ratings of expectations being met, the robot as a teammate, and their willingness to work with the robot again across four varying robot conditions (the robot being fluent and task competent, fluent and task incompetent, disfluent and task competent, and disfluent and task incompetent). Error bars 95% CI. The significant values are shown in * (∗ = p < .05, ∗∗ = p < .01, ∗ ∗ ∗ = p < .001).*

5.5.1 RQ1: Effects of Varying Robot Conditions on Perceptions

5.5.1.1 Verbal Competence

Separate univariate ANOVAs on the dependent variables revealed significant effects of varying robot conditions on perceptions of verbal competence, ($F(3, 56) = 16.46$, $p < .001$, $\eta^2 = .469$). The results (Fig.

1-A) suggest participants in the fluent-competent condition rated verbal competence (M = 4.53, SD = 1.060) higher than participants in the disfluent-competent group (M = 2.67, SD = 1.397), ($p < .001$). Participants in the fluent-competent condition rated verbal competence (M = 4.53, SD = 1.060) higher than participants in the disfluent-incompetent group (M = 2.00, SD = 1.604), ($p < .001$). Participants in the fluent-incompetent condition (M = 4.60, SD = 0.828) rated verbal competence higher than participants in the disfluent- competent group (M = 2.67, SD = 1.397), ($p < .001$). Participants in the fluent-incompetent condition rated verbal competence (M = 4.60, SD = 0.828) higher than participants in the disfluent-incompetent group (M = 2.00, SD = 1.604), ($p < .001$).

Discussion: The results indicate that there was no significant difference between the perception of verbal competence in the fluent-competent and fluent-incompetent robots. The results also suggest that participants found both the fluent robots (fluent-competent and fluent-incompetent) to be significantly more verbally competent than the two disfluent robots (disfluent-competent and disfluent-incompetent). This was expected as the disfluent robots made grammatical errors, which may influence the perception of the verbal competence of those robot conditions. Additionally, the fluent-incompetent robot's verbal competence was rated significantly higher than that of the disfluent-competent robot. Finally, the fluent-component and fluent-incompetent robots received higher participant ratings than the disfluent-component and disfluent-incompetent robots within verbal competence, respectively. The results may indicate that language fluency plays a greater role than task competency on the perception of verbal competence of a robot for monolingual individuals.

5.5.1.2 Intelligence

Separate univariate ANOVAs on the dependent variables revealed significant effects of varying robot conditions on perceptions of intelligence ($F(3, 56) = 6.425$, $p < .001$, $\eta^2 = .256$). The results (Fig. 1-B) suggest participants in the fluent-competent condition rated intelligence (M = 4.07, SD = 1.534) higher than participants in the disfluent-incompetent condition (M = 1.93, SD =1.534), ($p < .001$). Additionally, participants in the fluent-incompetent condition rated intelligence (M = 3.27, SD = 1.163) higher than participants in the disfluent-incompetent condition (M = 1.93, SD = 1.534), (p = .047). Participants in the disfluent-competent condition rated intelligence (M = 3.40, SD = 1.183) higher than participants in the disfluent-incompetent condition (M = 1.93, SD = 1.534), (p = .024).

Discussion: The results suggest that the participants perceived the fluent-competent, fluent-incompetent, and disfluent-competent robots to be more intelligent than the disfluent-incompetent robot. Additionally, the fluent-incompetent robot received higher participant ratings than the disfluent-incompetent robot within the perception category of intelligence. However, in the fluent-incompetent and disfluent-competent conditions, at least one of the variables (language fluency or task competency) was compromised, yet robots in both conditions were still perceived to be more intelligent than the disfluent-incompetent robot, which had both variables compromised. These results may suggest that both language fluency and task competency have an impact on the perceived intelligence of the robot.

5.5.1.3 Reliability

Separate univariate ANOVAs on the dependent variables revealed significant effects of varying robot conditions on perceptions of reliability ($F(3, 56) = 6.463$, $p < .001$, $\eta^2 = .257$). The results (Fig. 1-C) suggest participants in the fluent-competent condition rated reliability (M = 3.87, SD = 1.807) higher than participants in the disfluent-incompetent condition (M = 1.73, SD = 1.534), (p = .001). Additionally, participants in the fluent-incompetent condition rated reliability (M = 3.20, SD = 1.207) higher than participants in the disfluent-incompetent group (M = 1.73, SD = 1.534), (p = .044). Furthermore, participants in the disfluent-competent condition rated reliability (M = 3.73, SD = 1.335) higher than participants in the disfluent-incompetent condition (M = 1.73, SD = 1.534), (p = .003).

Discussion: The results indicate that participants in the fluent-competent, fluent-incompetent, and disfluent-competent conditions all perceived the robot to be significantly more reliable than how the participants in the disfluent-incompetent condition perceived the robots. Additionally, the fluent-incompetent robot received higher participant ratings for reliability than the disfluent-incompetent robot.

However, in the fluent-incompetent and disfluent-competent conditions, at least one of the variables (language fluency or task competency) was compromised, yet robots in both conditions were still perceived to be more reliable than the disfluent-incompetent robot, which had both variables compromised. The results may imply that both language fluency and task competency have an impact on the perceived reliability of the robot.

5.5.2 RQ2: Effects of Varying Robot Conditions on Participants' Expectations Being Met

Separate univariate ANOVAs on the dependent variables revealed significant differences between the varying robot conditions on ratings of expectations being met ($F(3, 56) = 5.716$, $p = .002$, $\eta^2 = .234$). Specifically, the results (Fig.1-D) suggest participants in the fluent-competent condition rated that their expectations were met ($M = 4.00$, $SD = 1.732$) significantly higher than participants in the disfluent-incompetent condition ($M = 1.73$, $SD = 1.534$), ($p = .001$). Additionally, participants in the disfluent-competent condition rated that their expectations were met ($M = 3.47$, $SD = 1.356$) significantly higher than participants in the disfluent-incompetent conditions ($M = 1.73$, $SD = 1.534$), ($p = .019$).

Discussion: The results indicate that participants in the fluent-competent and disfluent-competent conditions perceived the robot to have met their expectations more than the participants in the disfluent-incompetent condition did. Additionally, there was no significant difference in the rating between fluent-competent and disfluent-competent or fluent-incompetent and disfluent-incompetent robot conditions for this perception category. This result may suggest that task competency has a greater impact on the perception of the robot meeting expectations than language fluency.

5.5.3 RQ3: Effects of Varying Robot Conditions on Perception of Robot as a Teammate

Separate univariate ANOVAs on the dependent variables revealed significant differences between the varying robot conditions on ratings of the robot as a teammate ($F(3, 56) = 7.188$, $p < .0005$, $\eta^2 = .278$). Specifically, the results (Fig. 1-E) suggest participants in the fluent-competent condition rated the robot as a teammate ($M = 4.13$, $SD = 1.356$) higher than participants in the disfluent-incompetent condition did (M

= 1.87, SD = 1.552), ($p < .001$). Additionally, participants in the disfluent-competent condition (M = 3.40, SD = 1.242) rated the robot as a teammate higher than participants in the disfluent-incompetent condition did (M = 1.87, SD = 1.552), ($p = .019$).

Discussion: Participants in the fluent-competent and disfluent-competent conditions rated the robot more highly for being a good teammate than participants in the disfluent-incompetent condition did. Additionally, there was no significant difference in the rating between fluent-competent and disfluent-competent or fluent-incompetent and disfluent-incompetent robot conditions for this perception category. This result may indicate that task competency has a greater impact on the perception of the robot as a teammate than language fluency.

5.5.4 RQ4: Effects of Varying Robot Conditions on Participants' Willingness to Work with Robot Again

There was no statistical significance of the effects of varying robot conditions on participants' willingness to work with the robot again ($F(3, 56) = 1.927$, $p = .136$, $\eta^2 = .094$) (Fig. 1-F).

Discussion: The results may suggest that neither language fluency nor task competency impacts whether monolingual participants would be willing to work with the robot again.

Chapter 6

STUDY II

6.1 Procedure

The procedure is replicated from that of Study I in Section 5.1.

6.2 Participant Classification

To control the participants' linguistic bias in the perception of the language fluency and task competency of the robot, we only selected multilingual participants to participate in this study. The multilingual categorization was based on participants' responses to a pre-task survey that collected biographical, demographic, and linguistic information and was completed upon arrival. The linguistic portion comprised two questions where the first asked how many languages the participant had verbal native fluency in, and the second followed up by asking to list the language(s) from the previous question. Native fluency was defined to the participants on the survey as being a native language that does not contain unnatural grammatical errors, pauses, stutters, repetitions, or self-corrections when speaking in the language. The participants of this study identified themselves as multilingual speakers. In order to be classified as a multilingual native speaker, the participant had to self-report their linguistic abilities and answer the first question with a number greater than or equal to “2” and the second question with two or more languages including “English”.

6.3 Participants

A total of 60 adults participated in the study (53.3% female (n = 32), 46.7% male (n = 28)). The mean age of participants was 21 years (SD = 3.09). All participants were required to be fluent in at least two languages with at least one of the fluently spoken languages being English, available to participate in-person, and at least 18 years of age or older. The mean number of languages fluently spoken by the participants was 2.38 (SD = 0.490) and the linguistic profile of the participants included 61.7% bilingual (n = 37) and 38.3%

trilingual (n = 23). For the purpose of this study, we classified bilingual individuals as multilinguals, so we can also say that 100% of participants were multilingual (n = 60). The participants consisted of undergraduate and graduate students, and high school and college graduates. During the study, participants also reported their experience with robots in general (M = 1.89, SD = 0.940). Participants were compensated with a $5 gift card for participating in the study. The study was approved by the Institutional Review Board.

6.4 Measures

The measures are the same as that of Study I in Section 5.4.

6.5 Results & Discussion

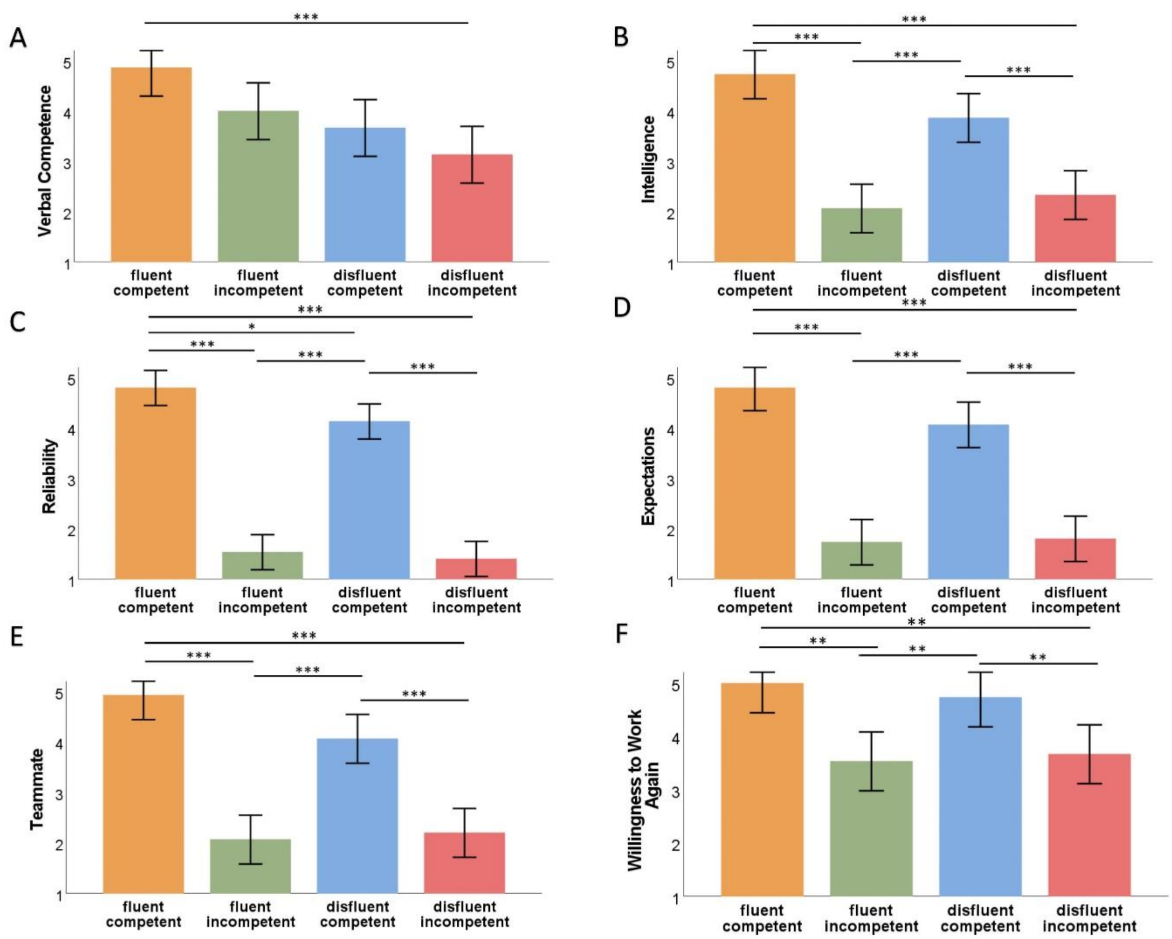

Figure 2: Bar graphs showing multilingual participant perceptions of verbal competence, intelligence, and reliability as well as their ratings of expectations being met, the robot as a teammate, and their willingness to work with the robot again across four

*varying robot conditions (the robot being fluent and task competent, fluent and task incompetent, disfluent and task competent, and disfluent and task incompetent). Error bars 95% CI. The significant values are shown in * (* = p < .05, ** = p < .01, * * * = p < .001).*

To address our research questions, we evaluated the effects of varying language fluency and task competency on the perceptions of a robot partner. A series of multivariate ANOVAs were conducted with four groups that made up the varying robot conditions as independent variables and six dependent variables (i.e., perceptions of verbal competence, intelligence, reliability, expectations being met, the robot as a teammate, and willingness to work with the robot again). For the analysis, an inspection of the data and Levene's test provided no strong evidence against the assumption of constant variance. Furthermore, we used Tukey's Honestly Significant Difference (HSD) test for post-hoc comparisons. The multivariate analysis highlighted a statistically significant interaction effect between varying robot conditions on the combined perceptions of the robot partner, ($F(18, 144) = 10.715$, $p < .001$, Wilks' $\lambda = .091$, $\eta^2 = .550$).

6.5.1 RQ1: Effects of Varying Robot Conditions on Perceptions

6.5.1.1 Verbal Competence

Separate univariate ANOVAs on the dependent variables revealed significant effects of varying robot conditions on perceptions of verbal competence, ($F(3, 56) = 6.646$, $p < .001$, $\eta^2 = .263$). The results (Fig. 2-A) suggest participants in the fluent-competent condition rated verbal competence ($M = 4.87$, $SD = 0.352$) higher than participants in the disfluent-incompetent group ($M = 3.13$, $SD = 1.356$), ($p < .001$).

Discussion: The results indicate that participants found the fluent-competent robot's verbal competence to be rated significantly higher than that of the disfluent-incompetent robot. There was no significant difference between the rating of the robot's verbal competence between the fluent-competent robot and the fluent-incompetent robot. Additionally, there was no significant difference between the rating of the robot's verbal competence between the disfluent-competent robot and the disfluent-incompetent robot. The unexpected findings were that there was also no significant difference between the perception of a fluent-incompetent robot's verbal competence and that of a disfluent-competent robot. Moreover, there was no significant difference between the rating of the robot's verbal competence between the fluent-competent robot and the disfluent-competent robot. This could be due to the variables of language fluency and task competency having different levels of impact within the varying robot conditions for the perception of verbal competence.

6.5.1.2 Intelligence

Separate univariate ANOVAs on the dependent variables revealed significant effects of varying robot conditions on perceptions of intelligence ($F(3, 56) = 27.587$, $p < .001$, $\eta^2 = .596$). The results (Fig. 2-B) suggest participants in the fluent-competent condition rated intelligence (M = 4.73, SD = 0.458) higher than participants in the fluent-incompetent group (M = 2.07, SD = 0.961), ($p < .001$). Participants in the fluent-competent condition rated intelligence (M = 4.73, SD = 0.458) higher than participants in the disfluent-incompetent group (M = 2.33, SD = 1.291), ($p < .001$). Participants in the disfluent-competent group rated intelligence (M = 3.87, SD = 0.834) higher than participants in the fluent-incompetent condition (M = 2.07, SD = 0.961), ($p < .001$). Participants in the disfluent-competent condition rated intelligence (M = 3.87, SD = 0.834) higher than participants in the disfluent-incompetent group (M = 2.33, SD = 1.291), ($p < .001$).

Discussion: The results suggest that the participants perceived both the competent robots (fluent-competent and disfluent-competent) to be more intelligent than the incompetent robots (fluent-incompetent and disfluent-incompetent robots). Participants also found the fluent-competent robot to be more intelligent than the fluent-incompetent robot. Similarly, the disfluent-competent robot was rated to have a significantly higher intelligence than the disfluent-incompetent robot. These results may indicate that task competency has a greater impact than language fluency on the perceived intelligence of the robot.

6.5.1.3 Reliability

Separate univariate ANOVAs on the dependent variables revealed significant effects of varying robot conditions on perceptions of reliability ($F(3, 56) = 101.494$, $p < .001$, $\eta^2 = .845$). The results (Fig. 2-C) suggest participants in the fluent-competent condition rated reliability (M = 4.73, SD = 0.458) higher than participants in the fluent-incompetent group (M = 2.07, SD = 0.961), ($p < .001$). Participants in the fluent-competent condition rated reliability (M = 4.80, SD = 0.561) higher than participants in the disfluent-incompetent group (M = 1.40, SD = 0.828), ($p < .001$). Participants in the fluent-competent group rated reliability (M = 4.80, SD = 0.561) higher than participants in the disfluent-competent condition (M = 4.13, SD = 0.743), ($p < .0043$). Participants in the disfluent-competent group rated reliability (M = 4.13, SD = 0.743) higher than participants in the fluent-incompetent condition (M = 2.07, SD = 0.961), ($p < .001$). Participants in the disfluent-competent condition rated reliability (M = 4.13, SD = 0.743) higher than participants in the disfluent-incompetent group (M = 1.40, SD = 0.828), ($p < .001$).

Discussion: The results indicate that the participants perceived both the competent robots (fluent-competent and disfluent-competent) to be more reliable than the incompetent robots (fluent-incompetent and disfluent-incompetent robots). Participants also found the fluent-competent robot to be more reliable than the fluent-incompetent robot. Similarly, the disfluent-competent robot was rated to be significantly more reliable than the disfluent-incompetent robot. However, an unexpected finding was that multilingual participants perceived the fluent-competent robot to be significantly more reliable than the disfluent-competent robot. These results need to be further explored in order to determine whether task competency has a greater impact on the perceived reliability of the robot.

6.5.2 RQ2: Effects of Varying Robot Conditions on Participants' Expectations Being Met

Separate univariate ANOVAs on the dependent variables revealed significant differences between the varying robot conditions on ratings of expectations being met ($F(3, 56) = 48.446$, $p < .001$, $\eta^2 = .722$). Specifically, the results (Fig. 2-D) suggest participants in the fluent-competent condition rated that their expectations were met (M = 4.80, SD = 0.561) significantly higher than participants in the fluent-incompetent group (M = 1.73, SD = 0.704), ($p < .001$). Participants in the fluent-competent condition rated that their expectations were met (M = 4.80, SD = 0.561) significantly higher than participants in the disfluent-incompetent group (M = 1.80, SD = 0.941), ($p < .001$). Participants in the disfluent-competent group rated that their expectations were met (M = 4.07, SD = 1.163) significantly higher than participants in the fluent-incompetent condition (M = 1.73, SD = 0.704), ($p < .001$). Participants in the disfluent-competent condition rated that their expectations were met (M = 4.07, SD = 1.163) significantly higher than participants in the disfluent-incompetent group (M = 1.80, SD = 0.941), ($p < .001$).

Discussion: The results suggest that the participants perceived both the competent robots (fluent-competent and disfluent-competent) to have met their expectations more than the incompetent robots (fluent-incompetent and disfluent-incompetent robots). Participants also found the fluent-competent robot to meet their expectations more than the fluent-incompetent robot. Similarly, the disfluent-competent robot was rated to have met expectations significantly more than the disfluent-incompetent robot. These results may indicate that task competency has a greater impact than language fluency on how well the robot meets expectations.

6.5.3 RQ3: Effects of Varying Robot Conditions on Perception of Robot as a Teammate

Separate univariate ANOVAs on the dependent variables revealed significant differences between the varying robot conditions on ratings of the robot as a teammate ($F(3, 56) = 34.064$, $p < .001$, $\eta^2 = .646$). Specifically, the results (Fig. 2-E) suggest participants in the fluent-competent condition rated the robot as a teammate (M = 4.93, SD = 0.258) higher than participants in the fluent-incompetent group (M = 2.07, SD = 1.033), ($p < .001$). Participants in the fluent-competent condition rated the robot as a teammate (M = 4.93, SD = 0.258) higher than participants in the disfluent-incompetent group (M = 2.20, SD = 1.320), ($p < .001$). Participants in the disfluent-competent group rated the robot as a teammate (M = 4.07, SD = 0.799) higher than participants in the fluent-incompetent condition (M = 2.07, SD = 1.033), ($p < .001$).Participants in the disfluent-competent condition rated the robot as a teammate (M = 4.07, SD = 0.799) higher than participants in the disfluent-incompetent group (M = 2.20, SD = 1.320), ($p < .001$).

Discussion: The results suggest that the participants perceived both the competent robots (fluent-competent and disfluent-competent) to be better teammates than the incompetent robots (fluent-incompetent and disfluent-incompetent robots). Participants also found the fluent-competent robot to be more of a good teammate than the fluent-incompetent robot. Similarly, the disfluent-competent robot was rated as a good teammate significantly higher than the disfluent-incompetent robot. These results may indicate that task competency has a greater impact than language fluency on the perception of the robot being a good teammate.

6.5.4 RQ4: Effects of Varying Robot Conditions on Participants' Willingness to Work with Robot Again

Separate univariate ANOVAs on the dependent variables revealed significant differences between the varying robot conditions on participants' willingness to work with the robot again ($F(3, 56) = 7.124$, $p < .001$, $\eta^2 = .279$) The results (Fig. 2-F) suggest participants in the fluent-competent condition rated their willingness to work with the robot again ($M = 5.00$, $SD = 0.00$) higher than participants in the fluent-incompetent group ($M = 3.53$, $SD = 1.246$), ($p = .002$). Participants in the fluent-competent condition rated their willingness to work with the robot again ($M = 5.00$, $SD = 0.00$) higher than participants in the disfluent-incompetent group ($M = 3.67$, $SD = 1.633$), ($p = .006$). Participants in the disfluent-competent group rated their willingness to work with the robot again ($M = 4.73$, $SD = 0.594$) higher than participants in the fluent-incompetent condition ($M = 3.53$, $SD = 1.246$), ($p = .017$). Participants in the disfluent-competent condition rated their willingness to work with the robot again ($M = 4.73$, $SD = 0.594$) higher than participants in the disfluent-incompetent group ($M = 3.67$, $SD = 1.633$), ($p = .041$).

Discussion: The results suggest that the multilingual participants were more willing to work with both the competent robots (fluent-competent and disfluent-competent) again than the incompetent robots (fluent-incompetent and disfluent-incompetent robots). Participants also had a significantly higher willingness to work with the fluent-competent robot again than the fluent-incompetent robot. Similarly, participants rated their willingness to work with the disfluent-competent robot again significantly higher than the disfluent-incompetent robot. These results may indicate that task competency has a greater impact than language fluency on participants' willingness to work with the robot again.

Chapter 7

STUDY III

7.1 Procedure

For Study III, we compared the data from Study I (see Chapter 5) and Study II (see Chapter 6) with one another.

7.2 Participants

Combining Study I and Study II, a total of 120 adults were included in this study (61.7% female (n = 74), 38.3% male (n = 46)). The mean age of participants was 21 years (SD = 2.92). 50% of participants were monolingual (n = 60), 30.8% were bilingual (n = 37), and 19.2% were trilingual (n = 23). For the purpose of this study, we classified bilingual individuals as multilinguals, so we can also say that 50% of participants were multilingual (n = 60). The participants consisted of undergraduate and graduate students, and high school and college graduates. Participants also reported their experience with robots in general (M = 1.65, SD = 0.823). Participants were compensated with a $5 gift card for participating in the study. The study was approved by the Institutional Review Board.

7.3 Measures

The measures are the same as that of Study I in Section 5.4.

7.4 Results & Discussion

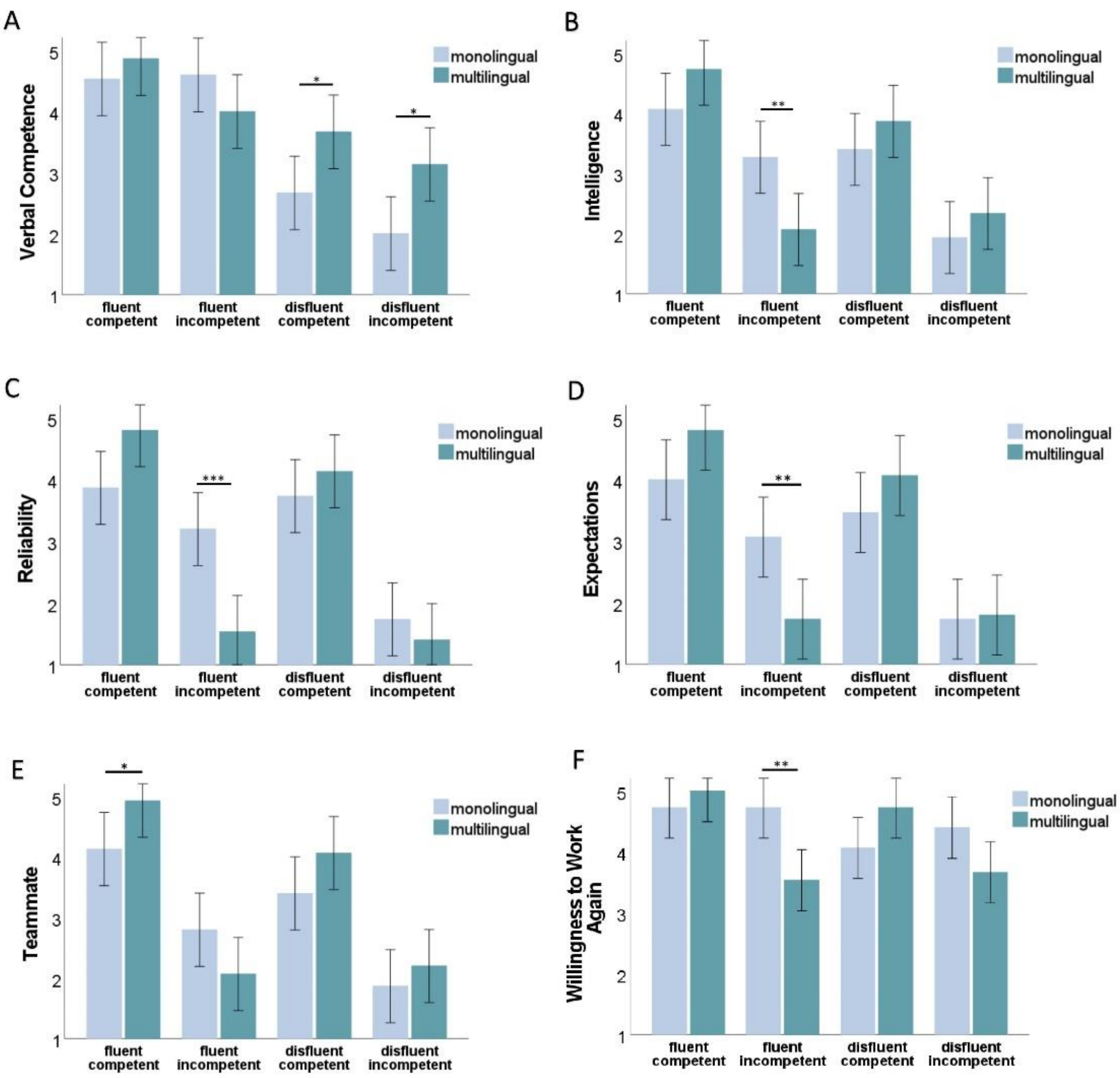

*Figure 3: Bar graphs showing the interaction between monolingual and multilingual participant perceptions of verbal competence, intelligence, and reliability as well as their ratings of expectations being met, the robot as a teammate, and their willingness to work with the robot again across four varying robot conditions (the robot being fluent and task competent, fluent and task incompetent, disfluent and task competent, and disfluent and task incompetent). Error bars 95% CI. The significant values are shown in * (∗ = p < .05, ∗∗ = p < .01, ∗ ∗ ∗ = p < .001).*

To address our research questions, we compared the effects of varying language fluency and task competency on the perceptions of a robot partner between monolingual and multilingual participant groups. A series of multivariate ANOVAs were conducted with four groups that make up the varying robot conditions as independent variables and six dependent variables (i.e., perceptions of verbal competence, intelligence, reliability, expectations being met, the robot as a teammate, and willingness to work with the robot again). For the analysis, an inspection of the data and Levene's test provided no strong evidence against the assumption of constant variance. Furthermore, we used Tukey's Honestly Significant Difference (HSD) test for post-hoc comparisons. The multivariate analysis highlighted a statistically significant interaction effect between the participant groups on the combined perceptions of the robot partner in the varying robot conditions, ($F(18, 303) = 2.371$, $p <= .002$, Wilks' $\lambda = .689$, $\eta^2 = .117$).

7.5.1 RQ1: Effects of Varying Robot Conditions on Perceptions

7.5.1.1 Verbal Competence

Separate univariate ANOVAs on the dependent variables revealed significant interactions between the participant groups and varying task conditions on perceptions of the robot's verbal competence ($F(3, 112) = 3.394$, $p = .020$, $\eta^2 = .083$). The results (Fig. 3-A) suggest multilingual participants rated verbal competence ($M= 3.67$, $SD= 1.175$) significantly higher in the disfluent-competent robot condition than their monolingual counterparts ($M= 2.67$, $SD= 1.397$), ($p=.043$). The multilingual participant group also rated verbal competence ($M= 3.13$, $SD= 1.356$) significantly higher in the disfluent-incompetent robot condition than the monolingual participants ($M= 2.00$, $SD= 1.604$), ($p=.046$).

Discussion: The results indicate that multilingual participants perceived both the disfluent robots (disfluent-competent and disfluent incompetent) to be more verbally competent than monolingual participants perceived them to be. This may imply that multilingual individuals are more empathetic than monolingual individuals towards disfluency in the perception of verbal competence.

7.5.1.2 Intelligence

Separate univariate ANOVAs on the dependent variables revealed significant interactions between the participant groups and varying task conditions on perceptions of the robot's intelligence ($F(3, 112) = 4.080$, $p = .009$, $\eta^2 = .099$). The results (Fig. 3-B) suggest monolingual participants rated intelligence ($M= 3.27$, $SD= 1.163$) significantly higher in the fluent-incompetent robot condition than the multilingual participant group ($M= 2.07$, $SD= 0.961$), ($p=.005$).

Discussion: The results may indicate that language fluency may have a greater impact on monolingual people's perception of intelligence for fluent-incompetent robots. On the other hand, multilingual people's perception of intelligence may be attributed to task competency rather than to or with language fluency for fluent-incompetent robots.

7.5.1.3 Reliability

Separate univariate ANOVAs on the dependent variables revealed significant interactions between the participant groups and varying task conditions on perceptions of the robot's reliability ($F(3, 112) = 7.135$, $p < .001$, $\eta^2 = .160$). The results (Fig. 3-C) suggest monolingual participants rated reliability ($M= 3.20$, $SD= 1.207$) significantly higher in the fluent-incompetent robot condition than the multilingual participant group ($M= 1.53$, $SD= 0.516$), ($p=.066$).

Discussion: Similar to the results on the perception of intelligence, these results may indicate that language fluency may have a greater impact on monolingual people's perception of reliability for fluent-incompetent robots while multilingual people's perception of reliability may be more impacted by task competency than language fluency.

7.5.2 RQ2: Effects of Varying Robot Conditions on Participants' Expectations Being Met

Separate univariate ANOVAs on the dependent variables revealed significant interactions between the participant groups and varying task conditions on perceptions of the robot meeting expectations (F(3, 112)

= 4.315, $p = .0006$, $\eta^2 = .104$). The results (Fig. 3-D) suggest monolingual participants rated the robot meeting their expectations (M= 3.07, SD= 1.624) significantly higher in the fluent-incompetent robot condition than the multilingual participant group (M= 1.73, SD= 0.704), ($p=.007$).

Discussion: The results may imply that language fluency played a greater role than task competency in monolingual participants' rating of the robot meeting their expectations for the fluent-incompetent robot condition. On the contrary, task competency may have impacted multilingual people's perception of the robot meeting their expectations for the fluent-incompetent robot more than language fluency did.

7.5.3 RQ3: Effects of Varying Robot Conditions on Perception of Robot as a Teammate

Separate univariate ANOVAs on the dependent variables revealed significant interactions between the participant groups and varying task conditions on perceptions of the robot as a good teammate ($F(3, 112) = 2.590$, $p = .056$, $\eta^2 = .065$). The results (Fig. 3-E) suggest multilingual participants rated the robot as a good teammate (M= 4.93, SD= 0.258) significantly higher in the fluent-competent robot condition than the monolingual participant group (M= 4.13, SD= 1.356), ($p=.033$).

Discussion: The results suggest that multilingual individuals rated the fluent-competent robot as being a good teammate significantly higher than monolingual individuals. This may imply that both language fluency and task competency play a greater role in multilingual people perceiving how good of a teammate a fluent-competent robot is.

7.5.4 RQ4: Effects of Varying Robot Conditions on Participants' Willingness to Work with Robot Again

Separate univariate ANOVAs on the dependent variables revealed significant interactions between the participant groups and varying task conditions on participants' willingness to work with the robot again ($F(3, 112) = 5.795$, $p = .001$, $\eta^2 = .134$). The results (Fig. 3-F) suggest monolingual participants rated willingness to work with the robot again (M= 4.73, SD= 1.033) significantly higher in the fluent-incompetent robot condition than the multilingual participant group (M= 3.53, SD= 1.246), ($p=.008$).

Discussion: The results may indicate that language fluency may have a greater impact on monolingual people's willingness to work with the robot again in the fluent-incompetent robot condition. Conversely, task competency may have a greater impact than language fluency on multilingual people's willingness to work with the robot again in the fluent-incompetent robot condition.

Chapter 8

FINDINGS

In Study I, for RQ1, we observed a significant interaction between varying robot conditions on perceptions of the robot's verbal competence, intelligence, and reliability. An important interaction to highlight is that the fluent-incompetent robot was perceived to be significantly more verbally competent, intelligent, and reliable than the disfluent-incompetent robot. For RQ2 and RQ3, we explored the effect of varying robot conditions on ratings of the robot meeting expectations and the robot as a teammate. We discerned a significant interaction between varying robot conditions on perceptions of the robot meeting expectations and on the rating of the robot as a teammate. The results indicate that both the competent robots (fluent and disfluent) were perceived as better at meeting expectations and to being good teammate than the disfluent-incompetent robot. Therefore, we can reason that task competency may be a factor of greater influence on the perceptions of the robot meeting expectations and the robot as a teammate. There were no statistically significant differences in rating between the fluent-competent and fluent-incompetent robots across all the dependent measures, which did not support the first part of our hypothesis (H1) but supported the second part of H1 which was concerned with the perception of verbal competence. This may suggest that monolingual individuals are impacted more by uncompromised language fluency (fluent) than varying task competency (competent and incompetent) in their ratings of the robot for all perception categories except willingness to work with the robot again. Furthermore, an interesting finding in this study was that the disfluent-competent robot received significantly higher ratings than the disfluent-incompetent robot in the perception categories of intelligence, reliability, meeting expectations, and being a good teammate. These results may suggest that between compromised language fluency (disfluent) and varying task competency (competent and incompetent), task competency plays a greater role in the monolingual perception of the robot's intelligence, reliability, it meeting expectations, and it being a good teammate. The fluent-competent robot was rated higher than the disfluent-incompetent robot in every perception category except willingness to work again. Thus, the monolingual component of our hypothesis (H2) was supported by RQ1, RQ2, and

RQ3 since fluent-competent robots were rated more highly than disfluent-incompetent robots from the monolingual perspective in those five perception categories. These results can be attributed to the fact that the fluent-competent group had neither the language fluency nor task competency variable compromised while the disfluent-incompetent condition had both the variables compromised, so it is not surprising that the fluent-competent group would receive higher ratings in certain perception categories. Another essential finding is that the monolingual component of H3 was supported by RQ1 in the perception category of verbal competence since the fluent-incompetent robot was rated significantly higher than the disfluent-competent robot. It can be reasoned that language fluency plays a greater role than task competency on the impact of the robot's perceived verbal competence. Additionally, for the perceptions of intelligence and reliability, both language fluency and task competency have an impact on the participants' perceptions. The monolingual component of H4 was supported by RQ1 because the fluent-incompetent robot received higher participant ratings than the disfluent-incompetent robot within the general perception categories, which may imply language fluency played a greater role than task competency on these perceptions. An interesting observation within RQ1, RQ2, and RQ3 is that for the perception categories of intelligence, reliability, expectations being met, and the robot being a good teammate, the disfluent-competent robot was rated significantly higher than the disfluent-incompetent robot. It can be interpreted that in these perception categories, if the robot has compromised language fluency, then task competency plays a greater role in the perception of the robot. For RQ4, there were no statistically significant differences between the effects of varying robot conditions on ratings of willingness to work with the robot again. This means neither language fluency nor task competency alone had a significant impact on the participant's rating of willingness to work with the robot again. We posit that this may be due to the compromised variables in the task design being relatively low stakes in their impact on participants. For example, the robot failing to perform the task correctly may not have been consequential enough in this context to the participant for them not to want to work with the robot again. There is a need for further exploration of the confounding factors that may impact the willingness of participants to work with the robot again. Overall, the first part of the monolingual component of H1 was not supported in any perception category, but the second part of H1 was supported in the category of verbal competence. H2 was supported in five of the six perception categories. The monolingual aspect of H3 was supported by the category of verbal competence within general perception but not supported in the other five categories. The first part of the monolingual component of H4 was supported by the perception category of verbal competence, and the second part of the monolingual aspect of H4 was supported by all categories in general perception (verbal competence, intelligence, and reliability) but not by the other three perception categories. It is imperative to understand the complete implications of these findings through further research on the impact of varying language fluency and task competency on perceptions of the robot through a monolingual perspective.

In Study II, for RQ5, we observed a significant interaction between varying robot conditions on perceptions of the robot's verbal competence, intelligence, and reliability. An important interaction to highlight is that multilingual participants did not perceive the verbal competence of the fluent-competent and fluent-incompetent robots differently. This was expected since both robots were fluent, and this perception category was specifically looking at the verbal competence of the robot. Additionally, the only significant difference found between the robot conditions in the perception of verbal competence was between the fluent-competent robot and disfluent-incompetent robot where the first was perceived to be significantly more verbally competent than the latter. Both the language fluency and task competency variables were not compromised in the fluent-competent condition while both were compromised in the disfluent-incompetent condition. Otherwise, when only one variable was compromised such as in the fluent-incompetent (task competency compromised) and disfluent-competent (language fluency compromised) robot conditions, there was no significant difference between these groups. Thus, it can be implied that language fluency and task competency may play a joint role in the perception of a robot's verbal competence from a multilingual perspective, and this may indicate that the perception of verbal competence by multilingual individuals is less varied between groups when each of those robot groups possesses only one uncompromised variable between language fluency and task competency. For RQ6 and RQ7, we explored the effect of varying robot conditions on ratings of the robot meeting expectations and the robot being a teammate. We discerned a significant interaction between varying robot conditions on perceptions of the robot meeting expectations and on the rating of the robot being a teammate. For RQ8, we investigated the effect of varying robot conditions on the participants' willingness to work with the robot again and found significant interactions between the varying robot conditions and this perception category. RQ5, RQ6, RQ7, and RQ8 all support our hypothesis (H1) because between the fluent-competent and fluent-incompetent robots, although both robots were fluent, multilingual participants perceived the one that was also task-competent to be more highly rated in all the perception categories except verbal competence. This implies that task competency has a greater impact on the perceptions of intelligence, reliability, robot meeting expectations, robot being a good teammate, and participant willingness to work with the robot again. Next, RQ5, RQ6, RQ7, and RQ8 also support H2 because multilingual participants rated the fluent-competent robot to be significantly higher than the disfluent-incompetent robot in all perception categories. These results indicate that multilingual individuals prefer a robot that is uncompromised in both language fluency and task competency over a robot that is compromised in both those variables. However, when the language fluency was swapped between those two robot conditions so that the robot conditions were disfluent-competent and fluent-incompetent, we find that participants rated the disfluent-competent robot significantly higher than the fluent-incompetent robot in all perception categories except for verbal competence. Therefore, RQ5, RQ6, RQ7, and RQ8 all support the multilingual aspect of H3. This indicates again that between compromised task competency and compromised language fluency, multilingual individuals prefer compromised language fluency and uncompromised task competency which indicates that task competency has a greater impact on the perceptions of intelligence, reliability, robot meeting expectations, robot being a good teammate, and participant willingness to work with the robot again. Aside from the perception category of reliability, RQ5, RQ6, RQ7, and RQ8 support the multilingual component of H4. When two robots maintained uncompromised task competency and varied language fluency so that one robot was fluent-competent and the other disfluent-competent, the results indicated that multilingual participants' ratings between the two robot conditions did not significantly differ in all perception categories except reliability. This shows how there was no difference in rating between both competent robots which indicates task competency impacts the perception of multilingual individuals more than language fluency in the realms of intelligence, expectations, teammate, and willingness to work again. Similarly, when two robots maintained compromised task competency and varied language fluence so that one robot was fluent-incompetent and the other disfluent-incompetent, the results indicated that multilingual participants' ratings between the two robot conditions in all perception categories did not significantly differ. This further suggests that task competency has a greater impact than language fluency on the perception of multilingual individuals in the perception categories of the robot's intelligence, reliability the robot meeting expectations, the robot being a good teammate, and participants' willingness to work with the robot again. Nevertheless, H4 needs to be further explored within the context of reliability because this perception category had a higher rating for the fluent-competent robot than for the disfluent-competent robot. More research needs to be conducted in order to better understand this result since the discrepancy in the ratings of reliability between the two robot condition groups weakens the argument that task competency has a greater impact than language fluency on the perception of reliability from the multilingual perspective. Further investigation is also necessary to explore the individual and joint impacts that task competency and language fluency can have on the perception of a robot's verbal competence from the perspective of multilingual individuals. Overall, H1, H2, and H3 were supported by all six perception categories (verbal competence, intelligence, reliability, robot meeting expectations, robot being a good teammate, and participant willingness to work with the robot again). H4 can be split into two parts where the first part stating that multilingual individuals did not significantly differ in their ratings between the fluent-competent and disfluent-competent robots was supported by all the perception categories except reliability. The second part of H4 predicting multilingual individuals would not significantly differ in their ratings between the fluent-incompetent and disfluent-incompetent robots was supported by all six perception categories (verbal competence, intelligence, reliability, robot meeting expectations, robot being a good teammate, and participant willingness to work with the robot again). It is essential to understand the complete implications of these findings through further research on the impact of varying language fluency and task competency on perceptions of the robot through a multilingual perspective.

In Study III, for RQ9, we observed a significant interaction between participant groups and varying robot conditions on the general perception (verbal competence, intelligence, and reliability) of the robot. RQ9 supports our hypothesis (H5) since multilingual individuals perceived both the disfluent robots (disfluent-competent and disfluent incompetent) to be more verbally competent than monolingual participants perceived them to be. This may imply that multilingual individuals are more empathetic towards non-native English speakers and disfluent English than monolingual individuals are in the perception of verbal competence due to their linguistic background. There were no significant interactions between participant groups in the fluent-incompetent robot condition on the perception of verbal competence; therefore, H6 was not supported for that perception category. For RQ10, RQ11, and RQ12 we found significant interactions between participant groups and varying robot conditions on ratings of the robot meeting expectations, robot as a teammate, and participants' willingness to work with the robot again. RQ9, RQ10, and RQ12 support H6 because monolingual participants rated the robot's intelligence, reliability, the robot meetings expectations, and participants' willingness to work with the robot again significantly higher in the fluent-incompetent robot condition than the multilingual participant group. Language fluency was not compromised in the fluent-incompetent robot, but the task competency was, so it can be reasoned that task competency may have impacted multilingual people's perception of the robot's intelligence, reliability, the robot meeting expectations, and participants' willingness to work with the robot again for the fluent-incompetent robot more than language fluency did. There were no significant interactions between participant groups on the disfluent-competent and disfluent-incompetent robot conditions on the perceptions of the robot's intelligence, reliability, the robot meeting expectations, and participants' willingness to work with the robot again; therefore, H5 was not supported in these four perception categories. For RQ11, the results suggest that multilingual individuals rated the fluent-competent robot as being a good teammate significantly higher than monolingual individuals. This was an unexpected finding as there was no statistical difference between multilingual and monolingual ratings of the robot in the fluent-competent condition for any of the other dependent variables. Additionally, there were no significant interactions between participant groups on the disfluent-competent and disfluent-incompetent robot conditions on the perception of the robot as a good teammate; therefore, H5 was not supported. The results also do not show any significant interaction between the participant groups on perceptions of the robot being a good teammate in the fluent-incompetent robot condition, which means that H6 was not supported in this perception category. This result may imply that both language fluency and task competency play a greater role in multilingual people perceiving how good of a teammate a fluent-competent robot is, but further research is necessary to discern these relationships. Overall, H5 was supported by the interaction between participant groups in one perception category (verbal competence) and H6 was supported by four perception categories (intelligence, reliability, robot meeting expectations, and participants' willingness to work with the robot again). Future investigations surrounding the impact of task competency on perceptions of robots and empathy towards non-native speakers from the perspectives of multilingual and monolingual individuals will be helpful in better understanding the implications of these results.

From the literature presented in Section 2.2 surrounding implications of varying language fluency in human-human interactions, we found that monolingual individuals may have negative biases and perceptions towards disfluent speakers and that multilingual individuals may demonstrate greater cognitive empathy towards non-native speakers. For example, studies found monolingual individuals to find non-native disfluent individuals as less intelligent, trustworthy, and verbally competent [48], [82]-[84]. In parallel, we can see that the results from our human-robot interaction studies may present similarities to the findings of disfluency perceptions by monolinguals. We observed that monolingual participants perceived disfluent robots more negatively compared to fluent robots in ratings of verbal competence and intelligence amongst other measures. This suggests that there may be negative biases towards disfluent robots by monolingual people. This potential correlation between the implications of varying language fluency in human and robot perceptions further highlights the need for additional exploration surrounding language fluency in robots and biases that humans may have towards them.

Chapter 9

FUTURE WORK

With my research being the first of its kind, there are substantial opportunities to expand on this work and further explore the implications of my findings. Although our studies produced some novel results, there were limitations that can serve as inspiration for future work. With globalization and the increasing exposure of different languages to human populations within and across societies, it is becoming more difficult to define and evaluate a person's linguistic background and language capabilities. Thus, it would be beneficial in future research to use standardized verbal language fluency assessments [85]-[89] to better distinguish between monolingual, bilingual, and multilingual participant groups. Another approach to control for confounding variables would be to develop a within-subjects design to explore the effects of varying language fluency and task competency on monolingual and multilingual participant groups. Our studies deployed a between-subjects design which may explain some variability within results due to the aforementioned linguistic inconsistencies that may have been present within participant groups and our data stemming from subjective ratings of the dependent measures. There are also various interpretations and attempted measures of speech fluency, some of which were addressed in Section 2.1, so it is necessary to investigate the outcomes of our study if language fluency was accounted for differently, such as by stutters, pauses, and hesitations rather than grammatical accuracy [37], [90]-[92].

Despite the measure of task competency in our studies being rather straightforward, there is potential to make it more robust in future work by lengthening the task so that the competencies of the robot are obvious to participants and expanding the scope of the task by making participants feel like they have some stake in the robot's performance. This would allow for a more practical understanding of how humans perceive robots when they interact with them. Additionally, there are considerable implications that our findings addressed that need to be further investigated. In our results, we reasoned that there might be some implications for language fluency and task competency having different levels of impact on perceptions of the varying robot conditions from a monolingual and multilingual perspective, but a clearer and more quantitative approach is vital for these inferences to be more concretely explored. Similarly, we know that multilingual individuals have higher cognitive empathy than their monolingual counterparts (see Section 2.2), but we cannot be sure that the reasons why multilingual perceptions of the robot differed from monolingual perceptions are attributed to multilingual individuals being more empathetic than monolingual individuals. Thus, we would recommend future research to include measures of cognitive empathy within their related studies to assess whether it plays a role in the perception of robots from the perspective of multilingual individuals [93]-[96].

Along with cognitive empathy, greater cultural empathy has also been found to correlate with multilingualism [71], [72]. Since certain languages are associated with certain cultural backgrounds, knowing the languages that people speak may provide insight into how those linguistic and cultural backgrounds may shape their perceptions of others. Therefore, collecting and analyzing data surrounding the different languages that multilingual individuals speak and how that may impact their perceptions of robots may also be valuable in future HRI research. Finally, with the existence of biases towards non-native humans established in human-human interaction studies, our work sets the foundation to explore if such biases exist in HRI as well.

Chapter 10

CONCLUSION

My thesis comprised of three main goals: understanding the perception of robots with varying language fluency and task competency from a monolingual perspective, multilingual perspective, and conducting a comparison between the robot perceptions by the monolingual and multilingual groups. In Study I, we implemented a human-robot collaboration task in the form of a guessing game in which the robot assumed different pairings of varying language fluency and task competency capabilities. We found there to be a significant difference between monolingual individuals' perceptions of robots in varying robot conditions in specific perception categories. For instance, monolingual participants may prioritize language fluency when perceiving verbal competence, both language fluency and task competency when perceiving intelligence and reliability, and task competency when determining if their expectations were met and if the robot was a good teammate. In Study II, we executed the same study design from Study I, except in this second study, we enlisted multilingual participants instead of monolingual participants. The results indicated a significant difference between multilingual individuals' perceptions of robots in varying robot conditions in certain perception categories. For example, multilinguals may be influenced by task competency more than language fluency in their perceptions of the robot's intelligence, ability to be a good teammate, meet expectations, and their willingness to work with the robot again while both language fluency and task competency may play a role in their perception of the robot's reliability.

To facilitate the research on understanding how monolingual participants' perception of a robot with varying language fluency and task competency compares to that of multilingual participants, we compared the results from Study I with Study II. The results showed that there were significant interactions between the participant groups and their perceptions of the robot in certain conditions. The implications of these findings may be that language fluency plays a greater role in the perception of a disfluent robot's verbal competence for monolingual participants than multilingual participants. In addition, monolingual individuals perceived fluent-incompetent robots more favorably in the perception categories of intelligence, reliability, meeting expectations, and willingness to work with the robot again.

The findings of this work underscore the importance of language fluency and task competency in the context of social HRI and will enable the development of more intelligent robots in the future. Following these studies, we hope for future research to investigate why perceptions between robot conditions and participant groups differed while also identifying and expanding on necessary robot capabilities to uphold positive human-robot interactions.

References

[1] A. A. Bahishti, “Humanoid robots and human society,” Advanced Journal of Social Science, vol. 1, no. 1, p. 60–63, Nov. 2017. [Online].

[2] T. Iqbal, S. Li, C. Fourie, B. Hayes, and J. A. Shah, “Fast online segmentation of activities from partial trajectories,” in ICRA, 2019.

[3] M. S. Yasar and T. Iqbal, “A scalable approach to predict multi-agent motion for human-robot collaboration,” in IEEE Robotics and Automation Letters (RA-L), 2021.

[4] M. M. Islam and T. Iqbal, “Multi-gat: A graphical attention-based hierarchical multimodal representation learning approach for human activity recognition,” in IEEE Robotics and Automation Letters (RA-L), 2021.

[5] V. V. Unhelkar, S. Li, and J. A. Shah, “Semi-supervised learning of decision-making models for human-robot collaboration,” in Conference on Robot Learning, 2020.

[6] T. Iqbal and L. D. Riek, “Coordination dynamics in multi-human multi-robot teams,” IEEE Rob. and Auto. Letters (RA-L), 2017.

[7] T. Kanda, H. Ishiguro, M. Imai, and T. Ono, “Development and evaluation of interactive humanoid robots,” Proceedings of the IEEE, vol. 92, no. 11, pp. 1839–1850, 2004.

[8] T. Iqbal and L. D. Riek, “Human robot teaming: Approaches from joint action and dynamical systems,” Humanoid Robotics, 2017.

[9] M. M. Islam and T. Iqbal, “Hamlet: A hierarchical multimodal attention-based human activity recognition algorithm,” in 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2020, pp. 10 285–10 292.

[10] M. de Graaf, “Living with robots: investigating the user acceptance of social robots in domestic environments,” Ph.D. dissertation, University of Twente, June 2015.

[11] M. M. Islam and T. Iqbal, “Mumu: Cooperative multitask learning-based guided multimodal fusion,” Proceedings of the AAAI Conference on Artificial Intelligence, pp. 1043–1051, 2022.

[12] M. S. Yasar and T. Iqbal, “Robots that can anticipate and learn in human-robot teams,” in 2022 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI), 2022.

[13] Y. Wang and J. E. Young, “2022 17th acm/ieee international conference on human-robot interaction (hri),” 2014.

[14] T. Iqbal, S. Rack, and L. D. Riek, “Movement coordination in human-robot teams: A dynamical systems approach,” IEEE Transactions on Robotics, vol. 32, no. 4, pp. 909–919, 2016.

[15] S. M. Anzalone, S. Boucenna, S. Ivaldi, and M. Chetouani, “Evaluating the engagement with social robots,” International Journal of Social Robotics, vol. 7, pp. 465–478, 2015.

[16] I. Leite, C. Martinho, and Paiva, “Evaluating the engagement with social robots,” International Journal of Social Robotics, vol. 5, pp. 291–308, 2013.

[17] F. Eyssel, D. Kuchenbrandt, F. Hegel, and L. de Ruiter, “Activating elicited agent knowledge: How robot and user features shape the perception of social robots,” in 2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication, 2012, pp. 851–857.

[18] S. Nikolaidis, M. Kwon, J. Forlizzi, and S. Srinivasa, “Planning with verbal communication for human-robot collaboration,” vol. 7, no. 3, 2018. [Online].

[19] A. Marin Vargas, L. Cominelli, F. Dell’Orletta, and E. P. Scilingo, “Verbal communication in robotics: A study on salient terms, research fields and trends in the last decades based on a computational linguistic analysis,” Frontiers in Computer Science, vol. 2, 2021.

[20] J. Woo, J. Botzheim, and N. Kubota, “Conversation system for natural communication with robot partner,” in 2014 10th France-Japan/ 8th Europe-Asia Congress on Mecatronics (MECATRONICS2014-Tokyo), 2014, pp. 349–354.

[21] H. N. Green, M. M. Islam, S. Ali, and T. Iqbal, “ispy a humorous robot: Evaluating the perceptions of humor types in a robot partner,” in AAAI Spring Symposium on Putting AI in the Critical Loop: Assured Trust and Autonomy in Human-Machine Teams, 2022.

[22] S. Zhao, “Humanoid social robots as a medium of communication,” New Media & Society, vol. 8, no. 3, pp. 401–419, 2006.

[23] H. N. Green, M. M. Islam, S. Ali, and T. Iqbal, “Who’s laughing nao? examining perceptions of failure in a humorous robot partner,” in ACM/IEEE International Conference on Human-Robot Interaction (HRI), 2022, p. 313–322.

[24] C. Breazeal, K. Dautenhahn, and T. Kanda, “Social robotics,” Springer handbook of robotics, pp. 1935–1972, 2016.

[25] E. C. Grigore, A. Pereira, I. Zhou, D. Wang, and B. Scassellati, “Talk to me: Verbal communication improves perceptions of friendship and social presence in human-robot interaction,” in Intelligent Virtual Agents: 16th International Conference, IVA 2016, Los Angeles, CA, USA, September 20–23, 2016, Proceedings 16. Springer, 2016, pp. 51–63.

[26] A. Gluszek and J. F. Dovidio, “Speaking with a nonnative accent: Perceptions of bias, communication difficulties, and belonging in the united states.”

[27] A. Pavlenko and A. Blackledge, Negotiation of identities in multilin-gual contexts. Multilingual Matters, 2004, vol. 45.

[28] O. Kang, D. Rubin, and S. Lindemann, “Mitigating us undergraduates’ attitudes toward international teaching assistants,” Tesol Quarterly, vol. 49, no. 4, pp. 681–706, 2015.

[29] S. van Waveren, E. J. Carter, and I. Leite, "Take one for the team: The effects of error severity in collaborative tasks with social robots," in Proceedings of the 19th ACM International Conference on Intelligent Virtual Agents, 2019, pp. 151–158.

[30] N. Walker, C. Mavrogiannis, S. Srinivasa, and M. Cakmak, "Influencing behavioral attributions to robot motion during task execution," in Proceedings of the 5th Conference on Robot Learning, ser. Proceedings of Machine Learning Research, A. Faust, D. Hsu, and G. Neumann, Eds., vol. 164. PMLR, 08–11 Nov 2022, pp. 169–179. [Online]. Available: https://proceedings.mlr.press/v164/walker22a.html

[31] M. H¨aring, D. Kuchenbrandt, and E. Andr´e, "Would you like to play with me? how robots' group membership and task features influence human-robot interaction," in Proceedings of the 2014 ACM/IEEE International Conference on Human-Robot Interaction, ser. HRI '14. New York, NY, USA: Association for Computing Machinery, 2014, p. 9–16. [Online]. Available: https://doi.org/10.1145/2559636.2559673

[32] B. Kim, K. S. Haring, H. J. Schellin, T. N. Oberley, K. M. Patterson, E. Phillips, E. J. de Visser, and C. C. Tossell, "How early task success affects attitudes toward social robots," in Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction, ser. HRI '20. New York, NY, USA: Association for Computing Machinery, 2020, p. 287–289. [Online]. Available: https://doi.org/10.1145/3371382.3378241

[33] M. F. Jung, J. J. Lee, N. DePalma, S. O. Adalgeirsson, P. J. Hinds, and C. Breazeal, "Engaging robots: Easing complex human-robot teamwork using backchanneling," in Proceedings of the 2013 Conference on Computer Supported Cooperative Work, ser. CSCW '13. New York, NY, USA: Association for Computing Machinery, 2013, p. 1555–1566. [Online]. Available: https://doi.org/10.1145/2441776.2441954

[34] S. van Waveren, E. J. Carter, and I. Leite, "Take one for the team: The effects of error severity in collaborative tasks with social robots," in Proceedings of the 19th ACM International Conference on Intelligent Virtual Agents, ser. IVA '19. New York, NY, USA: Association for Computing Machinery, 2019, p. 151–158. [Online]. Available: https://doi.org/10.1145/3308532.3329475

[35] E. J. Carter, L. M. Hiatt, and S. Rosenthal, "You're delaying my task?! the impact of task order and motive on perceptions of a robot," in 2022 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI), 2022, pp. 304–312.

[36] M. Salem, G. Lakatos, F. Amirabdollahian, and K. Dautenhahn, "Would you trust a (faulty) robot? effects of error, task type and personality on human-robot cooperation and trust," ser. HRI '15. New York, NY, USA: Association for Computing Machinery, 2015, p. 141–148. [Online]. Available: https://doi.org/10.1145/2696454.2696497

[37] P. Lennon, "Investigating fluency in efl: A quantitative approach."

[38] B. VanPatten and J. Williams, "Skill acquisition theory Robert dekeyser," in Theories in Second Language Acquisition. Routledge, 2014, pp. 106–124.

[39] J. Kormos and M. D´enes, "Exploring measures and perceptions of fluency in the speech of second language learners," System, vol. 32, no. 2, pp. 145–164, 2004.

[40] H. R. Bosker, A.-F. Pinget, H. Quen´e, T. Sanders, and N. H. de Jong, "What makes speech sound fluent? the contributions of pauses, speed and repairs," Language Testing, vol. 30, no. 2, pp. 159–175, 2013.

[41] M. J. White and Y. Li, "Second-language fluency and person perception in china and the united states," Journal of Language and Social Psychology, vol. 10, no. 2, pp. 99–113, 1991.

[42] T. Bongaerts, C. van Summeren, B. Planken, and E. Schils, "Age and ultimate attainment in the pronunciation of a foreign language," Studies in second language acquisition, pp. 447–465, 1997.

[43] A. Hagi-Mohamed, "Perceptions of nonnative english-speaking graduate teaching assistants: Identity issues, successes, and challenges in the field of tesl/tesol," 2018.

[44] D. Barona, "Native and non-native speakers' perceptions of non-native accents," LL Journal, vol. 3, no. 2, pp. 1–18, 2008.

[45] M. J. Munro and T. M. Derwing, "The functional load principle in esl pronunciation instruction: An exploratory study," System, vol. 34, no. 4, pp. 520–531, 2006.

[46] D. L. Rubin and K. A. Smith, "Effects of accent, ethnicity, and lecture topic on undergraduates' perceptions of nonnative english-speaking teaching assistants," International journal of intercultural relations, vol. 14, no. 3, pp. 337–353, 1990.

[47] R. W. Todd and P. Pojanapunya, "Implicit attitudes towards native and non-native speaker teachers," System, vol. 37, no. 1, pp. 23–33, 2009.

[48] A. James and J. Leather, Second-language speech: structure and process. Walter de Gruyter, 2011, vol. 13.

[49] C. Groff, W. Zwaanswijk, A. Wilson, and N. Saab, "Language diversity as resource or as problem? educator discourses and language policy at high schools in the netherlands," International Multilingual Research Journal, pp. 1–19, 2023.

[50] R. Jackson and T. Williams, "On perceived social and moral agency in natural language capable robots," 03 2020.

[51] M. Scheutz, R. Cantrell, and P. Schermerhorn, "Toward humanlike task-based dialogue processing for human robot interaction," AI Magazine, vol. 32, no. 4, pp. 77–84, Dec. 2011. [Online]. Available: https://ojs.aaai.org/aimagazine/index.php/aimagazine/article/view/2381

[52] Y. Bisk, D. Yuret, and D. Marcu, "Natural language communication with robots," in Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. San Diego, California: Association for Computational Linguistics, June 2016, pp. 751–761. [Online]. Available: https://aclanthology.org/N16-1089

[53] N. Wang, D. V. Pynadath, S. G. Hill, and A. P. Ground, “Building trust in a human-robot team with automatically generated explanations,” in Proceedings of the interservice/industry training, simulation and education conference (I/ITSEC), vol. 15315, 2015, pp. 1–12.

[54] E. Wenger, M. Bronckers, C. Cianfarani, J. Cryan, A. Sha, H. Zheng, and B. Y. Zhao, “” hello, it’s me”: Deep learning-based speech synthesis attacks in the real world,” in Proceedings of the 2021ACM SIGSAC Conference on Computer and Communications Security, 2021, pp. 235–251.

[55] A. St. Clair and M. Mataric, “How robot verbal feedback can improve team performance in human-robot task collaborations,” ser. HRI ’15. New York, NY, USA: Association for Computing Machinery, 2015, p. 213–220. [Online]. Available: https://doi.org/10.1145/2696454.2696491

[56] A.-F. Pinget, H. R. Bosker, H. Quen´e, and N. H. De Jong, “Native speakers’ perceptions of fluency and accent in l2 speech,” Language Testing, vol. 31, no. 3, pp. 349–365, 2014.

[57] N. G¨urb¨uz, “Understanding fluency and disfluency in non-native speakers’ conversational english,” 2017.

[58] M. Bakti, “Speech disfluencies in simultaneous interpretation,” De Crom, Dies (ed.), 2009.

[59] T. J. Setiyorini, P. Dewi, and E. S. Masykuri, “The grammatical error analysis found in students’ composition,” Journal Lensa, vol. 10, no. 2, 2020.

[60] Y. B. Al-Shujairi and H. Tan, “The grammatical error analysis found in students’ composition,” International Journal of Education Literacy Studies, vol. 5, no. 4, 2017.

[61] K. Nonkukhetkhong, “Grammatical errors analysis of the first year english major students, udon thani rajabhat university,” in ACLL2013 Conference Proceedings. IAFOR, 2013.

[62] Amazon Web Services, Inc, “Amazon polly,” https://aws.amazon.com/polly/, 2021, retrieved: June, 2021.

[63] Fillmore, Charles J. "On fluency." *Individual differences in language ability and language behavior*. Academic Press, 1979. 85-101.

[64] Brumfit, Christopher. *Communicative methodology in language teaching: The roles of fluency and accuracy*. Vol. 129. Cambridge: Cambridge University Press, 1984.

[65] Nation, Paul. "Improving speaking fluency." *System* 17.3 (1989): 377-384.

[66] Swain, Merrill. "Communicative competence: Some roles of comprehensible input and comprehensible output in its development." *Input in second language acquisition* 15 (1985): 165-179.

[67] Cenoz, Jasone. “Defining Multilingualism.” Annual Review of Applied Linguistics, vol. 33, 2013, pp. 3–18.

[68] Ellis, Elizabeth. "Defining and investigating monolingualism." *Sociolinguistic Studies* 2.3 (2008): 311-330.

[69] Richards, Jack C., and Richard W. Schmidt. *Longman dictionary of language teaching and applied linguistics*. Routledge, 2013.

[70] Dictionary, Merriam-Webster. *Merriam-webster dictionary*. *On-line at http://www. mw. com/home. htm* 8.2 (2002).

[71] Dewaele, Jean-Marc. "Bilingualism and multilingualism." *The international encyclopedia of language and social interaction* (2015): 1-11.

[72] Dewaele, Jean–Marc. "The link between foreign language classroom anxiety and psychoticism, extraversion, and neuroticism among adult bi-and multilinguals." *The Modern Language Journal* 97.3 (2013): 670-684.

[73] Dewaele, Jean-Marc, and Li Wei. "Is multilingualism linked to a higher tolerance of ambiguity?." *Bilingualism: Language and Cognition* 16.1 (2013): 231-240.

[74] Rothman, Jason. "Linguistic epistemology and the notion of monolingualism." *Sociolinguistic Studies* 2.3 (2008): 441-458.

[75] Kachru, Braj B. "Bilingualism." *Annual Review of Applied Linguistics* 1 (1980): 2-18.

[76] Clyne, Michael. "Multilingualism." *The handbook of sociolinguistics* (2017): 301-314.

[77] Dewaele, Jean-Marc, and Li Wei. "Multilingualism, empathy and multicompetence." *International Journal of Multilingualism* 9.4 (2012): 352-366.

[78] Bianco, Joseph Lo. *National policy on languages*. Canberra: Australian Government Publishing Service, 1987.

[79] Lambert, Wallace E., and G. Richard Tucker. "Bilingual education of children: The St. Lambert experiment." (1972).

[80] Hawkins, Eric W. "Foreign language study and language awareness." *Language awareness* 8.3-4 (1999): 124-142.

[81] Met, Myriam, and Michael Byram. "Standards for foreign language learning and the teaching of culture." *Language Learning Journal* 19.1 (1999): 61-68.

[82] Volz, Sarah, Marc-André Reinhard, and Patrick Müller. "Why don't you believe me? Detecting deception in messages written by nonnative and native speakers." *Applied Cognitive Psychology* 34.1 (2020): 256-269.82].

[83] Storova, Lucie. *Non-native Speakers in Call Centers. The Influence of Language on the Perceived Service Quality (Available on Internet)*. MS thesis. Hanken School of Economics, 2010.

[84] Byers-Heinlein, Krista, et al. "Monolingual and bilingual children's social preferences for monolingual and bilingual speakers." *Developmental science* 20.4 (2017): e12392.

[85] Huhta, Ari, et al. "Fluency in language assessment." *Fluency in L2 learning and use* (2019): 129-145.

[86] Bethlehem, Deborah, Janet De Picciotto, and Nola Watt. "Assessment of verbal fluency in bilingual Zulu-English speakers." *South African Journal of Psychology* 33.4 (2003): 236-240.

[87] Aita, Stephen L., et al. "Executive, language, or both? An examination of the construct validity of verbal fluency measures." *Applied Neuropsychology: Adult* (2018).

[88] Abwender, David A., et al. "Qualitative analysis of verbal fluency output: Review and comparison of several scoring methods." *Assessment* 8.3 (2001): 323-338.

[89] Rosselli, Monica, et al. "A cross-linguistic comparison of verbal fluency tests." *International Journal of Neuroscience* 112.6 (2002): 759-776.

[90] Chambers, Francine. "What do we mean by fluency?." *System* 25.4 (1997): 535-544.

[91] Hilton, Heather. "Oral fluency and spoken proficiency: Considerations for research and testing." *Measuring L2 proficiency: Perspectives from SLA* 27 (2014): 53.

[92] Kormos, Judit, and Mariann Dénes. "Exploring measures and perceptions of fluency in the speech of second language learners." *System* 32.2 (2004): 145-164.

[93] Spaulding, Shannon. "Cognitive empathy." *The Routledge handbook of philosophy of empathy*. Routledge, 2017. 13-21.

[94] Khanjani, Zeinab, et al. "Comparison of cognitive empathy, emotional empathy, and social functioning in different age groups." *Australian Psychologist* 50.1 (2015): 80-85.

[95] Murphy, Brett A., and Scott O. Lilienfeld. "Are self-report cognitive empathy ratings valid proxies for cognitive empathy ability? Negligible meta-analytic relations with behavioral task performance." *Psychological Assessment* 31.8 (2019): 1062.

[96] Ze, Oksana, Patrizia Thoma, and Boris Suchan. "Cognitive and affective empathy in younger and older individuals." *Aging & mental health* 18.7 (2014): 929-935.

[97] Prasad, Rajkishore, Hiroshi Saruwatari, and Kiyohiro Shikano. "Robots that can hear, understand and talk." *Advanced Robotics* 18.5 (2004): 533-564.

[98] Chung, Hanna, Hyunmin Kang, and Soojin Jun. "Verbal anthropomorphism design of social robots: Investigating users' privacy perception." *Computers in Human Behavior* (2023): 107640.

[99] Iqbal, Tariq, and Laurel D. Riek. "A method for automatic detection of psychomotor entrainment." *IEEE Transactions on affective computing* 7.1 (2015): 3-16.

[100] Iqbal, Tariq, Maryam Moosaei, and Laurel D. Riek. "Tempo adaptation and anticipation methods for human-robot teams." *RSS, Planning HRI: Shared Autonomy Collab. Robot. Workshop*. 2016.

[101] Iqbal, Tariq, and Laurel D. Riek. "Temporal anticipation and adaptation methods for fluent human-robot teaming." *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2021.

[102] Islam, Md Mofijul, Reza M. Mirzaiee, Alexi Gladstone, and Tariq Iqbal. "Caesar: An embodied simulator for generating multimodal referring expression datasets." *Advances in Neural Information Processing Systems* 35 (2022): 21001-21015.

[103] Islam, Md Mofijul, Mohammad SAMIN Yasar, and Tariq Iqbal. "MAVEN: A Memory Augmented Recurrent Approach for Multimodal Fusion." *IEEE Transactions on Multimedia* (2022).

[104] Walker, Nick, et al. "Human perceptions of a curious robot that performs off-task actions." Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction. 2020.

[105] Neuteboom, Sven Y., and Maartje MA de Graaf. "People's perceptions of gendered robots performing gender stereotypical tasks." Social Robotics: 13th International Conference, ICSR 2021, Singapore, Singapore, November 10–13, 2021, Proceedings 13. Springer International Publishing, 2021.

[106] McGinn, Conor, and Ilaria Torre. "Can you tell the robot by the voice? An exploratory study on the role of voice in the perception of robots." 2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI). IEEE, 2019.