THE INFLUENCE OF ETHNOCENTRISM ON TRAIT ATTRIBUTION PERCEPTIONS TOWARDS STANDARD AMERICAN, FOREIGN AND MACHINE ACCENTS

A THESIS SUBMITTED TO THE GRADUATE DIVISION OF THE UNIVERSITY OF HAWAI'I AT MĀNOA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF ARTS

IN

COMMUNICOLOGY

December 2020

by

Marguerite Liechtenstein

Thesis Committee:

Min-Sun Kim (Chair)

Jinguang Zhang

Soo Shin

ACKNOWLEDGEMENTS

I would like to express my sincerest gratitude to Dr. Min-Sun Kim (Committee Chair), who has continuously motivated and inspired me to work my hardest. Her enthusiasm and guiding hand have allowed me to be creative in telling my "story" from beginning to end. Without her patience, understanding, and brilliant notion of adding machine voices, this thesis would not have been possible.

I would also like to thank my committee member, Dr. Soo Shin, who fueled my interest in human and machine interactions. Dr. Shin has given me helpful information relating to machine interactions. Her class on human communication and Tech has made me even more excited about my thesis's machine interaction aspect.

My gratitude also goes to my committee member, Dr. Jinguang Zhang, whose wisdom in analyzing the data was invaluable in completing this thesis. For his support in analyzing and interpreting the results, I am incredibly grateful. His passion for statistics always incites excitement for the use of SPSS and R. His enthusiasm for my thesis topic was also very motivational.

Finally, many thanks to my family, who made it possible to join this Master's program and continuously supported me during my journey. A special thanks to my father, who was very enthusiastic about my thesis topic. He has given me many ideas and examples from his own experiences. The beginning quote of my introduction was contributed by him. The idea came to him late one evening and stole his sleep. My family's support and encouragement were invaluable.

ABSTRACT

The main purpose of this study was to examine the influence of ethnocentrism on American-English speakers' ratings of standard American, foreign, and computer-synthesized accents on the task and social dimension of trait attributions. The 351 MTurk participants were asked to listen to nine accents separately and consequently fill out the measure for perceptions of trait attributions (task and social dimensions) and perceived voice impressions. This was followed by the generalized ethnocentrism scale and a demographic questionnaire. The results indicate that in general American-English speakers rate the standard American accent higher on the task and social dimensions of trait attributions than a low prestige foreign, followed by computer-synthesized accents. Secondly, the findings indicate that as American-English speakers' ethnocentric tendencies increase, the computer-synthesized accent is increasingly perceived more positively than a low prestige foreign accent on the task and social dimensions of trait attributions. These results highlight the role of high levels of ethnocentrism on perceptions of foreign accents, which is perceived significantly less positive than machine accents. As previous notions of in-group favoritism among humans are challenged in our increasingly technological world, future research needs to engage in theoretical and ethical implications of communication with intelligent machines.

Keywords: Ethnocentrism, Attribution Perceptions, task dimension, social dimension, standard accent, foreign accent, computer-synthesized voice, text-to-speech system.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

"The rain in Spain stays mainly in the plain" is the famous lyric played in the movie adaptation of My Fair Lady (Warner & Cukor, 1964). The plot revolves around a phonology professor trying to teach a woman possessing a cockney-accent to speak with a proper English accent, and having the upper class accept her. This movie perfectly illustrates the social influence accents can have on status and how 'idealized' pronunciation can change a listeners' perception of a speaker. Hence, the human accent, meaning phonological or intonation features dependent on the space we inhabit (Lippi-Green, 1994), instigates various language attitudes. An accent can be the source of strong social influence, wherein a listener's inferences about a speakers' accent can influence the listeners' attitudes towards that speaker (Cargile et al., 1994).

Accented speech, focusing on English-accented speech as most of the research was conducted in that area, has put forth two distinctions, namely standard and non-standard accent. The former is spoken by the dominant culture and held in high regard, whereas the latter encompasses dialect, regional and foreign accents; and is considered a deviation from the norm (Dragojevic, Giles, & Watson, 2013; Burgoon, Guerrero, & Floyd, 2010). Hence, research on language attitudes infers that listeners make distinctions based on their perceptions of accented speech.

So far, no study has examined computer-synthesized accents in comparison to standard and non-standard accents. Human-machine interaction has developed extensively over the years, wherein we interact with advanced AI or computer-synthesized machine voices regularly in our daily life. For example, either through Google Voice, Alexa, Echo, NUGU, and older versions of text-to-speech programs on laptops, i.e., Microsoft, Windows, or Apple. Technology has achieved remarkably human-like text-to-speech voices, one of whom is the text-to-speech system of an Apple MacBook, which the current study uses. To that extent, doubly disembodied language concerns itself with computer-synthesized voices, which considers that humans are

predisposed to process voices from an evolutionary perspective and attribute social characteristics (Lee & Nass, 2004), even computer-synthesized voices (Nass & Gong, 2000).

Research of accented speech perceptions often includes Accent Prestige Theory, wherein standard accents are perceived more prestigious on two dimensions, namely, status and solidarity (Giles, 1970 as cited in Fuertes, Potere & Ramirez, 2002). Although status is continuously rated higher for standard accents, Fuertes, Potere, and Ramirez (2002) stated that solidarity offers inconsistent findings. Previous findings indicate that accented speech is rated lower on dimensions of attractiveness, expertness, and trustworthiness (Fuertes & Gelso, 2000); dimensions of social attractiveness and prestige (Coupland & Bishop, 2007); and less credible (Lev-Ari & Keysar, 2010). As these variables have somewhat consistently indicated perceptions of trait attributions toward accented speech, they were used in this study under the name of task (i.e., competence, intelligence, etc.) and social (i.e., pleasantness, likeability, etc.) dimensions.

In recent years, research has also focused on attitudes towards computer-synthesized machine voices. Research indicates that listeners' perceptions of computer-synthesized speech are less preferred compared to human voices. Computer-synthesized voices are rated lower on truthfulness and knowledgeability (Stern, Mullennix, Dyson & Wilson, 1999); and on trust, likability, and credibility (Rau, Li & Li, 2009). These findings indicate that the task and social dimensions of trait attributions are readily applicable to computer-synthesized machine voices.

Perceptions of trait attributions can be investigated from an ethnocentrism-based approach as language attitudes can influence the categorization of accented speech. Ethnocentrism operates based on categorizing the social world into in- and out-groups (Hopkins, 2015). Additionally, members of the in-group epitomize morality and superiority compared to out-group members. The latter exuding the opposite, such as immorality and weakness (Neuliep & McCroskey, 1997), consequently leading to various attitudes towards out-group members. The

attitudes towards out-group members have been examined extensively, within the context of research concentrating on attitudes towards accented speech (Neuliep & McCroskey, 1997).

A study focusing on accents influencing the group categorization process was conducted by Bresnahan, Ohashi, Nebashi, Liu, and Shearman (2002). Their study indicated that individuals with a stronger ethnic identity rated American English more favorable on dynamism, status, and attractiveness than foreign accents (Bresnahan et al., 2002). Additionally, research on own-accent bias found that the own-accent or an accent similar to oneself evokes favorable and trustworthy perception (Coupland & Bishop, 2007; Lev-Ari & Keysar, 2010). Applied research continuously supports own-accent bias wherein infants show preferences early on in development (Kinzler, Corriveau & Harris, 2011). Additionally, Bestelmeyer, Belin, and Ladd (2015) stated that due to neural activity in specific parts of the brain, own-accent bias could be explained by affective processing, meaning emotional reactions. The research indicates that categorization, and by extension, ethnocentrism is important in accent research.

Research indicates that individuals will attribute human characteristics and social rules to nonhuman entities during interactions. More specifically, findings indicate that individuals will attribute same-ethnicity favoritism when taking suggestions from computers (Nass, Isbister & Lee, 2000). Furthermore, in- and out-group categorization can be seen with human and computer interactions and perceiving in-group entities as friendlier and more intelligent (Nass, Fogg & Moon, 1996). According to Gong (2008), white participants with high levels of prejudice selected computer-synthesized robot characters over black human characters as virtual friends.

To the best of my knowledge, no other study aimed to investigate ethnocentrism's influence on perceptions of trait attributions (task and social dimensions) towards standard American accents, foreign accents, and computer-synthesized accents. The novelty of this research is that not only does it compare these three established accents (standard accents,

foreign accents, and computer-synthesized accents) on the task and social dimensions of trait attributions but also associates it directly with ethnocentrism.

This study investigates how ethnocentrism influences the perceptions of trait attributions towards standard American accents, foreign accents, and computer-synthesized machine accents. To do this, we firstly examined human accents and machine accents; and how accents will be defined. Then, an overview of previous research of perceptions of trait attributions in human and machine accents is examined. Subsequently, we will examine the research of the influence of ethnocentrism on perceptions of trait attributions, specifically focusing on ethnocentrism, own-accent bias, and the categorization of nonhuman entities.

CHAPTER 2: LITERATURE REVIEW

Human Accent

Previous research has investigated human accents at length within various fields of study, including psychology, linguistics, and communication. Hence, various theoretical frameworks have been forwarded to explain attitudes and perceptions towards accented speech. Language is not solely used to refer information but is also a compelling social force, according to Cargile et al. (1994). Another individual's language features can be utilized to make inferences about their presumed competencies, attributes, and beliefs. As such, we might assume that a particular accent (e.g., British) portrays an individual as being 'cultured' or 'refined.' Language can also influence individuals' decisions regarding opportunities (e.g., employment) and social welfare. Moreover, Cargile et al. (1994) stated that it is essential to study language attitudes as social interactions often use cues from language behaviors, which are considered the most salient. They further explained that it is crucial to establish how the speaker is perceived to determine listeners' attitudes.

Accessibility to and practice of language attitudes can, according to Cargile et al. (1994), be affected by three components, namely goals, mood, and expertise. Firstly, a hearer can attend to various language behaviors depending on his/her goals. Consequently, evaluations or behaviors towards a speaker are guided by various language attitudes (i.e., goals). Thus, if a listener's goal is to employ an individual, the interviewee's particular language behavior (i.e., British accent) may elicit a particular attitude toward said speaker (i.e., educated or cultured), consequently increasing the chances of getting hired. Secondly, attitudes are naturally affective and not solely conceptualized by beliefs and cognitions; thus, hearers' mood can influence attitudes and perceptions. Thirdly, a hearers' level of expertise in a situation might evoke different language attitudes. Thus, the level of expertise will allow an expert to focus on other cues as cognitively; it would be less effortful to understand accented speech. As accents would

be considered a language behavior that can elicit particular language attitudes from a listener, it is essential to define an accent and how they will be used in this study.

An accent can generally encompass various aspects of speech, such as pitch or stress, when observed through phonetics. Although, more often, it refers to phonological or intonational characteristics that differ depending on the social or geographic space we inhabit. An accent can also refer to particular phonology or intonation, which might remain with a non-native speaker of any language (Lippi-Green, 1994). In this study, an accent refers to accented English speech of native and non-native speakers and computer-synthesized accents (i.e., English US).

The dominant culture of a country speaks in a standard accent. These accents are usually perceived as prestigious and the standard. Another term used in the United Kingdom is 'received pronunciation' and refers to the 'official' or 'first-class' accent spoken in said country. This received pronunciation is perceived as signifying an individual having intellect and distinction. Standard accents consist of a combination of accents, prosody, grammar, and vocabulary that are perceived as prestigious, even by individuals who possess non-standard accents. Moreover, the speech form and writing usually represented in grammar books and dictionaries reflect this 'ideal' accent (Dragojevic, Giles, & Watson, 2013; Burgoon, Guerrero, & Floyd, 2010).

Thus, a non-standard accent refers to any speech that deviates from the 'idealized' standard accent and includes dialects and regional and foreign accents (Dragojevic et al. 2013). This investigation focuses on foreign accents, which, according to Munro and Derwing (1995), is defined as "non-pathological speech that differs in some noticeable respect from native speaker pronunciation norm" (p. 289). Moreover, individuals who learn another language later in life would inevitably have an accent due to listeners' inability to recognize or understand the accented pronunciation of certain phonetic (speech sounds) segments, words, or sentences. However, even when the accent is understood, communication can still be impacted due to certain prejudices (Munro & Derwing, 1995). This prejudice might be due to an individual's

accent providing or conveying background information on ethnicity, socioeconomics, and geography (Labov, 2006 as cited in Bestelmeyer, Belin & Ladd, 2015). Furthermore, listeners are sensitive to phonetic (a speech sound) and prosody (language stress and intonation) variations accented speech provides and subsequently allows for social judgments (e.g., personality) (Dailey, Giles, & Jansma, 2005; Magen, 1998).

Machine Accent

In our daily lives, we partake in an increased amount of human and machine speech due to the evolution of technology, such as Alexa, Echo, Google Voice, NUGU, and other devices. However, before overviewing machine entities' research, the meaning of machine accents for this study must be defined. To that extent, Lee and Nass (2004) explored doubly disembodied language, which is an evolvement of disembodied language. Disembodied language consists of written language and pre-recorded human speech. For the latter form, the virtual speakers' perception is affected by vocalic (e.g., speech rate, volume, pitch) and linguistic cues received by the listeners' (Lee & Nass, 2004).

As technology has evolved, a new theory was developed to include computer-synthesized machine voices, which was termed doubly disembodied language (Lee & Nass, 2004). This term has been chosen as the synthesized voice would be doubly disembodied. The synthesized voice does not have an actual speaker when the message is interpreted (i.e., disembodied language). The double disembodiment occurs due to a disassociation between the source and the speech characteristic (Lee & Nass, 2004). In other words, there is firstly no speaker physically present when the message is received. Secondly, the computer-synthesized voice has no association between the speaker (i.e., source) and characteristics of speech.

Additionally, it has been proposed that there is an evolutionary response to people's reactions and attitudes to computer-synthesized voices. The human mind has not evolved to distinguish computer-synthesized machine voices when contrasted to human voices in their

responses (Lee & Nass, 2004; Nass & Gong, 2000). Humans are predisposed to discriminate speech from other sounds through the left-brain hemisphere (i.e., right ear), whereas the right brain hemisphere is more likely to pick up other sounds (e.g., music). Hence, the predisposition provides advantages of processing speech, such as native and foreign languages, backward speech, and nonsensical syllables. As the intonation and pronunciation of computer-synthesized machine voices are a reminder of non-humanness, Nass and Gong (2000) questioned whether people process evolutionary vocal patterns of social characteristics such as gender, personality, and emotion similarly as they would humans.

A study examining the "gender" of machine voices and their effect on individuals' perceptions of, and conformity to, recommendations from the computer, attempted to illustrate computer-synthesized speech evaluation. According to Lee, Nass, and Brave (2000), more conformity is elicited by males than females in persuasion. Moreover, communicators are considered to have more competence and social status if they are male rather than female communicators. The findings indicated that individuals treated the computer-synthesized voice as real humans, and thus the gender of the voice influenced the personality perceptions. Moreover, societal rules regarding gender were attributed to machine voices. Hence, the results indicated that females were less persuasive than males. Additionally, males and females preferred their same gender voices when it came to social attraction, conformity, and trust (Lee et al. 2000). Thus, on an evolutionary basis, it could be suggested that individuals seem to process computer-synthesized speech similarly to human speech, as questioned by Nass and Gong (2000).

Additionally, a study illustrated that individuals would anthropomorphize robots differently depending on whether it is a spontaneous reaction in a social context or the careful consideration of robots' abstract conceptualization (Fussell, Kiesler, Setlock & Yew, 2008). The study is supportive of the notion that people attribute human characteristics to non-human

entities. Thus, it could be argued that this further supports the idea of listeners' reacting similarly to accents, whether human or not. The human and robot interviewers were either polite or impolite. The findings indicate that in their spontaneous reaction to robots, human characteristics were attributed as often and quickly to the robot and human interviewers.

In continuation, both interviewers were perceived positively as being curious and organized rather than distractible or nervous (Fussell et al., 2008). Furthermore, the impolite robot and human interviewers were regarded as obnoxious and rude, whereas the polite interviewers were viewed as polite and friendly. The findings provide strong evidence that people assess robot personalities similarly to humans.

Although, Fussell et al. (2008) indicated that when people carefully considered their responses, they viewed robot interviewers more mechanically. For instance, people did not believe that robots had moods, feelings, or experienced frustrations. Thus, it is important to note that the findings indicate that such human attributions happen primarily through spontaneous reactions rather than carefully considered responses. Nonetheless, the findings support speech being processed via an evolutionary predisposition, wherein people will anthropomorphize non-human entities.

There are currently three types of speech technologies: speech and audio coding, text-to-speech synthesis, speech recognition, and spoken language understanding (Kamm, Walker, & Rabiner, 1997). This study will be using and defining text-to-speech synthesis. The text-to-speech synthesis, also referred to as the TTS system, mainly refers to synthesized speech that relies on a machine speaking information to a user (Kamm et al., 1997). Additionally, a TTS system is evaluated on two dimensions, namely intelligibility and naturalness of the produced speech. While current systems are evaluated as being in the fair-to-good range concerning quality on these dimensions, their quality and prosody still do not match natural speech (Kamm et al., 1997). Moreover, Parr (2012) stated that a TTS system decodes any printed text format and

reads it back to an individual aloud using a computer-synthesized voice. A computer-synthesized voice is an output of speech that is identifiable as not being produced by an actual person, individuals, or human (Lee & Nass, 2004).

Perceptions of Trait Attributions

Perceptions of Standard and Foreign Accents

Data collected via an extensive online survey on accent evaluations from across the UK, which was geographically distributed and diverse, indicated that certain accents were held in higher prestige than others. Coupland and Bishop (2007) found that two ethnic accents (Asian and Afro-Caribbean) and three regional accents (e.g., Birmingham, Glasgow, Liverpool) were attributed to lower social attractiveness and prestige, compared to received pronunciation, meaning the standard British accent. Although, results also seemed to indicate that received pronunciation was evaluated less positively by younger generations. Additionally, accents were rated less heavily when individuals reported being more open to linguistic diversities. The research supports the task and social dimensions of trait attributions as individuals' rate what is perceived as the standard accent as more prestigious than any other non-standard accent. Simultaneously, the findings indicate that other factors could influence individuals' perceptions.

Furthermore, the perception of Hispanic counselors' race and speech accents and the connection to the cultural awareness of participants, who were Euro-Americans, indicated that non-accented speakers are rated more favorably than accented speakers (Fuertes & Gelso, 2000). Cultural awareness was examined by measuring participant Universal-diverse orientation (UDO), meaning their progressive and tolerant attitudes. Thus, individuals who score high in UDO are more open, tolerant, and tolerant of similarities and differences towards the self and others. The results indicate that low UDO participants rated the accented counselor lower in attractiveness, expertness, and trustworthiness than the non-accented counselor. Additionally, participants were more inclined towards having a long-term counselor without accented speech

compared to non-accented. Hence, the findings indicate that cultural awareness is a mediator of the effect accents had on listeners. Moreover, in support of the traits attributed to the task and social dimensions, accented counselors were rated less favorably on items included in the dimensions than their non-accented counterparts.

In line with the perception of trait attributions on the task and social dimensions, Lev-Ari and Keysar (2010) suggested that accented speakers might sound less credible because non-native speech signals being part of the 'other' and fluency is harder to process. Their study used mild- accented speech (Polish, Turkish, and Austrian-German) and heavily accented speech (Korean, Turkish, and Italian). Findings indicate that accented speech was rated as less credible on account of statements being perceived as more untruthful than difficult to understand. Furthermore, when participants were aware that the difficulty of processing accented speech could impair their credibility judgments, they were unable to avoid misattributing heavy-accented speech but succeeded with mild-accents. The study further supports the evaluation of accented speech regarding the trait attributions of the task dimension.

The standard accent received more favorable ratings on both dimensions than non-standard, regional, and foreign-accented speech (Coupland & Bishop, 2007). Individuals, even those with non-standard accents, rated standard accented speech for status higher (Fuertes et al., 2002). Fuertes et al. (2002) proposed that this might be due to mainstream views of the standard way of speech. However, there is less consistency regarding the solidarity dimension, wherein firstly, higher ratings were given to similar accented speech by non-standard accented individuals, and secondly, the standard and non-standard speech were rated similarly by standard accented individuals (Fuertes et al., 2002). The observation of accented speech is perceived as more prestigious in interpersonal evaluations on two dimensions, namely status and solidarity, which lead to the development of the Accent Prestige Theory (Fuertes, Potere & Ramirez, 2002). The status dimension (i.e., task) includes aspects related to the speaker's perceived competence,

such as intelligence, success, education, and social class. The solidarity dimension (i.e., social) includes aspects related to a speaker's perceived friendliness, trustworthiness, and kindness (Fuertes et al., 2002). Therefore, this study uses standard American and foreign accents to assess individuals' perceptions of trait attributions categorized into the task and social dimensions.

Perceptions of Machine Voices

This section explores previous research on attitudes and perceptions towards computer-synthesized speech. To the best of my knowledge, only one study has previously examined various accents in computer-synthesized accents, which will be reviewed first. Other research in this section will focus on studies examining human and machine voices and individuals' perceptions or preferences.

The study examined listeners' perceptions of robotic accents to develop a healthcare robot in New Zealand (NZ). The investigation of Tamagawa, Watson, Kuo, MacDonald, and Broadbent (2011) was separated into two studies of English accents from the United States, New Zealand, and the United Kingdom. The first study examined the preference of listeners towards the robot accents and how robotic they sounded. The findings suggested that the accent, which was rated as significantly less robotic, was NZ's accent rather than the US. In contrast, the UK accent did not rate significantly in comparison to other voices on roboticness. Regarding preferences, it was found that the UK accent was significantly preferred over the US accent, whereas the NZ accent compared to other voices did not show any significant differences. It was suggested that individuals being familiar with the NZ accent and the accent being similar to their own; explained their preference. Moreover, the UK accent might have been preferred over the US accent due to NZ history with colonialism (Tamagawa et al. 2011).

Continuously, in their second study, Tamagawa et al. (2011) combined the robot voice and a physical robot' body' to examine whether ratings and feelings towards the robot would be affected. Participants were randomly assigned to a robot with either an NZ or US accent, as they

were rated less robotic in the first study. The findings indicate that the participants with US-accented robots had fewer positive emotions during the interaction than participants assigned to the NZ-accented robot. Additionally, the robot with the NZ accent was rated higher in overall performance than the US-accented robot. As the study was performed in NZ, this could indicate that the NZ's accent similarity and familiarity made participants prefer it to the US accent (Tamagawa et al., 2011). It was suggested that preference could be due to similarity or familiarity with their own accent. In line with the theory of doubly disembodied language, accented computer-synthesized speech would be processed similarly to real human speech. Hence, it could be suggested that the own-accent bias could be applied here in line with ethnocentrism. Wherein, the in-group is made up of the NZ accent, and the US accent is part of the out-group.

Additionally, support for the processing of text-to-speech voices comes from research of attitudes and perceptions of computer-synthesized speech. Stern, Mullennix, Dyson, and Wilson (1999) examined individuals' attitudes towards human and synthesized speech by examining speaker perceptions. They used two different synthetic speech systems, which they defined as speech, which was made solely by a machine and five pre-recorded male speakers. Their findings indicated that an individual would consider the speech more truthful, knowledgeable, and involved if the speaker was human compared to synthesized speech, which was comparatively considered more powerful.

Additionally, it was proposed that communication style and culture could affect whether subjects accept robotic recommendations (Rau, Li & Li, 2009). The robotic recommendation was measured through a significant change in opinion from a participant group (i.e., Chinese or German) when the robot appeared to disagree with price selections. Moreover, participants were asked to rate the robot on trust, likability, and credibility. On the premise of Hofstede's (2001, as cited in Rau et al., 2009) cultural dimensions, Germans believe in being more independent in

their decision making and are more decisive as their culture is more masculine and individualistic. Comparatively, the Chinese are on the other end of these two cultural dimensions (i.e., feminine and collectivistic) and thus, reflect more and prefer group decisions. Thus, it was hypothesized that the robot would be rated lower in trust, likability, and credibility; and robot recommendations would be less accepted by German rather than Chinese participants, which was supported by results (Rau, Li & Li, 2009). The findings indicate that the computer-synthesized speech, depending on cultural values, might be perceived differently.

Another study assessed whether practical knowledge (i.e., knowledge and experience) and linguistical rhetoric cues (i.e., speaking prowess, organization, fluency) would affect the persuasiveness and perceived credibility of a robot in western (US) and Arabic cultures (Lebanon) (Andrist, Ziadee, Boukaram, Mutlu, & Sakr, 2015). The findings, specifically those referencing the robots' credibility, are required to us in this study. Thus, according to Andrist et al. (2015), the findings indicate that robots who had high knowledge were rated as more credible than low knowledge robots. Participants from America gave higher ratings of credibility than Arabic ones. Furthermore, high knowledge robots were rated higher in sociability than low knowledge robots. Additionally, the American participants rated the robots higher in sociability than the Arabic participants (Andrist et al., 2015). The findings suggest that robots are perceived differently between cultures, and one could argue that the task and social dimensions of trait attributions seem to be applicable toward computer-synthesized machine voices.

The research indicates that human accents and computer-synthesized machine voices can be rated on the task and social dimensions of trait attributions. The theory indicates that standard accents, which are believed to be the typical or 'idealized' way of speaking, will be afforded higher prestige either because of similarities to an individual's own accent (Lindemann, 2003) or because of existing mainstream views regarding the standard way of speaking (Fuertes et al., 2002).

Research on the trait attributions of the task dimension (i.e., intelligence or competence) when investigating human accents seems to consistently find similar results. More specifically, the standard accent is rated higher on expertness than a Hispanic accent (Fuertes & Gelso, 2000); on credibility compared to heavy accented Korean, Turkish, and Italian speech (Lev-Ari & Keysar, 2010). Besides, the dominant group in society has been rated to have more status compared to foreign accents (Nesdale & Rooney, 1996; Lindemann, 2003). Another idealized accent used in the UK, namely perceived pronunciation, is afforded more prestige than Asian and Afro-Caribbean accents (Coupland & Bishop, 2007).

Furthermore, the same studies indicated that when it comes to trait attributions of the social dimension, the findings are more inconsistent, on which accent is rated higher. More specifically, the standard accent is rated higher on trustworthiness (Fuertes & Gelso, 2000). According to Coupland and Bishop (2007), the received pronunciation was attributed to more social attractiveness than Asian or Afro-Caribbean accents. Contrastingly, studies have indicated that a non-standard accent or foreign accent is also rated high on solidarity (Nesdale & Rooney, 1996; Lindemann, 2003).

Similarly, research of attitudes and perceptions of machine voices has indicated that human voices are preferred and rated higher on task and social dimensions over machine voices (Stern et al., 1999). There were not many inconsistencies for the task and social dimensions of trait attributions, whether the human or machine voice was preferred. More specifically, human speech was considered more knowledgeable and truthful than computer-synthesized speech (Stern et al., 1999). Additionally, Rau et al. (2009) found that robots were rated lower on trust, likability, and credibility. Therefore, findings indicate that the standard accent will be rated higher than foreign accents and machine voices concerning the task dimension. When it comes to the social dimension of trait attributions, while ratings have been inconsistent for foreign accents, it can be assumed that the standard accent would be rated higher than foreign accents and

computer-synthesized voice. Based on the argument above, the following hypotheses were proposed:

H1: In general, American-English speakers will rate the standard American accent higher on the task dimension of trait attributions than a low prestige foreign accent and computer-synthesized accent.

H2: In general, American-English speakers will rate the standard American accent higher on the social dimension of trait attributions than a low prestige foreign accent and computer-synthesized accent.

Influence of Ethnocentrism on Perceptions of Trait Attributions

Ethnocentrism

Accents can be examined using an ethnocentrism-based approach, which essentially allows individuals to reduce the complexity of the social world through order and predictions by categorizing people into in- and out-groups (Hopkins, 2015; Cargile, Giles, Ryan & Bradac, 1994). Categorization is achieved by highlighting similarities to oneself (e.g., that sounds like me) and others' dissimilarities toward oneself (e.g., that sounds different to me). Once people have been categorized, an individual will show a preference towards people who are similar to them, meaning part of the in-group. If another is dissimilar, they will become part of the out-group and viewed as the 'other' (Reddington, 2008).

Ethnocentric groups believe themselves to be the epitome of virtue, superiority, and upholding the universal standard of values. Subsequently, in-group members will behave positively towards one another through cooperation and obedience as they are willing to remain a member (Neuliep & McCroskey, 1997). Contrastingly, out-groups are perceived as the epitome of immorality, inferiority, and weakness, and are treated contemptuously. Additionally, the ethnocentric group will actively compete with and disobey members of the out-group (Neuliep &

McCroskey, 1997). Furthermore, through the increase of conformity, solidarity, loyalty, and cooperation, ethnocentrism could help secure the survival of the in-group (Sharma, Shimp & Shin, 1995).

Incidentally, in times of threat, ethnocentric tendencies can be advantageous as it might lead to patriotism or sacrifices for one's group (Neuliep & McCroskey, 1997). However, seeing the own group as solely acting in the right way can have detrimental effects on an out-group, such as prejudice, discrimination, and in extreme cases, ethnic cleansing (Neuliep & McCroskey, 1997). Moreover, concerning accented speech, ethnocentrism operates on the basis that the standard way of speaking comes from one's own culture. Thus, out-group speech is evaluated against that standard. If the speech (i.e., non-native accents) deviates from the standard, it can elicit negative views and provide evidence of inferiority (Neuliep & McCroskey, 1997). Additionally, it is suggested that the accent, which has specific speech characteristics, is an indicator of which group they will be categorized into (Bresnahan, Ohashi, Nebashi, Liu, & Shearman, 2002). Thus, if the accent is categorized as part of the out-group, it can initiate various stereotypes associated with social, ethnic, or national groups (Bresnahan et al., 2002; Ryan, 1983).

Furthermore, ethnocentric individuals will view others less favorably should they deviate from self- and socially constructed norms. Consequently, according to Chakraborty (2017), these deviators will be perceived as inferior, extending to individuals speaking different languages or with accents. Moreover, the author stated that ethnocentrism is not specific to any one social situation or domain but is prevalent across domains; and, as such, can even pervade accented speech. According to Gagnon and Bourhis (1996), out-group discrimination can be observed through ethnocentrism as it leads individuals to identify their in-group. Once in-group identification took place, this will lead to in-group biases and, consequently, out-group discrimination (Gagnon & Bourhis, 1996). Thus, Chakraborty (2017) noted that individuals with

non-native accents would be perceived more negatively by listeners whose ethnocentricity is higher.

A study investigating individuals' behavior being influenced depending on their ethnocentricity was investigated in a study about group affinity (Heblich, Lameli, & Riener, 2015). In essence, the study incorporated an economic situation whereby a participant is asked to assess whether he/she believes to possess the ability to cognitively out-perform the speaker (standard accent vs. regional accents). They are then asked to follow one of two payment schemes, which indicates whether an ethnocentric group affinity exists. In other words, depending on the payment scheme chosen, they will cooperate with the speaker or compete with them. The findings indicate that when participants are presented with common tasks, they are more likely to cooperate with individuals of the same accent. Alternatively, when paired with an individual speaking in a different accent, more competitiveness can be observed. The findings led to Heblich, Lameli, and Riener (2015), suggesting that linguistic dissimilarities in accents create ethnocentric beliefs about different accents and influences how listeners perceive accented speakers. In other words, listeners might feel more superior about their cognitive abilities, the more dissimilar a speaker's accent sounds.

The idea of an ethnocentrism-based approach towards the evaluation of accented speech was supported by Neuliep and Speten-Hanse (2013), who explored ethnocentrism on perceptions of non-native accent speakers. Findings indicate that the relative degree of ethnocentrism does play a role in negative perceptions of non-native accented speech. More specifically, there is a decrease in the positive perception of non-native accented speakers on homophily, credibility, and attractiveness compared to native-accented speech.

Additionally, an individuals' degree of ethnic identity and its effect on feelings towards native and non-native accented speech was examined by Bresnahan et al. (2002). They hypothesized that individuals with a weak ethnic identity would exhibit a more positive attitude

and emotional response towards foreign-accented speech than individuals who strongly identify with their ethnic group. Their study explored and compared attitudes towards American English, intelligible, and unintelligible foreign accents. The findings supported the hypothesis, wherein individuals with relatively stronger ethnic identities attributed more dynamism, higher status, and attractiveness towards American English than unintelligible foreign accents (Bresnahan et al., 2002). Contrastingly, higher status and attractiveness were attributed to unintelligible foreign accents rather than American English by individuals with a weaker ethnic identity.

Moreover, Bresnahan et al. (2002) found that American English was rated more dominant, arousing, and pleasant by individuals with stronger ethnic identities than weaker ethnic identities. American English and intelligible foreign accents compared to unintelligible foreign accents also elicited an emotional response, which was significantly more positive when individuals had a strong ethnic identity. Weaker ethnic identity individuals had an affective response in which foreign accents were preferred for pleasantness, arousal, and dominance compared to American English (Bresnahan et al., 2002). These findings reinforce the role of ethnocentrism in understanding in-group favoritism. A strong ethnic identity, equated to strong in-group membership, led to a more favorable emotional response towards American English.

Furthermore, by aiming to explore judgments made by children who are part of the majority group on the voices of majority and minority group children, Nesdale and Rooney (1996) further supported the role of ethnocentrism in understanding in-group favoritism. The study asked Anglo-Australian children who only spoke one language to listen to children with Anglo-, Italian-, and Viet-Australian accents. The researchers were interested in pre-adolescent children's language attitudes, wherein the study explored which accent was attributed more status or solidarity. The findings are consistent with previously reported results; namely, the dominant group's accents were attributed more status; and more solidarity was attributed to foreign accents. Previous research has indicated that results for solidarity are inconsistent. Additionally,

consistent with ethnocentrism, the dominant group was partially rated more favorably on the dimensions as opposed to the foreign-accented out-group.

Another study examined Korean-accented speakers' evaluations and attempted to determine whether listeners could identify the speaker's ethnicity and status (Lindemann, 2003). American-English speaking listeners viewed speakers of the same accent as part of the in-group and as more dominant, compared to the Korean speakers who are associated as out-group members (Lindemann, 2003). As such, the findings are in line with previous accented speech research; in this case, Koreans are rated more negatively on status than the standard accent. Moreover, as previously indicated, solidarity ratings may be inconsistent in whether they are attributed to the standard or non-standard accent; these findings rated the Korean accent relatively high (Lindemann, 2003). These findings indicate that the in-group is rated more favorable than the out-group; this offers further support for ethnocentrism in understanding in-group favoritism.

Own-Accent Bias

Additional support of the categorization of accented speech and ethnocentrism's role in understanding in-group favoritism comes from research investigating the concept of own-accent bias. In essence, the own-accent bias functions on the premise that when accented speech is judged, a similar or own accent is perceived as more favorable (Coupland & Bishop, 2007) and trustworthy (Lev-Ari & Keysar, 2010) by an individual.

A study illustrated that own-race bias does not solely regard faces of the same race but can also be applied to voices. The study suggested that exposure to various dialects is advantageous for listeners as it aids in differentiating between own-race voices and those of another race (Perrachione, Chiao & Wong, 2010). In the study, the authors analyzed the acoustics of numerous dialects and discovered that categorizing speakers' race was accurately done by white and black Americans. Additionally, own-group bias was demonstrated by

participants seemingly having an advantage in identifying their racial group (Perrachione, Chiao, & Wong, 2010). Although, Bestelmeyer et al. (2015) stated that the study provides evidence of own-accent bias rather than own-race bias as the speakers' dialect was the focal point of the study rather than vocal structures.

Moreover, it has been proposed by Kinzler, Dupoux, and Spelke (2007) that social preferences are connected to linguistic in-group interactions in early development. Newborn infants show a preference for their language or what is familiar (i.e., mothers' voices). Infants are sensitive to native speech and can distinguish it from non-native language (Kinzler, Dupoux & Spelke, 2007). Thus, Kinzler et al. (2007) investigated whether infants and young children exhibited visual and social preferences towards speakers who spoke their native language. They firstly, had infants watch an American-English and French-speaking adult woman. Infants looked equally long at both women during the exposure trial but longer at American-English-speaking women during silent trials.

In another test, Bostonian and Parisian infants were offered toys simultaneously from women speaking English and French, with infants reaching more for the toy coming from the woman speaking their native language (Kinzler et al., 2007). The last experiment consisted of 5-year-old English speakers being shown two unfamiliar pictures of children while simultaneously hearing French or English. The children chose the English-speaking child over the French one when asked whom they would instead befriend. The results support that social preference for group-members of the native language compared to foreign language develops early on (Kinzler et al., 2007), thus supporting the own-accent bias.

Further support of the own-accent bias was provided by Kinzler, Corriveau, and Harris (2011). They found a connection to social preference when native-accented speakers (i.e., American English) and foreign-accented speakers (i.e., Spanish) interacted with children. The findings indicated that children seem to learn selectively, depending on the accent. More

specifically, children are more likely to ask for more information on a novel object introduced if it came from a native English speaker than a foreign-accented speaker. Additionally, children prefer to endorse information received from a native-accented speaker rather than a foreign-accented speaker (Kinzler, Corriveau & Harris, 2011). The findings seem to be in line with the own-accent bias, wherein native English speakers will show a preference for the in-group or the own-accent instead of the 'other' or out-group.

To explore own-accent bias and whether affective processing or prototype representation could best explain it, Bestelmeyer et al. (2015) attempted to identify neural markers through fMRIs. In this case, affective processing would light up brain areas, such as the amygdala or the basal ganglia, if emotional reactions drive a positive own-accent bias. Alternatively, the own-accent bias might be better explained through prototype representation, meaning the speaker's accent has familiar prototypes or representations, while 'other' accents deviate from those. Moreover, if prototype representation better explains one's own-accent bias, then specialized auditory regions in the brain, such as the bilateral primary and secondary auditory cortices, would light up (Bestelmeyer et al., 2015). The participants listened to Southern English, Scottish, and American accents, while the fMRI measured their blood oxygenation level, indirectly indicating neural activity. The findings indicate that own-accent bias has a neural signature. The repetition of accents similar to the participant was associated with increased activation in the bilateral amygdala, anterior cingulum, and the right Rolandic operculum. Comparatively, accents associated with the 'other' showed reduced activity in those regions. The findings led to Bestelmeyer et al. (2015) concluding that own-accent bias could be explained by affective processing instead of prototype representation.

Categorization of Nonhuman Machine Voices

The human accent research indicates that ethnocentrism plays a significant role in understanding in-group favoritism, which influences perceptions of trait attributions on the task

and social dimensions towards accented speakers. However, this indicates that ethnocentrism levels can be influential toward perceptions of not only human voices but also machine-generated voices. According to Gong (2008), there has not been a large amount of empirical testing connected to human in-group favoritism. The reason is mainly due to no nonhuman out-group (i.e., robot) previously existing in society's everyday interactions. Although, further statements indicate that testing could still be done because of popular culture continuously representing them (Gong, 2008), such as video games or movies. Additionally, with the advancement of technology since then, a nonhuman out-group seems to have become commonplace for individuals due to their interaction with 'machines,' such as automated customer services, computers, or smartphones.

Aligned with the proposition of ethnocentrism playing a role in understanding in-group favoritism is the idea of mindless behavior, meaning that the attention toward a subset of contextual cues leads to the triggering of various scripts or expectations, thus splitting an individual's focus (Nass & Moon, 2000). In essence, mindless responses towards computers contend that individuals will apply pre-existing social scripts used in human-human interactions to a computer, thus ignoring cues indicative of a computer's nonhuman nature. Hence, individuals will use their past experiences when responding, instead of consciously absorbing all the characteristics of the situation to construct further categories or distinctions (Nass & Moon, 2000).

In line with that, a study attempted to examine whether individuals would depend on assigned social categories of in- and out-group in their computer interactions. The study used the "color war" concept, which, according to Nass, Fogg, and Moon (1996), led to feeling loyal and dependent on the group and to perceptions of superiority regarding the own team rather than the other team. Individuals in the first condition were reminded of their dependency on the computer, given blue armbands, had blue borders on the computer monitor, and were named the

"blue team" to create feelings of shared identity between humans and machines. Contrastingly, the second condition encouraged people to concentrate on individual responsibility and referred to individuals as the "blue person" who had to work with the "green computer."

In a bigger context, Robertson (2007) noted that the melding between humans and machines is actively pursued to establish a superior species, which she named Robo sapiens japanicus. Robots are created to be used in the service, research, and entertainment (includes advertisement) sectors. Furthermore, the Japanese government's active promotion and investment in robotic technology are due to a decline in birthrates and increased elderlies. Hence, the aim is to provide robots for childcare, eldercare, and performing household tasks. Additionally, Robertson (2007) states that robots are preferred over foreign labor because there are no cultural differences. Moreover, robots do not possess wartime memories of Japan's past imperialistic transgressions towards other countries (e.g., Asian). The same cannot be said for foreign or migrant workers. Thus, using robots instead of foreign labor could establish Japan as a homogenous country (Robertson, 2007). Wright (2019) also noted that the robot could be used to breach cultural and linguistic barriers and facilitate migrant workers' introduction at the cost of quality caregiving.

The integration of robots into society (e.g., globally) or the family (e.g., in Japan) has begun establishing and differentiating between human and robot rights. Robertson (2014) notes that Japan officially supports universal human rights. However, the author believes the concept and reality of human and civil rights distribution indicate a division between Japanese and non-Japanese residents. Consequently, Robertson (2014) notes that this divide can be observed in human and robot rights, with the differential treatment of robots and non-Japanese residents. Hence, the state (through policies) and the public embrace robots, as robots can be considered family members and can even receive Japanese citizenship. Contrastingly, non-Japanese residents, minorities, migrant workers, or foreigners do not receive the same treatment.

Robertson (2014) states that this might indicate that ethnonationalism (i.e., Japaneseness) might be privileged over mere humanness.

Consistent with such shift of allegiance away from foreigners see in Japan, the findings of Nass et al. (1996) indicate that individuals in the team condition were significantly more likely to cooperate and conform to the computer's suggestion compared to individuals from the different team conditions (Nass et al., 1996). Additionally, computers were assessed as more friendly and intelligent and perceived as similar to the self by individuals in the same-team than the different-team conditions. Even though individuals claimed afterward that the labels did not affect their attitudes and behaviors, the results indicate that individuals, regardless of whether the social categories make sense, will rely on them during computer interactions (Nass et al., 1996). It is important to establish that the study did not use actual computer-synthesized voices but rather text-based scripts.

Another study using an ethnic cue as a social category and presenting Korean participants with hypothetical choice-dilemma situations seems to support the attribution of pre-existing social scripts (Nass, Isbister, & Lee, 2000). The study examined individuals' perceptions and quality of the agents' decision and argument and participants' decisions on the various dilemmas. Individuals had to read the dilemma situation, make a decision, and ask for the computer agents' (i.e., Caucasian or Korean video face) decision (Nass et al., 2000). Consistent with mindless stereotyping, the findings indicated that regardless of the context, the ethnicity cue triggered certain expectations and attributions. Notably, individuals would perceive the computer agent as more intelligent, trustworthy, persuasive, and attractive as well as similar in decision making, when they were in the same-ethnicity condition compared to the different-ethnicity condition (Nass, Isbister & Lee, 2000).

Additionally, a second experiment was conducted by Nass et al. (2000), where individuals were told they would videoconference with another individual to ascertain whether

the computer was mindlessly categorized as a social actor. According to Nass, Isbister, and Lee (2000), individuals rely strongly on an actual person's ethnicity when assessing advice. The findings suggest that the effect of the ethnic video face was as strong for human agents as for computer agents. The study indicates that a social actor's humanness or non-humanness was inconsequential once an individual had ethnically categorized them. Therefore, this would suggest that when hearing a voice or accent that is ethnically similar to the own, it might be rated or perceived more favorably than the accent of different ethnicities.

Furthermore, another study investigated racial prejudice, applying not only to humans but also to robot characters and human characters, which had been computer-synthesized (Gong, 2008). The researcher investigated individuals' preferences for white, black, and robot characters based on measures of explicit and implicit racial prejudices. A unique research question pondered whether preferences between black and robot characters could be predicted through racial prejudice when participants are not interested in robots. Participants were then asked to eliminate or select avatars according to their preferences in connection to fighting games, and virtual friends and tutors (Gong, 2008). The findings indicated that when selecting virtual friends and possibly tutors, the robot characters were preferred over the black characters when the white participants are highly prejudiced. Moreover, highly prejudiced individuals liked the robot characters over the black characters, even though they indicated a low interest in robots (Gong, 2008). In discussing the findings, Gong (2008) posited that the results challenged the notion of human in-group favoritism because robots were not assumed to be considered part of the nonhuman out-group.

This study investigated how ethnocentrism measures influence ratings on task and social dimensions of trait attributions for foreign and computer-synthesized accents. Research into human accents has established that accents are strong social indicators for categorizing speakers into in- and out-groups (Neuliep & McCroskey, 1997). In keeping with categorization,

ethnocentrism research indicates that native accented speech is perceived more positively on similarities, credibility, and attractiveness than non-native accented speech (Neuliep & Speten-Hanse, 2013). Furthermore, Bresnahan et al. (2002) found that individuals who strongly identify with their ethnic group would have more negative attitudes and emotional responses toward foreign-accented speech. Moreover, studies indicate that children exhibit a social preference and bias for the own-accent from early development compared to that of the 'other,' meaning foreign-accented speech (Kinzler et al., 2007; Kinzler et al., 2011). Therefore, in line with the role of ethnocentrism in understanding in-group favoritism, the own accent will be favored over the accent of the 'other.'

Furthermore, research into machines has proposed that individuals can attribute social categories and human characteristics to nonhuman entities (Nass & Moon, 2000). More specifically, research indicates that individuals who believed the computer was on their team would perceive it as more friendly and intelligent (Nass et al., 1996). Moreover, according to Nass et al. (2000), individuals who showed an ethnic video frame on a computer would perceive the same-ethnicity entities as more intelligent, trustworthy, and persuasive than different-ethnicities. Another study found that individuals who measured high on prejudices would like a computer-synthesized robot character over a black character and prefer to have the robot character as a virtual friend over the black character (Gong, 2008). Therefore, research indicates that for an individual, a foreign accent and computer-synthesized accent could be either seen as part of the in- or out-group when it comes to human and nonhuman categorization. However, the computer-synthesized accent is that of US English, wherein this could be a social cue and category of the same ethnic group. Thus, an individual with a standard accent who indicates higher levels of ethnocentrism might perceive the nonhuman entity as part of the in-group. Based on this argument, the following hypotheses were proposed:

H3: As the levels of ethnocentrism among American-English speakers increase, the positive perceptions on the task dimension of trait attributions toward the computer-synthesized accent in comparison to a low prestige foreign accent will increase.

H4: As the levels of ethnocentrism among American-English speakers increase, the positive perceptions on the social dimension of trait attributions toward the computer-synthesized accent in comparison to a low prestige foreign accent will increase.

CHAPTER 3: METHODOLOGY

Participants

This study had a total of 351 American-English speaking participants recruited via Amazon MTurk, which would lead them to a Qualtrics link. Overall, 60.2 % of the participants identified as being female, 39.2 % identified as being male, and the remaining 0.6 % identified as other. The ethnic and racial background composition were mainly Caucasian (79%) with the remaining ethnic and racial backgrounds being divided among Black (7.4%), Asian (5.4%), Hispanic (3.4%), Mixed (4.0%), and lastly, American Indian/ Native American (0.9%). The participants' ages ranged from 19 years to 72 years (M = 38.03, SD = 12.26). Additionally, participants were asked whether they spoke any language other than English fluently taken from the Center for the Study of Adult Literacy (n.d.); and if yes, to list them. A large number of participants (85.2%) did not speak another language fluently, compared to 14.8% who spoke various other languages fluently. The languages participants self-reported to speaking fluently other than English were Arabic, Chinese and Chinese Mandarin, Japanese, French, Spanish, Italian, Portuguese, German, Russian, Hebrew, Hindi, Korean, Turkish, Uzbek, Vietnamese, and Cambodian.

Procedure

Participants used the Qualtrics link made available through Amazon MTurk by having been given a short description of the study (Appendix A). Participants were only able to access the study if they met the worker qualifications provided on MTurk, namely being located in the US and having approval ratings for HITs greater than or equal to 95%. Additionally, three screening questions (born in the US; time spent in the US; speaking standard American-English) were specifically coded as an added qualification test, which had to be answered with "yes" (Appendix B). Having answered "yes" to the three screening questions to qualify them for the study, participants were then shown the consent form (Appendix C). The first part of the study

pertained to the task and social dimensions scale. Participants were asked nine times to listen to an audio recording (three from each accent category) and fill out the measure for perceptions of trait attributions (task and social dimension) and an abbreviated version of the Perceived Voice Impressions measure for each. The second part of the study asked participants to fill out the Generalized Ethnocentrism Scale. Lastly, participants were given the demographic survey and thanked for their participation, which signifies the end of the survey.

Study Stimuli

Accents

The standard American accents and the foreign accents were obtained from the International Dialects of English Archive (IDEA), which has a vast collection of pre-recorded accents. The recordings used in this project are used by special permission of the International Dialects of English Archive, online at http://www.dialectsarchive.com. This study used a verbal guise method, wherein different speakers pre-recorded their voices. Lindemann (2003) noted that this presents less methodological issues present in the matched guise method, where participants might become aware that the same speaker is performing a standard and foreign accent.

The computer-synthesized machine voices were created using the text-to-speech system available through the Apple MacBook and Amazon Polly. All the audio recordings used the same transcript; Comma gets a Cure (Honorof, McCullough, & Somerville, 2000), used by IDEA (see Appendix D).

The transcript was created specifically for IDEA and was used for two reasons. Firstly, because of availability as the human accents have already been recorded ahead of time and IDEA provides many accents from across the world. Secondly, transcripts from previous research had either statement (e.g., see Perrachione et al., 2010), or individuals talking about themselves (e.g., see Fuertes & Gelso, 2000), or two different types of texts for friendship and teaching purposes (e.g., see Bresnahan et al., 2002). Thus, the transcript, which is a story narration, seems to be

neutral on the task and social dimensions of trait attributions, and according to IDEA (2019), uses the standard lexical sets of words established by J.C. Wells. Using Comma Gets a Cure allows participants to share their attitudes and perceptions of the voices without being unnecessarily influenced by the context of the speech. Additionally, participants were not exposed to the full-length text, but only the first paragraph, which should be around 30-second recording. Although, it is important to note that the length might have varied depending on the speaker's speaking rate.

Selection of the Stimulus Tapes. The human accents for this study were taken from the International Dialects of English Archive (IDEA). For the purpose of this study, only 'male' recordings were selected from the human and machine voices as gender stereotypes may apply when it comes to perceptions of trait attributions on the task and social dimensions (see Lee et al., 2000), hence, to circumvent this only male voices were used. The standard accent was retrieved from a special collection named general American, which was generated by trained speech teachers. There is a selection of males and females; however, only the males were important for this study.

As for the foreign accent, IDEA categorized the accents into their respective continents, namely Asian, South American, and Middle Eastern. Next, within the random continents, the accents were divided into countries. By randomly selecting countries and male recordings within each, it was possible to attain a selection of accents per continent from various countries. Accents were selected by listening to the recordings and ascertaining whether the recording was comprehensible, clear, and the background noise or microphone noise was kept to a minimum. Also, the stimulus tapes should have a discernible accent, as certain individuals' speech was heavily influenced by having lived in an English-speaking country for some time.

The computer-synthesized accent was generated using the text-to-speech system provided by the Apple MacBook. This is a free system as long as one owns an Apple device. An audio

recording can be produced by merely highlighting the transcript, choosing a voice, and generating an mp3 file. For the Amazon Polly recordings taken from Amazon Web Services (AWS), the transcript has to be pasted into the system and generated to be downloaded.

Once all the accents have been selected or generated, a selection of 21 stimulus tapes, as shown in Table 1, were used for informal pre-testing to determine which ones should be used in the study. The standard accents which were chosen are General American 1, General American 5, General American 9, General American 10, General American 13. The foreign accents which were chosen are from Asia (Beijing 2, Anhui 2, India 3, Japan 9, Vietnam 2); South America (Chile 4, Colombia 4, Peru 6, Venezuela 5); and the Middle East (Turkey 3). For human accents, a total of five Standard American and ten foreign accents were chosen. The computer-synthesized accents which were used in the informal pre-testing had a total of four from the TTS system in the Apple MacBook as well as two from Amazon Polly.

Table 1

List of Names of Stimulus Tapes by Accent Category

Standard American*	**Foreign***	**Computer-Synthesized**
General American 1	Beijing 2	Alex**
General American 5	Anhui 2	Bruce**
General American 9	India 3	Fred**
General American 10	Japan 9	Tom**
General American 13	Vietnam 2	Matthew***
	Chile 4	Joey***
	Peru 6	
	Colombia 4	
	Venezuela 5	
	Turkey 3	

* The stimulus tapes for both the Standard American and Foreign Accents came from IDEA
** The stimulus tapes were generated using Apple MacBook
*** The stimulus tapes were generated using Amazon Polly

The informal pre-testing (N = 11) aided in selecting nine stimulus tapes out of the 21 stimulus tapes (see Appendix E). The informal pre-testing was conducted using opportunity sampling. Overall, 72.7 % of the participants identified as being female, 27.3 % identified as being male. The ethnic and racial background composition were mainly Caucasian (45.5%) with the remaining ethnic and racial backgrounds being divided among Asian (18.2%), Mixed (27.3%), and lastly, Hawaiian or another Pacific Islander (9.1%). The participants' ages ranged from 23 years to 37 years (M = 26.91, SD = 4.13). The participants received a Qualtrics link asking them to listen to 21 stimulus tapes and fill out the measure for Perceived Voice Impression. For the pre-testing, participants had to randomly listen to 5 audio-recordings from each accent category. Each recording was followed by three statements from the Perceived Voice Impression measure related to the respective accent category (see Appendix I). Then the data were analyzed for mean scores and reliability to ensure the sound of the speakers' accent was in line with their accent category. The accents indicating the highest mean scores were used in the study. The findings indicate that the averages indicated these tapes sounded the most similar to the accent category they belonged to. Table 2 present the results from the reliability test, with the Cronbach alphas ranging from .717 to 1.00.

Table 2

Mean scores and Reliability for the 9 Selected Stimulus Tapes

Accent Stimulus Tapes	Mean Scores *(SD)*	Reliability
General American 1	6.42 (.86)	.96
General American 10	6.23 (.80)	.82
General American 13	6.19 (1.08)	.94
Tom	6.79 (.59)	.96
Fred	7.00 (.00)	1.00
Bruce	6.83 (.36)	.89
Vietnam 2	6.53 (.51)	.72
Turkey 3	5.93 (1.28)	.97
Anhui 2	5.87 (1.26)	.96

Note. Reliability reflects Cronbach's alpha (α) for selected accent types on the Perceived Voice Impression measure.

Measures

Generalized Ethnocentrism Scale

The revised general ethnocentrism scale (GENE) was used (Neuliep & McCroskey, 2013) to ascertain individuals' levels of ethnocentrism by scoring 15 items resulting in a composite ethnocentrism score (See Appendix F). Participants were asked to indicate on a 7-point scale from how strongly they disagree (1) to agree (7) with certain statements. Some example statements would be; "Most other cultures are backward compared to my culture" and "My culture should be the role model for other cultures." As per the scales' instructions, items 4, 7, and 9 were reverse coded. Additionally, items 3, 6, 12, 15, 16, 17, and 19 were dropped before ascertaining the composite ethnocentrism score (see Appendix E). The reliability for the Generalized Ethnocentrism Scale was calculated, which resulted in a Cronbach alpha of .95, indicated excellent internal consistency.

Perceptions of Trait Attributions: Task and Social Dimensions

For this study, participants' perceptions of trait attributions were measured on a task and social dimension. The specific task and social dimension traits were selected from items used in previous research to provide a variety of trait attributions associated with the dimensions (Tamagawa et al., 2011; Lindemann, 2003; Pantos & Perkins, 2013; Dailey et al., 2005; Lev-Ari & Keysar, 2010; Nesdale & Rooney, 1996; Rau et al., 2009). The scale was developed from these traits by identifying and using items that seemed to overlap across studies on human accents and machine voices (see Appendix G). We have included four items associated with the task dimension, namely knowledgeable, intelligent, competent, and persuasive, and four associated with the social dimension, namely pleasant, likable, friendly, and trustworthy. Participants were asked to indicate the degree to how strongly they disagree (1) or agree (7) on a 7-point scale.

The reliability for the task and social dimensions of trait attributions was calculated for each accent type separately. The Cronbach alpha for the standard American accent for the task and social dimension were .90 and .89, respectively. The Cronbach alpha for the foreign accent for the task and social dimension were .96 and .97, respectively. The Cronbach alpha for the computer-synthesized accent for the task and social dimension were .94 and .94, respectively.

Perceived Voice Impressions

A measure for perceived voice impressions was used for the informal pre-testing of the accents to determine whether the recordings sound like standard American, foreign, and computer-synthesized (Appendix I). Additionally, the measure for perceived voice impressions was used in the actual study as a manipulation check. Participants were asked to indicate on a scale from 1 to 7 how strongly they disagree or agree with each statement from each category of accents. For example, "I think that this was a standard American speaker," "I think that this was a

foreign speaker," and "I think that this was a machine voice." The overall average of responses was then used to determine which audio-recordings were used in the study.

The actual study included an item each from the for scale, with each item relating to each accent sounding either "like typical American English," seemingly having a "foreign accent" or sounding "robotic" to the participant. The Cronbach alphas for the standard American accents sounding like standard American, foreign, and computer-synthesized accents were .69, .90, and .77, respectively. The Cronbach alphas for the foreign accents sounding like standard American, foreign, and computer-synthesized accents were .80, .69, and .62, respectively. The Cronbach alphas for the computer-synthesized accents sounding like standard American, foreign, and computer-synthesized accents were .97, .69, and .86, respectively.

CHAPTER 4: RESULTS

Descriptive Analysis

For the purpose of the analysis, six new variables needed to be computed by separately categorizing the ratings for the task and social dimensions of trait attributions as well as grouping together the accents into their respective accent category. That is, out of the nine accents in the study, the three accents connected to the standard American accent, the foreign accent, and the computer-synthesized accent were each computed into two new variables representing the perceptions of trait attributions on the task dimension and the social dimension for each accent respectively.

The generalized scale of ethnocentrism showed considerable variance with a range of 15 to 86 out of a total score of 105 (M = 37.82, SD = 18.78). Thus, American-English speakers, on average self-reported having weaker ethnocentric beliefs. A histogram was used to examine whether the ethnocentrism scores were normally distributed. The analysis of the distribution of participants' scores indicated that the results were skewed to the right.

The descriptive statistics of the perceptions of trait attributions on the task dimension indicate that the standard American accent (M = 5.75, SD = .75) was rated highest, followed by the foreign accent (M = 4.17, SD = 1.15). In contrast, the computer-synthesized accent (M = 3.27, SD = 1.40) was rated lowest. This indicates that American-English speaking participants perceive the standard American accent more likely to possess traits associated with tasks, such as intelligence or competence, followed by the foreign accent. Conversely, the computer-synthesized accent is less likely to be attributed to such task-orientated traits out of the three accent types.

Furthermore, the descriptive statistics related to the perceptions of trait attributions on the social dimension indicate similar results as the standard American accent (M = 5.77, SD = .80) was rated highest, followed by the foreign accent (M = 4.62, SD = 1.17). In contrast, the

computer-synthesized accent (*M* = 2.76, *SD* = 1.48) was rated lowest. In this instance, American-English speaking participants perceive the standard American accent to be more likely to possess social traits, such as likability or friendliness. The computer-synthesized accent seems to be least favorably associated with these social traits than the other accent types. Table 3 presents a summary of the descriptive statistics and the intercorrelation matrix.

Table 3

Correlations of Mean Scores of Ethnocentrism and Accent Types on Trait Attribution Perceptions (Task and Social Dimension)

	Ethnocentrism	US-T	US-S	FA-T	FA-S	CA-T	CA-S
Ethnocentrism	–						
US-T	-.10	–					
US-S	-.03	.80*	–				
FA-T	.06	.23*	.21*	–			
FA-S	-.06	.23	.22*	.84*	–		
CA-T	.38*	.16*	.19*	.50*	.43*	–	
CA-S	.50*	.08	.17*	.49*	.43*	.88*	–
M	37.82	5.75	5.77	4.17	4.62	3.27	2.76
SD	18.78	.75	.80	1.15	1.17	1.40	1.48
Reliability	.95	.90	.89	.96	.97	.94	.94

Note. US-T = Standard American Accent on the Task Dimension; US-S = Standard American Accent on the Social Dimension; FA-T = Foreign Accent on the Task Dimension; FA-S = Foreign Accent on the Social Dimension; CA-T = Computer-synthesized Accent on the Task Dimension; CA-S = Computer-synthesized Accent on the Social Dimension. Reliability reflects Cronbach's alpha (α) for all measures.
* Correlation is significant at the 0.01 level (2-tailed)

Hypotheses 1 and 2

Hypotheses 1 and 2 proposed that in general, American-English speaking participants will rate the standard American accent higher on the task and social dimensions of trait attributions than a low prestige foreign accent and a computer-synthesized accent. Then, Table 4

presents a summary of the results from the repeated measures analysis of variance. The hypotheses were tested using a 3 (accent types) x 2 (dimensions) repeated measures analysis of variance. Furthermore, before running the analysis of variance, the syntax needed to be recoded to include a comparison of the accent types. A main effect for accent types showed that American-English speakers rated the standard American accents most favorably (M = 5.76, SD = .73), followed by the foreign accents (M = 4.39, SD = 1.11), and lastly the computer-synthesized accents (M = 3.01, SD = 1.40) were rated least favorably, F(1.87, 657.54) = 767.84, p < .001, η^2 = .69. The main effect for dimensions was not significant, F(1.87, 657.54) = .44, p = .506, η^2 = .00. The interaction between accent types and dimensions on the ratings of the accents was significant, F(1.87, 657.54) = 218.65, p < .001, η^2 = .38.

Table 4

Repeated Measures Analysis of Variance for Accent Types and Perceptions of Trait Attributions (Task and Social Dimensions)

	df	F	η^2	p
Accent Types	2	767.84	.69	.000
Task and Social Dimensions	1	.44	.001	.506
Accent Type x Task and Social Dimensions	2	218.65	.38	.000
Error (Accent Type x Task and Social Dimensions)	351			

Note. Significant at the p < 0.05 level; Accent Types = Standard American, Foreign, and Machine

* Within-Subjects (SS_w)

Table 5

Paired-samples T-test Results for the Standard American, Foreign, and Computer-Synthesized Accents on the Perceptions of Trait Attributions (Task and Social Dimension)

	Mean Ratings (*SD*)			T Statistics		
	US	FA	CA	US vs. FA*	US vs. CA*	FA vs. CA**
Task Dimension	5.75 (.75)	4.17 (1.15)	3.27 (1.40)	24.31 ($p < .001$)	31.34 ($p < .001$)	12.99 ($p < .001$)
Social Dimension	5.77 (.80)	4.61 (1.17)	2.76 (1.48)	17.19 ($p < .001$)	36.24 ($p < .001$)	24.12 ($p < .001$)

Note. US = standard American accent. FA = Foreign accent. CA = Computer-synthesized accent. Degree of freedom for the *t*-tests was 351.
* *t*-test results for hypotheses 1 and 2
** *t*-test results for additional findings (see hypotheses 3 and 4)

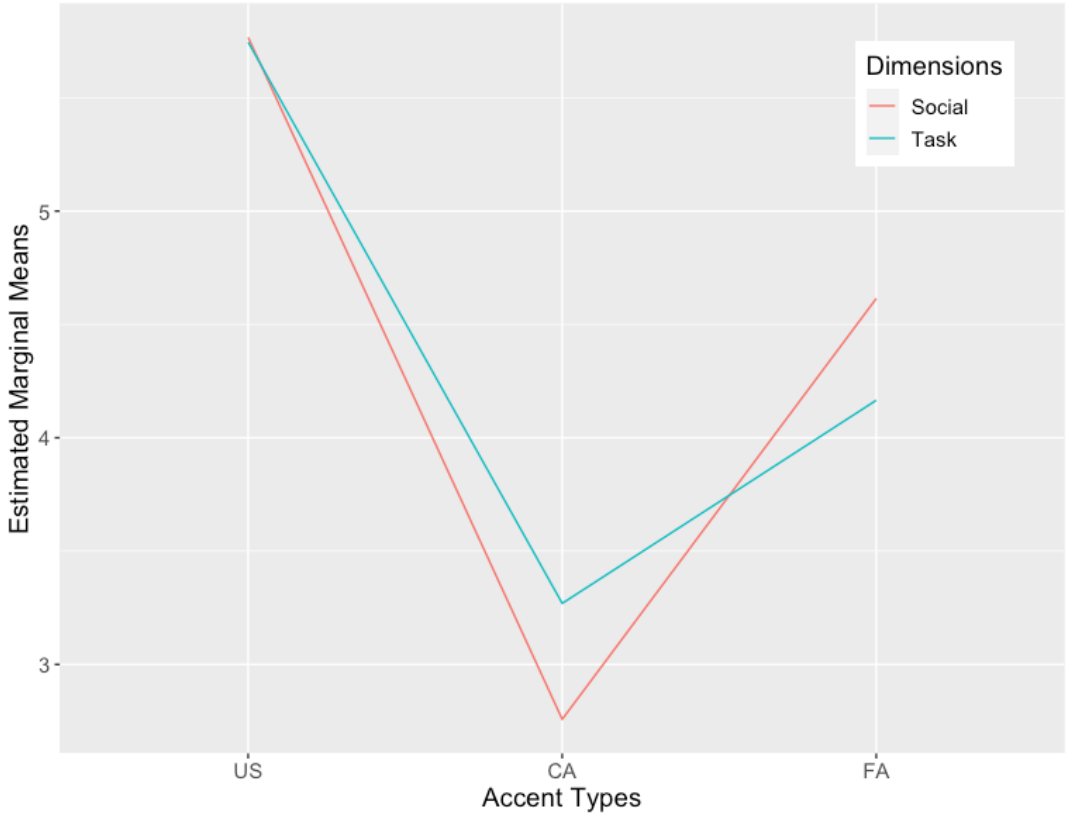
Table 5 presents a summary of the results obtained from the post hoc comparisons between the accent types on the perceptions of trait attributions (task and social dimensions). To test hypotheses 1 and 2, four paired samples t-tests were conducted. For hypothesis 1, two paired samples t-tests were used to make post hoc comparisons on the task dimension comparing the standard American accent against the foreign and computer-synthesized accent, respectively. The first paired samples t-test indicated that there was a significant difference in scores for the standard American accent ($M = 5.75$, $SD = 0.75$) and the foreign accent ($M = 4.17$, $SD = 1.15$) on the task dimension, $t(351) = 24.31$, $p < .001$. A second paired t-test indicated that there was a significant difference in scores for the standard American accent ($M = 5.75$, $SD = 0.75$) and the computer-synthesized accent ($M = 3.27$, $SD = 1.40$) on the task dimension, $t(351) = 31.34$, $p < .001$. These results suggest that in general American-English speakers rate the standard American accent significantly higher on the task dimension of trait attributions than a foreign accent and a computer-synthesized accent.

For hypothesis 2, another two paired samples t-tests were used to make post hoc comparisons on the social dimension comparing the standard American accent against the foreign and computer-synthesized accent, respectively. The third paired samples t-test indicated that there was a significant difference in scores for the standard American accent ($M = 5.77$, $SD = 0.8$) and the foreign accent ($M = 4.61$, $SD = 1.17$) on the social dimension, $t(351) = 17.19$, $p < .001$. A fourth t-test indicated that there was a significant difference in scores for the standard American accent ($M = 5.77$, $SD = 0.8$) and the computer-synthesized accent ($M = 2.76$, $SD = 1.48$) on the social dimension, $t(351) = 36.24$, $p < .001$. These results suggest that in general American-English speakers rate the standard American accent significantly higher on the social dimension of trait attributions than a foreign accent and a computer-synthesized accent.

Overall, hypotheses 1 and 2 were supported as the results found significant main effects on the perceptions of trait attributions (task and social dimensions) for the standard American, foreign, and machine accents. More specifically, the results support that American-English speakers, in general, rate the standard American accent higher than the foreign accent and the computer-synthesized accent on both the task and social dimensions of trait attributions. Figure 1 shows the estimated marginal means of the accent types on the perceptions of trait attributions for both the task and social dimensions, thus further illustrating support for the hypotheses.

Figure 1

Estimated Marginal Means of Accent Types on Perceptions of Trait Attributions (Task and Social Dimensions)



Note. US = standard American accent. CA = computer-synthesized accent. FA = foreign accent.

Hypothesis 3 and 4

Hypothesis 3 and 4 proposed that American-English speaking participants with increasing levels of ethnocentrism will perceive the computer-synthesized accent increasingly more positive on the task and social dimension of trait attributions in comparison to a low prestige foreign accent. For the purpose of testing hypotheses 3 and 4, a linear regression for each dimension was conducted with additionally computed variables. New variables were computed related to the accents and ethnocentrism.

Firstly, two new variables were computed, which calculated the differences in ratings of the accents for both the task and social dimensions of trait attributions. Hypothesis 3 and 4

specifically mentioned the foreign and computer-synthesized accents; thus, the two new variables indicated how much higher the computer-synthesized accent was rated compared to the foreign accent. Thus, the ratings on the perceptions of trait attributions (Task and Social Dimensions) for the foreign accent were subtracted from the computer-synthesized accent. Secondly, a histogram examining the normal distribution of the ethnocentrism scores indicated that results were skewed towards the right. A new variable was computed using the log function in SPSS to normalize the distribution of the ethnocentrism scores.

To test hypotheses 3 and 4, two separate simple linear regression analyses were conducted, which predicted that increased levels of ethnocentrism among American-English speaker would be associated with increasingly more positive perceptions on the task and social dimensions of trait attributions for the computer-synthesized accents than the low prestige foreign accents. Furthermore, two separate figures were examined to illustrate the linear relationship between levels of ethnocentrism's and the above-mentioned accent types on the task and social dimensions.

For hypothesis 3, which predicted that the increasing levels of ethnocentrism among American-English speakers would be associated with increasingly more positive perceptions on the task dimension for the computer-synthesized accent compared to the low prestige foreign accent, a simple linear regression analysis was conducted for the perceptions of trait attributions. Table 6 presents a summary of the results from the simple linear regression related to the task dimension. The overall regression was statistically significant, $F(1, 350) = 53.01$, $p < .001$, with an R^2 of .132. An increase in the levels of ethnocentrism among American-English speakers was a significant predictor of positive ratings of the computer-synthesized accent (minus the foreign accent) on the task dimension, $\beta = .36$, $p < .001$. In other words, this means that as participants' levels of ethnocentrism increases, the computer-synthesized accent is increasingly rated more positively on the task dimension compared to the foreign accent.

Contrastingly, additional results from the post hoc comparisons yielded noteworthy findings on the perceptions of trait attributions on the task dimension for the foreign and computer-synthesized accents. The paired samples t-test for the task dimension indicated that there was a significant difference in scores for the foreign accent ($M = 4.17$, $SD = 1.15$) and the computer-synthesized accent ($M = 3.27$, $SD = 1.40$), $t(351) = 12.99$, $p < .001$. These results suggest that American-English speakers rate the foreign accent higher on the task dimension than the computer-synthesized accent. Hence, the inclusion of ethnocentrism as a predictor, indicates that ethnocentrism has a neutralizing effect on the perceptions of the task dimension. In other words, ethnocentrism decreases the differences in ratings between the machine and foreign accent, as the machine accent is perceived increasingly more positive compared to the foreign accent. These findings suggest that an increase in the levels of ethnocentrism is a strong predictor for perceptions of trait attributions on the task dimension.

Table 6

Simple Linear Regression Results for Ethnocentrism Predicting the Differences in the Rating of the Computer-synthesized and Foreign Accent on the Task Dimension [a]

Variable	B	SE B	β	*t*	*p*
(Constant)	-4.28	.47		-9.12	.000
Ethnocentrism	2.22	.30	.36	7.28	.000

Note. Significant at the $p < 0.05$ level

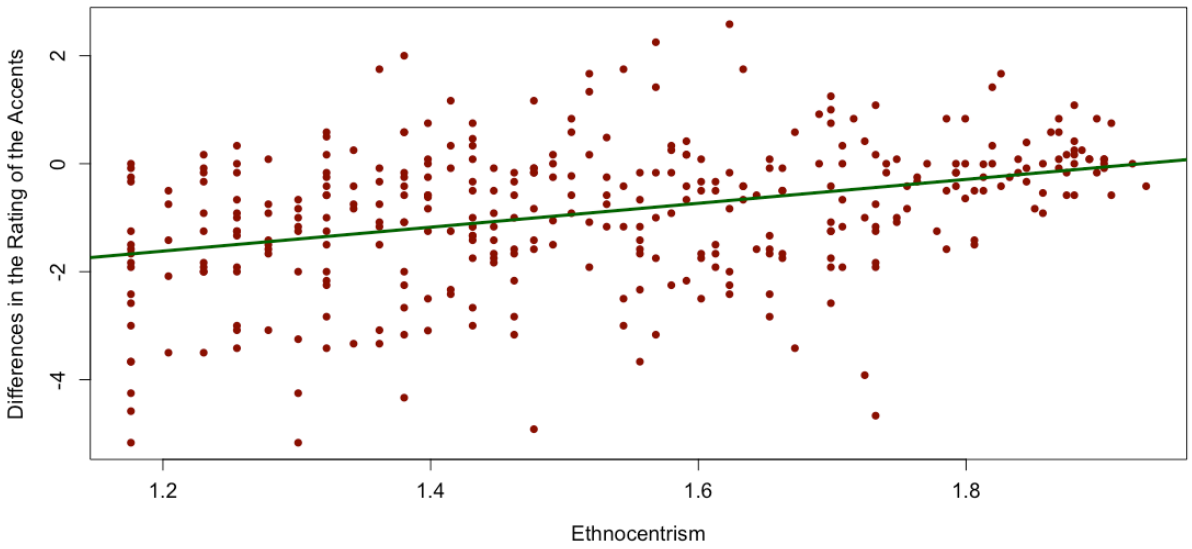
a. Dependent Variable: computer-synthesized accent minus foreign accent

Figure 2 illustrates the results from the simple linear regression for ethnocentrism predicting the differences in rating the computer-synthesized accent and the foreign accent on the task dimension. The figure shows a strong positive linear relationship between the levels of ethnocentrism and the differences in the perceptions of the computer-synthesized accent and low prestige foreign accents on the task dimension of trait attributions. More specifically, as levels of

ethnocentrism increase among American-English speakers, there is an increase in how much more positively the computer-synthesized accent is perceived compared to the foreign accent.

Figure 2

Simple Linear Regression for Ethnocentrism Predicting the Differences in the Rating of the Accents (Computer-Synthesized Accent and Foreign Accent) on the Task Dimension



Note. Differences in the Rating of the Accents = computer-synthesized accent minus foreign accent.

For hypothesis 4, which predicted that an increase in levels of ethnocentrism among American-English speakers would be associated with increasingly more positive perceptions on the social dimensions for the computer-synthesized accent compared to the low prestige foreign accent, a second simple linear regression analysis was conducted for the perceptions of trait attributions. Table 7 presents a summary of the results from the simple linear regression related to the social dimension. The overall regression was statistically significant, $F(1, 350) = 143.23$, $p < .001$, with an R^2 of .290. An increase in the levels of ethnocentrism in American-English speakers was a significant predictor of positive ratings of the computer-synthesized accent (minus the foreign accent) on the social dimension, $\beta = .54$, $p < .001$. In other words, this

means that as participants' levels of ethnocentrism increase, the computer-synthesized accent is increasingly rated more positively on the social dimension compared to the foreign accent.

Contrastingly, additional results from the post hoc comparisons yielded noteworthy findings on the perceptions of trait attributions for the social dimension for the foreign and computer-synthesized accents. The paired samples t-test for the social dimension indicated that there was a significant difference in scores for the foreign accent ($M = 4.61$, $SD = 1.17$) and the computer-synthesized accent ($M = 2.76$, $SD = 1.48$), $t(351) = 24.12$, $p < .001$. These results suggest that American-English speakers rate the foreign accent higher on the social dimension than the computer-synthesized accent. Hence, the inclusion of ethnocentrism as a predictor, indicates that ethnocentrism has a neutralizing effect on the perceptions of the social dimension. In other words, ethnocentrism decreases the differences in ratings between the machine and foreign accent, as the machine accent is perceived increasingly more positive compared to the foreign accent. These findings suggest that an increase in the levels of ethnocentrism is a strong predictor for perceptions of trait attributions on the social dimension.

Table 7

Simple Linear Regression Results for Ethnocentrism Predicting the Differences in the Rating of the Computer-synthesized and Foreign Accent on the Social Dimension [a]

Variable	B	SE B	β	*t*	*p*
(Constant)	-7.43	.47		-15.81	.000
Ethnocentrism	3.65	.31	.54	11.97	.000

Note. Significant at the $p < 0.05$ level

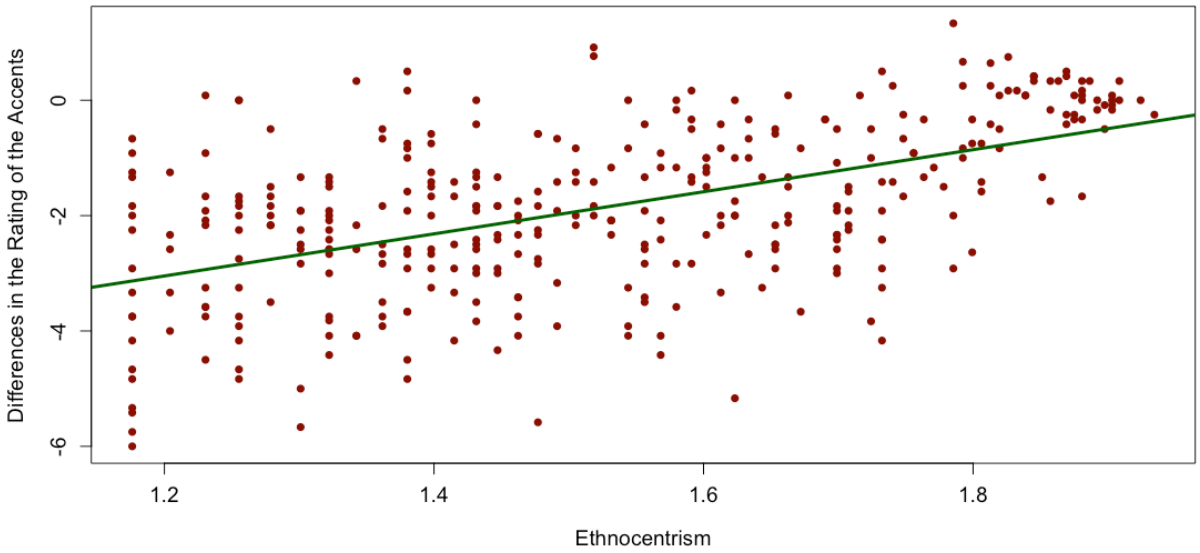
a. Dependent Variable: computer-synthesized accent minus foreign accent

Figure 3 illustrates the results from the simple linear regression for ethnocentrism predicting the differences in rating the computer-synthesized accent and the foreign accent on the social dimension. The figure shows a strong positive linear relationship between the levels of

ethnocentrism and the differences in perceptions of the computer-synthesized accent and low prestige foreign accents on the social dimension. More specifically, as the levels of ethnocentrism increase among American-English speakers, there is an increase in how much more positively the computer-synthesized accent is perceived compared to the foreign accent.

Figure 3

Simple Linear Regression for Ethnocentrism Predicting the Differences in the Rating of the Accents (Computer-Synthesized Accent and Foreign Accent) on the Social Dimension



Note. Differences in the Rating of the Accents = the computer-synthesized accent minus the foreign accent.

To summarize, hypotheses 3 and 4 proposed that American-English speaking participants with increasing levels of ethnocentrism will perceive the computer-synthesized accent more positively on the task and social dimensions compared to a low prestige foreign accent. The results indicate that ethnocentrism is a predictor for differences in increasingly more positive perceptions of the computer-synthesized and foreign accent on trait attributes (Task and Social dimensions). The results also indicate that the effect of ethnocentrism on the increasingly more positive perceptions seem to be larger for the social dimension of trait attributions than the task dimension (see β results in Table 6 and 7).The respective simple linear regressions for the task

and social dimensions found that as American-English speakers' levels of ethnocentrism increase, the computer-synthesized accent is increasingly perceived more positive compared to the low prestige foreign accent. The results of the simple linear regressions are statistically significant; thus, supporting hypotheses 3 and 4. Moreover, the additional findings from the post hoc comparison accentuate the influence of ethnocentrism on perceptions of trait attributions (Task and Social dimensions). The results from the paired samples t-tests indicate that in the absence of ethnocentrism as a variable, the foreign accent is perceived more favorably on trait attributions (task and social dimensions) than the computer-synthesized machine accent. Hence, these findings further support hypotheses 3 and 4, as the results emphasize that ethnocentrism is a predictor of differences in positive perceptions of the computer-synthesized and foreign accent on the task and social dimensions.

CHAPTER 5: DISCUSSION

The purpose of this study was to examine whether ethnocentrism influences the perception of traits in human and machine accents. To date, no study has examined the influence of ethnocentrism on the perception of standard American, foreign, and computer-synthesized accents (i.e., English US text-to-speech voice). Individuals are increasingly exposed to voice-generating machine entities. This study examines the role of ethnocentrism in in-group favoritism pertaining to human and non-human entities. Hence, American-English speakers were required to listen to the audio-recordings of the three accent categories and report their perceptions relating to traits. The trait attributions were separated into two categories, namely, task and social dimensions. The present study addressed and tested four hypotheses, which were both supported. This study expands the growing body of literature on accents and technology by merging them together.

Summary of Findings

Firstly, the present study posited that in general American-English speakers rate the standard American accent higher on the task and social dimensions of trait attributions than a low prestige foreign accent and a computer-synthesized accent, which the findings support. Results indicate that American-English speakers rated the standard American accent more favorably on the perceptions of trait attributions (task and social dimensions) compared to the foreign accent. Participation in the current study was only possible if respondents self-reported their fluency in American English. The results indicate that participants exhibited a bias toward the standard American accent since it was most similar to their own accent rather than the foreign accent. These results are congruent with research on the own-accent bias (e.g., Perrachione et al., 2010), whereby listeners' will rate their own accent more favorably than accents which sound dissimilar to themselves.

From the current study, it can be concluded that American-English speakers perceive the foreign accent to possess lower levels of task-orientated traits, namely intelligence, competence, knowledgeability, and persuasiveness; compared to a standard American accent, consistent with previous research (e.g., Lev-Ari & Keysar, 2010). Moreover, inconsistent with some previous research (e.g., Lindemann, 2003), this study suggests that the foreign accent is perceived to possess less social-orientated traits, namely pleasantness, likeability, friendliness, and trustworthiness; compared to the standard American accent by American-English speakers. These results add to the literature on attitudes towards standard and foreign accents, as the perception of attributions on task and social traits towards the accents could lead to negative attitudes toward foreign-accented speakers. Although, it is important to note that this study examined perceptions of trait attributions towards accented speech rather than attitudes towards accented speakers.

As posited by hypothesis 1 and 2, American-English speakers were also found to rate the standard American accent higher than a computer-synthesized accent on the task and social dimensions of trait attributions. The results suggest that human speech is rated more favorably compared to computer-synthesized speech, coinciding with previous research (e.g., Andrist et al., 2015). These findings are important, as previously used text-to-speech systems were not as advanced as the Apple MacBook system used for the current study. Essentially, although the computer-synthesized accent was perceived as sounding robotic by participants, compared to previous studies (e.g., Tamagawa et al., 2011), the audio recordings from this study sounded less robotic. As participants rated human speech more favorably than the computer-synthesized accent, the findings add to the literature on machine voices.

The current study's final prediction posited that as the levels of ethnocentrism increase among American-English speakers, the positive perceptions on the task and social dimensions of

trait attributions toward the computer-synthesized accent in comparison to a low prestige foreign accent will increase.

The study found that as American-English speakers' levels of ethnocentrism increase, the computer-synthesized accent was increasingly perceived more positively on the task and social dimensions of trait attributions than the low prestige foreign accent. The results indicate that ethnocentrism plays a remarkable role in American-English speakers' categorization of these accent types adding to our understanding of in-group favoritism; and consequent adjustments of individuals' perceptions of trait attributions towards computer-synthesized and foreign accents.

The results from hypotheses 3 and 4 indicate that American-English speakers perceive the computer-synthesized accent increasingly more positively on trait attributions on the task and social dimensions compared to the low prestige foreign accent, as one's levels of ethnocentrism increases. Contrastingly, the additional findings from the post hoc comparison of hypotheses 1 and 2 showed that ethnocentrism's absence led participants to perceive the foreign accent more favorably than the computer-synthesized accent. This study indicates that accents are strong indicators of in- and out-group categorization, congruent with previous research (e.g., Neuliep, & McCroskey, 1997). An explanation for the findings is that the computer-synthesized accent acted as a social cue for participants with higher levels of ethnocentrism to categorize the non-human machine accent as belonging to the same group, namely American-English accented speech. Moreover, the study provides empirical evidence of in-group categorization invoking a shift in participants' perceptions of trait attributions related to the computer-synthesized accent.

This study is essential to the growing body of literature on attitudes and perceptions toward not only synthesized speech but also foreign accents. These findings illustrate ethnocentrism's effect on changes in in-group favoritism (from humans toward robots) because the results show how easily ethnocentrism led to a shift in perceptions of trait attributions. Consequently, the shift in perception leads to more negative and even discriminatory perceptions

towards foreign accents over machine accents as it pertains to trait attributions, which the next section will explore in more detail.

Implications

This study offers insights into the investigation of machine accents. First, this study has implications for accent research and the design of machine voices. This study's far-reaching implication is the illustration of human temperament and beliefs that influence their perceptions towards other humans and machines. This section will explore the theoretical and practical implications of this research's findings as it concerns the investigation of machine accents.

Theoretically, the results of hypotheses 1 and 2 are congruent with previous research regarding own-group favoritism. More specifically, the study found that participants rated the standard American accent more favorably on the trait attributions than the foreign accent and the computer-synthesized accent. As participation was only possible if participants self-reported speaking American-English fluently, one can assume that they sounded similar to the standard American accent. Thus, the findings indicate that participants perceived the accent sounding most similar to their own more favorably on the trait attributions, thereby exhibiting an own-accent bias. This is similar to previous research (see Perrachione et al., 2010), as participants perceived the foreign and computer-synthesized accent less favorably as it sounded more dissimilar to themselves than the standard American accent provided.

Moreover, the results indicate that own-group favoritism presently still plays a role in perceptions of trait attributions. Hansen, Raki, and Steffens (2014) noted that the investigation of accent-based discrimination interventions is relatively scarce. Their investigation found that German participants in the control group would evaluate job applicants sounding similar to themselves (i.e., standard accent) as more competent than a non-standard accented job applicant (i.e., Turkish). However, participants in the experimental condition had to speak English before interviewing job applicants, hence evaluating the standard and non-standard accented speakers as

similarly competent. In this instance, the interviewer became a non-standard accented speaker, which allowed the participant to view the situation from a different perspective (Hansen, Raki, & Steffens, 2014).

This study indicates that own-accent favoritism plays an important role in person perception. Hence, research should investigate how own-accent favoritism could be utilized to decrease accent-based discrimination. Alternatively, as in Hansen, Raki, and Steffens's (2014) study shifting the group dynamics could also be a potential starting point for future investigations.

This study demonstrates that an individuals' accent plays a role in perceptions towards trait attributions of synthesized speech; thus, the findings of hypotheses 1 and 2 provide some practical implications. The results suggest that a standard American accent is rated more favorably on the task and social dimensions than a computer-synthesized accent. These perceptions lead to a computer-synthesized voice to be viewed as less knowledgeable, competent, intelligent, and persuasive as well as less friendly, pleasant, likable, and trustworthy than a human voice. Newman (2018) noted that the use of voice-activated devices such as Amazon Alexa or Google Assistant is rapidly increasing. According to findings in the US, 14% of adults regularly use such devices at a basic level (Newman, 2018). The use of devices is lower in European countries such as the UK (10%) and Germany (5%), although those numbers are quickly increasing (Newman, 2018). This study indicates a need to identify the reasoning behind individuals' favorability of standard accented speech over machine voices. Even more so as it extends not only to standard accented speech but includes non-standard accented speech. Although interactions between humans and machines have not currently advanced to include conversations, this might not be far in our future.

A study investigating using a chat and voice-based conversational agent for workplace reflection (Kocielnik, Avrahami, Marlow, Lu, & Hsieh, 2018) opens up another aspect of the

future relationship between humans and machine interactions. The authors noted that the added advantage of interacting with a conversational agent allowed individuals to engage more and be more personal in their reflections. As human speech is favored over synthesized speech, the design of the synthesized speech seems vital to increase user satisfaction and comfortability.

Hence, future research should consider exploring perceptions of the design of text-to-speech systems and the various types of text-to-speech technologies in circulation.

The results of hypotheses 3 and 4 demonstrate that as the levels of ethnocentrism increase, the computer-synthesized accent is increasingly perceived more positively compared to the foreign accent. This is contrary to the perception of these accents in the absence of ethnocentrism, as this more positive perception was not observable. Hence, ethnocentrism has a key role in influencing individuals' perceptions, whereby there is a shift in in-group perceptions towards humans. This shift in perceptions could indicate that ethnocentrism contributes to the dehumanization of human out-groups. Thus, this shift in in-group perceptions (from humans towards robots) could lead to negative perceptions and out-group discrimination towards foreign accents. In other words, individuals with stronger ethnocentrism perceive the non-human out-group (machine accent) increasingly more positively on trait attributions compared to the human out-group (foreign accent). Furthermore, ethnocentrism leading to prejudice and discrimination of human accented speech has been noted in previous research (e.g., Heblich et al., 2015; Neuliep & McCroskey, 1997; Gagnon & Bourhis, 1996). However, the influence of ethnocentrism is not solely bound to human groups, but instead extending to non-human entities indicates some theoretical implications.

These findings are concerning, as it indicates that individuals' beliefs can circumvent human solidarity and unity. Even more concerning, in this instance, that the neutral entity is a non-human synthesized accent which is perceived more favorably than a human foreign accent. Although, it is important to note that the study did not examine participants' explicit attitudes

towards these accents or establish preference. Future research could further examine individuals' attitudes and whether the synthesized accent was in actuality preferred over the foreign accent. In other words, the study provides empirical evidence of individuals' beliefs influencing perceptions, thus it is important to further investigate whether these beliefs affect actual speech preference and behavioral preference. Knowledge in both areas of preference would be useful for developing future human and machine relationships, and address discrimination because of ethnocentric beliefs.

Furthermore, hypotheses 3 and 4 bring implications when considering future relationships between humans and machines. The findings indicate that discrimination and prejudice can contribute to the deterioration of human-to-human relations as human-to-machine interactions are rapidly expanding in the world. Future research should attempt to replicate the findings in different countries and languages as computer-synthesized machine voices are being used globally. Since ethnocentrism is observable across countries or cultures, it would be essential to explore whether similar results can be found across various languages. Exploring the levels of ethnocentrism and perceptions of foreign accents and computer-synthesized accents across various languages would be beneficial in adding to the literature.

Future research should also investigate whether ethnocentrism or other phenomena such as race, ethnicity, or gender influence the perception of robot or AI devices. Consistent with the current findings, certain individuals may prefer interacting with robots over foreign workers or caregivers. Wright (2019) stated that this might not be an actual possibility as a robot would decrease care quality. However, as technology continually advances, there might be a time in the future when the skills of a robot could be equal or exceed those of human caregivers. Hence, this study contributes to our understanding of the societal impact of in-group favoritism shifting away from human towards machine. Future research should continue to explore whether discrimination, prejudices, or stereotypes influence individuals to shift their perceptions to

consider a non-human outgroup more positively than a human outgroup. As this study only used synthesized speech, it seems essential to explore whether similar results can be replicated with embodied robots or AI devices. This seems a critical future direction, as this is an underdeveloped research area with potentially significant societal impacts.

Limitations

This study has several limitations. One limitation of this study is the number of accents used. The study settled on using nine audio-recordings, evenly categorized into their respective accent categories. This number was chosen for generalizability reasons, although the repetitiveness of recordings and measures could have decreased the participants' attentiveness. There is also the chance of participants becoming aware of the purpose of the study, which could lead to social desirability. Thus, participants might have rated the accents regarding what they perceived the study wanted to study. Future studies could include a questionnaire measuring participants' social desirability to exclude possible biases influencing the study results.

This study only used male audio-recordings, as previous literature mentioned that gender stereotypes apply to the perception of synthesized speech (e.g., Lee et al., 2000). This reduces the generalizability of the findings as representing perceptions of human and synthesized speech across gender. Hence, this is a limitation of the current study, which future research should address. Future research should attempt to replicate the findings using both male and female audio-recordings. This would make the study not only more generalizable but would also add another aspect to the research. More specifically, it would explore the influence of ethnocentrism in general as well as across gendered audio-recordings. Another aspect that could be explored is the categorization and in-group favoritism of human and synthesized speech not only concerning ethnocentrism but also to gender.

This study is also limited as the majority of participants were Caucasian females. This means that the results are not representative of the population, hence decreasing the

generalizability. The limitations of mainly female participants listening to male audio recordings could mean that different aspects of the recording influenced their perceptions. A similar limitation comes from the participants being mostly Caucasian and the standard American accent being mostly associated with that ethnicity. Here the question becomes whether more minority ethnicities in the population pool would have had similar results in the findings. Thus, a larger and more diverse sample with equal numbers of males and females could result in the accents' perceptions to change on the task and social dimensions.

A potential limitation of this study was that the thickness of an accent was not considered. The selection of recordings only considered that the accent needed to be discernible. Thus, the future direction could examine the thickness of the accents. This would further our understanding of ethnocentrism's influence, as the research could explore whether the thickness of the accent would provide similar results or even more pronounced differences in ratings as the levels of ethnocentrism increases. Moreover, this could be further extended to include low prestige foreign accents, regional accents and dialects, and high prestige foreign accents. This research would add another dimension to the influence of ethnocentrism on perceptions of accented speech. Thus, such a study could explore whether ethnocentrism would be a predictor of differences in rating accents across different US regions. Additionally, whether foreign accents are categorized differently and consequently, ethnocentrism influences ratings of these accents differently compared to computer-synthesized accents.

This study concerned itself with perceptions of the accents as it pertains to the task and social traits, thus providing assumptions into individuals' attitudes of the accents. This could be a potential limitation that future research could explore by investigating individuals' actual attitudes of the various accent categories investigated in this study. Hence, expanding the research into individuals' explicit attitudes could provide further insight. In line with this proposition, research could investigate individuals' actual preferences by having participants

ranking the accents. This would allow researchers to determine whether individuals prefer computer-synthesized or foreign accents, depending on the levels of ethnocentrism.

This study used the transcript "Comma Gets a Cure" made available by IDEA, which could be a further limitation. The transcript could be a limitation in determining specific task- or social-orientated trait attributions. The results indicated that ethnocentrism's influence on perceptions had a more substantial effect on the social dimension than the task dimension. Future research could replicate this study to investigate whether transcripts with a situational context would change the findings. In other words, should audio-recordings use more knowledge or persuasion relevant terminology, would the ratings on the task dimension be more strongly influenced. Alternatively, having transcripts that are more sociable, rather than just a neutral story, might also influence the results differently. Hence, future research should include situational contexts to the transcripts as this would facilitate the perceptions of trait attributions.

Conclusion

This study has added to research on perceptions of accents related to human and non-human speech. More specifically, the findings demonstrated the importance of designing synthesized speech as sounding more human-like, as both standard American and foreign accents are perceived more favorably on task and social traits compared to computer-synthesized accents. One of the major findings from this study was that ethnocentrism was a predictor for perceiving the computer-synthesized accent increasingly more positive compared to the low prestige foreign accent. This study illustrates the influence of ethnocentrism shifting participants' perceptions of the low prestige accent and the computer-synthesized accent.

Overall, this investigation throws light on the practical importance of perceptions of speech, which includes computer-synthesized machine speech. The findings imply that biases toward foreign accents in preference for machine accents among those with higher levels of ethnocentrism can have significant impact on the human relations in the near future. Although,

the advancement of technology blurs the boundaries between human and machine sounding speech systems, this could change how we humans see ourselves, as well as other human beings in relation to intelligent machines.

REFERENCES

Andrist, S., Ziadee, M., Boukaram, H., Mutlu, B., & Sakr, M. (2015). Effects of Culture on the Credibility of Robot Speech: A Comparison between English and Arabic. *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction - HRI '15*, 157–164. https://doi.org/10.1145/2696454.2696464

Bestelmeyer, P. E. G., Belin, P., & Ladd, D. R. (2015). A Neural Marker for Social Bias Toward In-group Accents. *Cerebral Cortex (New York, NY)*, *25*(10), 3953–3961. https://doi.org/10.1093/cercor/bhu282

Bresnahan, M. J., Ohashi, R., Nebashi, R., Liu, W. Y., & Shearman, S. M. (2002). Attitudinal and affective response toward accented English. *Language & Communication*, *22*, 171–185. Retrieved from https://www.researchgate.net/publication/255983261

Burgoon, J. K., Guerrero, L. K., & Floyd, K. (2010). Visual and Auditory Codes: Kinesics and Vocalics. In J. K. Burgoon, L. K. Guerrero, & K. Floyd (Eds.), *Nonverbal Communication* (pp. 111-144). [Kindle DX Version]. Retrieved from Amazon.com

Cambre, J., & Kulkarni, C. (2019). One Voice Fits All?: Social Implications and Research Challenges of Designing Voices for Smart Devices. *Proceedings of the ACM on Human-Computer Interaction*, *3*(CSCW), 1–19. https://doi.org/10.1145/3359325

Cargile, A. C., Giles, H., Ryan, E. B., & Bradac, J. J. (1994). Language attitudes as a social process: A conceptual model and new directions. *Language & Communication*, *14*(3), 211–236. https://doi.org/10.1016/0271-5309(94)90001-9

Center for the Study of Adult Literacy (n.d.). *U.S. demographic survey.* Retrieved from http://csal.gsu.edu/content/us-demographic-survey

Chakraborty, R. (2017). A Short note on Accent–bias, Social Identity and Ethnocentrism. *Advances in Language and Literary Studies*, *8*(4), 57–64. https://doi.org/10.7575/aiac.alls.v.8n.4p.57

Coupland, N., & Bishop, H. (2007). Ideologised values for British accents. *Journal of Sociolinguistics*, *11*(1), 74–93. https://doi.org/10.1111/j.1467-9841.2007.00311.x

Dailey, R. M., Giles, H., & Jansma, L. L. (2005). Language attitudes in an Anglo-Hispanic context: the role of the linguistic landscape. *Language & Communication*, *25*(1), 27–38. https://doi.org/10.1016/j.langcom.2004.04.004

Dragojevic, M., Giles, H., & Watson, B. M. (2013) Language Ideologies and Attitudes: A Foundational Framework. In H. Giles & B. M. Watson (Eds). *The Social Meanings of Language, Dialect, and Accent: International Perspectives on Speech Styles* (pp. 1–25).

Fuertes, J. N., & Gelso, C. J. (2000). Hispanic counselors' race and accent and Euro Americans' universal-diverse orientation: A study of initial perceptions. *Cultural Diversity and Ethnic Minority Psychology*, *6*(2), 211. https://doi.org/10.1037/1099-9809.6.2.211

Fuertes, J. N., Potere, J. C., & Ramirez, K. Y. (2002). Effects of speech accents on interpersonal evaluations: implications for counseling practice and research. *Cultural Diversity and Ethnic Minority Psychology*, *8*(4), 346. https://doi.org/10.1037/1099-9809.8.4.347

Fussell, S. R., Kiesler, S., Setlock, L. D., & Yew, V. (2008). How people anthropomorphize robots. *Proceedings of the 3rd International Conference on Human Robot Interaction - HRI '08*, 145. https://doi.org/10.1145/1349822.1349842

Gagnon, A., & Bourhis, R. Y. (1996). Discrimination in the minimal group paradigm: Social identity or self-interest? *Personality and Social Psychology Bulletin, 22*(12), 1289-1301.

Gong, L. (2008). The boundary of racial prejudice: Comparing preferences for computer-synthesized White, Black, and robot characters. *Computers in Human Behavior*, *24*, 2074–2093.

Hansen, K., Raki, T., & Steffens, M. C. (2014). When Actions Speak Louder Than Words: Preventing Discrimination of Nonstandard Speakers. *Journal of Language and Social Psychology*, *33*(1), 68–77. https://doi.org/DOI: 10.1177/0261927X13499761

Heblich, S., Lameli, A., & Riener, G. (2015). The Effect of Perceived Regional Accents on Individual Economic Behavior: A Lab Experiment on Linguistic Performance, Cognitive Ratings and Economic Decisions. *PLOS ONE*, *10*(2), e0113475. https://doi.org/10.1371/journal.pone.0113475

Hopkins, D. J. (2015). The Upside of Accents: Language, Inter-group Difference, and Attitudes toward Immigration. *British Journal of Political Science; Cambridge*, *45*(3), 531–557. https://doi.org/10.1017/S0007123413000483

Kamm, C., Walker, M. & Rabiner, L. (1997, February). The role of speech processing in human-computer intelligent communication. Paper presented at NSF Workshop on human-centered systems: Information, interactivity, and intelligence. Retrieved October 24, 2019, from http://www.ifp.uiuc.edu/nsfhcs/talks/rabiner.html

Kinzler, K. D., Corriveau, K. H., & Harris, P. L. (2011). Children's selective trust in native-accented speakers. *Developmental Science*, *14*(1), 106–111. https://doi.org/10.1111/j.1467-7687.2010.00965.x

Kinzler, K. D., Dupoux, E., & Spelke, E. S. (2007). The native language of social cognition. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(30), 12577–12580. https://doi.org/10.1073/pnas.0705345104

Kocielnik, R., Avrahami, D., Marlow, J., Lu, D., & Hsieh, G. (2018). Designing for Workplace Reflection: A Chat and Voice-Based Conversational Agent. *Proceedings of the 2018 on Designing Interactive Systems Conference 2018 - DIS '18*, 881–894. https://doi.org/10.1145/3196709.3196784

Lee, K. M., & Nass, C. (2004). The Multiple Source Effect and Synthesized Speech. *Human Communication Research*, *30*(2), 182–207. https://doi.org/10.1111/j.1468-2958.2004.tb00730.x

Lee, E., Nass, C., & Brave, S. (2000). Can computer-generated speech have gender?: An experimental test of gender stereotype. *In CHI '00 Extended Abstracts on Human Factors in Computer Systems*, 289–290. ACM Press.

Lev-Ari, S., & Keysar, B. (2010). Why don't we believe non-native speakers? The influence of accent on credibility. *Journal of Experimental Social Psychology*, *46*(6), 1093–1096. https://doi.org/10.1016/j.jesp.2010.05.025

Lindemann, S. (2003). Koreans, Chinese or Indians? Attitudes and ideologies about non-native English speakers in the United States. *Journal of Sociolinguistics*, *7*(3), 348–364. https://doi.org/10.1111/1467-9481.00228

Lippi-Green, R. (1994). Accent, Standard Language Ideology, and Discriminatory Pretext in the Courts. *Language in Society*, *23*(2), 163–198.

Munro, M. J., Uriiversity, S. F., & Derwing, T. M. (1995). Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech. *Language and Speech, 38*(3), 289–306.

Nass, C., Fogg, B. J., & Moon, Y. (1996). Can computers be teammates? *International Journal of Human-Computer Studies*, *45*(6), 669–678. https://doi.org/10.1006/ijhc.1996.0073

Nass, C., & Gong, L. (2000). SPEECH INTERFACES from an EVOLUTIONARY PERSPECTIVE. *Communications of the ACM*, *43*(9), 36–43. https://doi.org/10.1145/348941.348976

Nass, C., Isbister, K., & Lee, E. J. (2000). Truth is beauty: Researching embodied conversational agents. In J. Castells (Ed.), *Embodied Conversational Agents*, 374-402. Cambridge, MA: MIT Press.

Nass, C., & Lee, K. (2000). Does computer-generated speech manifest personality? an experimental test of similarity-attraction. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* 329-336.

Nass, C., & Moon, Y. (2000). Machines and Mindlessness: Social Responses to Computers. *Journal of Social Issues*, *56*(1), 81–103. https://doi.org/10.1111/0022-4537.00153

Neuliep, J. W., & McCroskey, J. C. (2013). Ethnocentrism Scale. *Measurement Instrument Database for the Social Science*. Retrieved from www.midss.ie

Neuliep, J. W., & McCroskey, J. C. (1997). The Development of a U.S. and Generalized Ethnocentrism Scale. *Communication Research Reports*, *14*(4), 385–398. https://doi.org/10.1080/08824099709388682

Neuliep, J. W., & Speten-Hanse, K. M. (2013). The influence of ethnocentrism on social perceptions of nonnative accents. *Language & Communication*, *33*, 167–176. http://dx.doi.org/10.1016/j.langcom.2013.05.001

Nesdale, D., & Rooney, R. (1996). Evaluations And Stereotyping Of Accented Speakers By Pre-Adolescent Children. *Journal of Language and Social Psychology*, *15*(2), 133–154.

Newman, N. (2018). *The future of voice and the implications for news* (Reuters Institute for the Study of Journalism Reports). Reuters Institute for the Study of Journalism. https://ora.ox.ac.uk/objects/uuid:c1ab8cef-79d4-4be6-8c8b-3a490483741b

Parr, M. (2012). The future of text-to-speech technology: How long before it's just one more thing we do when teaching reading? *Procedia - Social and Behavioral Sciences*, *69*, 1420–1429. https://doi.org/doi:10.1016/j.sbspro.2012.12.081

Perrachione, T. K., Chiao, J. Y., & Wong, P. C. M. (2010). Asymmetric cultural effects on perceptual expertise underlie an own-race bias for voices. *Cognition*, *114*(1), 42–55. https://doi.org/10.1016/j.cognition.2009.08.012

Phinney, J. (1992). The Multigroup Ethnic Identity Measure: A new scale for use with adolescents and young adults from diverse groups. *Journal of Adolescent Research, 7*, 156-176.

Rau, P. L. P., Li, Y., & Li, D. (2009). Effects of communication style and culture on ability to accept recommendations from robots. *Computers in Human Behavior*, *25*(2), 587–595. https://doi.org/10.1016/j.chb.2008.12.025

Reddington, E. (2008). *Native Speaker Response to Non-Native Accent: A Review of Recent Research*. *8*(2), 1–27. Retrieved from https://doi.org/10.7916/D8988KK8

Robertson, J. (2007). Robo Sapiens Japanicus: Humanoid Robots and the Posthuman Family. *Critical Asian Studies*, *39*(3), 369–398. https://doi.org/10.1080/14672710701527378

Robertson, J. (2014). Human Rights vs. Robot Rights: Forecasts from Japan. *Critical Asian Studies*, *46*(4), 571–598. https://doi.org/10.1080/14672715.2014.960707

Ryan, E. B. (1983). Social Psychological Mechanisms Underlying Native Speaker Evaluations of Non-Native Speech. *Studies in Second Language Acquisition*, *5*(2), 148–159. Retrieved from https://www.jstor.org/stable/44487275

Sharma, S., Shimp, T. A., & Shin, J. (1995). Consumer Ethnocentrism: A Test of Antecedents and Moderators. *Journal of the Academy of Marketing Science*, *23*(1), 26–37. https://doi.org/10.1007/BF02894609

Stern, S. E. (2008). Computer-Synthesized Speech and Perceptions of the Social Influence of Disabled Users. *Journal of Language and Social Psychology*, *27*(3), 254–265. https://doi.org/10.1177/0261927X08318035

Stern, S. E., Mullennix, J. W., Dyson, C., & Wilson, S. J. (1999). The Persuasiveness of Synthetic Speech versus Human Speech. *Human Factors*, *41*(4), 588–595. https://doi.org/10.1518/001872099779656680

Suchman, L. A. (2009). *Human-Machine Reconfigurations: Plans and Situated Actions* (2nd ed.). New York: Cambridge University Press.

Tamagawa, R., Watson, C. I., Kuo, I. H., MacDonald, B. A., & Broadbent, E. (2011). The Effects of Synthesized Voice Accents on User Perceptions of Robots. *International Journal of Social Robotics*, *3*(3), 253–262. https://doi.org/10.1007/s12369-011-0100-4

Warner, J. L. (Producer), & Cukor, G. (Director). (1964). *My Fair Lady* [Motion picture]. United States: Warner Bros. Pictures.

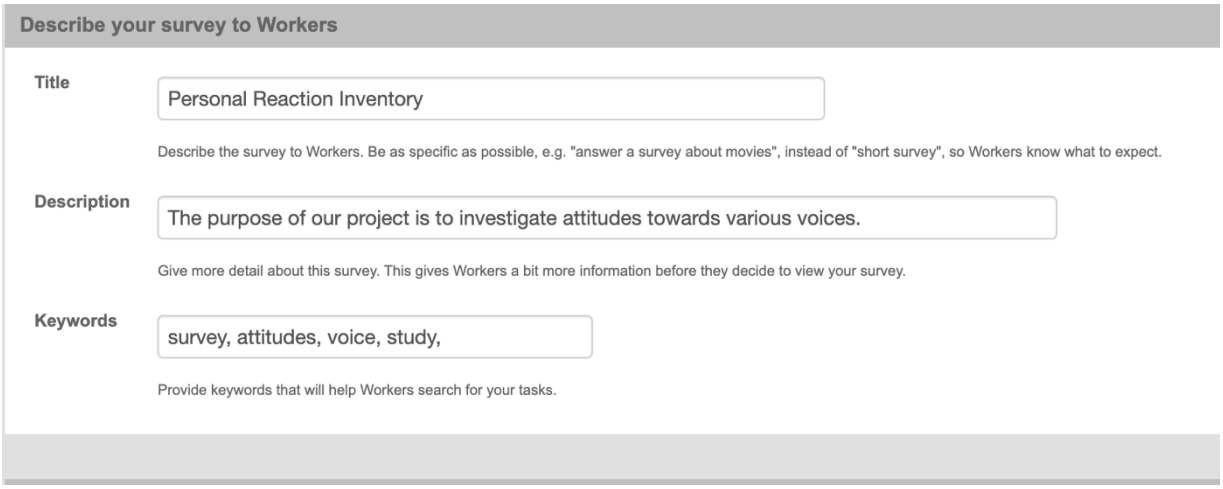
Wiese, E., Shaw, T., Lofaro, D., & Baldwin, C. (2017). Designing Artificial Agents as Social Companions. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *61*(1), 1604–1608. https://doi.org/10.1177/1541931213601764

Wright, J. (2019). Robots vs migrants? Reconfiguring the future of Japanese institutional eldercare. *Critical Asian Studies*, *51*(3), 331–354. https://doi.org/10.1080/14672715.2019.1612765

APPENDIX A: AMAZON MTURK POST

MTurk Posting

Below is a screenshot of the information provided to participants via Amazon MTurk.



APPENDIX B: SCREENING QUESTIONS

Screening Questions

These questions were specifically coded to appear as a worker qualification on Amazon MTurk. The screening questions were used to ascertain whether the participants were American participants.

1. Were you born in the United States of America?

    (1) Yes
    (2) No

2. During the last 10 years, have you lived in the USA for the majority of the time?

    (1) Yes
    (2) No

3. Would you consider yourself to be a fluent speaker of American English?

    (1) Yes
    (2) No

APPENDIX C: CONSENT FORM

Consent to Participate in a Research Project

You are being asked to participate in a research study conducted by the Department of Communicology at the University of Hawaii at Mānoa.

What am I being asked to do?
If you participate in this project, you will be asked to complete an online survey.

Taking part in this study is your choice.
You have the choice of taking part, or not taking part in this study. You can withdraw from the study at any time with no payment.

Why is this study being done?
The purpose of our project is to investigate attitudes towards various voices.

What will happen if I decide to take part in this study?
Part one will ask you to listen to audio-recordings and rate them on a scale from one to seven. The second part will ask your opinion on various statements. The survey will take about 30 minutes. After completing the survey, you will be given a random ID number to insert into MTurk for payment.

What are the risks and benefits of taking part in this study?
There are no foreseeable risks, discomforts, or inconveniences to the participant. There are no significant physical or psychological risks to participants that would cause the researchers to terminate the study.

By participating in this study, you will help expand our knowledge on human communication.

Privacy and Confidentiality:
Any information that is obtained from this study will remain confidential. Confidentiality will be maintained by means of extrapolating data and separating collected data from any personal information. Participants are not audio recorded nor videotaped during this study. The data from this study may be used for future research.

Compensation:
You will receive $3.00 as compensation for participating in and completing this survey.

Questions:
If you have any questions about this study, please email mliecht@hawaii.edu. You may contact the UH Human Studies Program at 808.956.5007 or uhirb@hawaii.edu to discuss problems, concerns and questions. Please visit http://go.hawaii.edu/jRd for more information on your rights as a research participant.

If you consent to participate in this project, please click the "Next" button.

APPENDIX D: TRANSCRIPT

This is the full transcript, however only the bolded paragraph was used for the stimulus tapes.

Comma gets a Cure

Well, here's a story for you: Sarah Perry was a veterinary nurse who had been working daily at an old zoo in a deserted district of the territory, so she was very happy to start a new job at a superb private practice in North Square near the Duke Street Tower. That area was much nearer for her and more to her liking. Even so, on her first morning, she felt stressed. She ate a bowl of porridge, checked herself in the mirror and washed her face in a hurry. Then she put on a plain yellow dress and a fleece jacket, picked up her kit and headed for work.

When she got there, there was a woman with a goose waiting for her. The woman gave Sarah an official letter from the vet. The letter implied that the animal could be suffering from a rare form of foot and mouth disease, which was surprising, because normally you would only expect to see it in a dog or a goat. Sarah was sentimental, so this made her feel sorry for the beautiful bird.

Before long, that itchy goose began to strut around the office like a lunatic, which made an unsanitary mess. The goose's owner, Mary Harrison, kept calling, "Comma, Comma," which Sarah thought was an odd choice for a name. Comma was strong and huge, so it would take some force to trap her, but Sarah had a different idea. First she tried gently stroking the goose's lower back with her palm, then singing a tune to her. Finally, she administered ether. Her efforts were not futile. In no time, the goose began to tire, so Sarah was able to hold onto Comma and give her a relaxing bath.

Once Sarah had managed to bathe the goose, she wiped her off with a cloth and laid her on her right side. Then Sarah confirmed the vet’s diagnosis. Almost immediately, she remembered an effective treatment that required her to measure out a lot of medicine. Sarah warned that this course of treatment might be expensive—either five or six times the cost of penicillin. I can’t imagine paying so much, but Mrs. Harrison—a millionaire lawyer— thought it was a fair price for a cure.

APPENDIX E: INFORMAL PRE-TESTING RESULTS

Here is a summary of the descriptive statistics from the informal pre-testing.

Table 8

Mean Scores, Standard Deviation and Reliability Results from the Informal Pre-testing

Accent	Mean Scores (*SD*)	Reliability
General American 1	6.42 (.86)	.96
General American 5	5.97 (1.54)	.98
General American 9	5.13 (1.89)	.99
General American 10	6.23 (.80)	.82
General American 13	6.19 (1.08)	.94
Alex (MacBook)	6.33 (1.08)	.99
Bruce (MacBook)	6.83 (.36)	.89
Fred (MacBook)	7.00 (.00)	1.00
Tom (MacBook)	6.79 (.59)	.96
Joey (Amazon Polly)	6.11 (1.54)	1.00
Matthew (Amazon Polly)	4.75 (1.21)	.97
Chile	5.73 (.89)	.88
Colombia	4.67 (1.49)	.70
Peru	5.80 (1.12)	.93
Venezuela	4.00 (.24)	-7.80
Turkey	5.93 (1.28)	.97
Beijing	4.33 (2.39)	.98
Anhui	5.87 (1.26)	.96
India	5.67 (1.31)	.83
Japan	5.13 (1.50)	.99
Vietnam	6.53 (.51)	.72

Note. Reliability reflects Cronbach's alpha (α) for all accents.

APPENDIX F: GENERALIZED ETHNOCENTRISM SCALE

Ethnocentrism Scale (Tamagawa et al., 2011)

Below are items that relate to the cultures of different parts of the world. Work quickly and record your first reaction to each item. There are no right or wrong answers.

Please indicate the degree to which you agree or disagree with each item using the following seven-point scale: Strongly Disagree (1) to Strongly Agree (7)

1. Most other cultures are backward compared to my culture.
2. My culture should be the role model for other cultures.
3. People from other cultures act strange when they come to my culture.
4. Lifestyles in other cultures are just as valid as those in my culture.
5. Other cultures should try to be more like my culture.
6. I am not interested in the values and customs of other cultures.
7. People in my culture could learn a lot from people in other cultures.
8. Most people from other cultures just don't know what's good for them.
9. I respect the values and customs of other cultures.
10. Other cultures are smart to look up to our culture.
11. Most people would be happier if they lived like people in my culture.
12. I have many friends from different cultures.
13. People in my culture have just about the best lifestyles of anywhere.
14. Lifestyles in other cultures are not as valid as those in my culture.
15. I am very interested in the values and customs of other cultures.
16. I apply my values when judging people who are different.
17. I see people who are similar to me as virtuous.
18. I do not cooperate with people who are different.
19. Most people in my culture just don't know what is good for them.
20. I do not trust people who are different.
21. I dislike interacting with people from different cultures.
22. I have little respect for the values and customs of other cultures.

Questions 4, 7, and 9 will be recoded, following the format: 1 = 7; 2 = 6; 3 = 5; 4 = 4; 5 = 3; 6 = 2; 7 = 1.

Questions 3, 6, 12, 15, 16, 17, and 19 will be dropped.

After the previous questions have been recoded, the responses for all remaining 15 items will be added together to get a composite ethnocentrism score.

APPENDIX G: PERCEPTIONS OF TRAIT ATTRIBUTIONS

Perceptions of Trait Attributions: Task and Social Dimensions

Instructions

In part 1, you will be asked to listen to audio recordings. Please make sure to be in an environment without distractions before pressing the play button. After each recording, please indicate the degree to which you agree or disagree with each statement

After having listened to the recording, indicate to what degree you agree/disagree with each statement provided.

Task Dimension

1. This speaker sounds knowledgeable.
2. This speaker sounds intelligent.
3. This speaker sounds competent.
4. This speaker sounds persuasive.

Social Dimension

1. This speaker sounds pleasant.
2. This speaker sounds likeable.
3. This speaker sounds friendly.
4. This speaker sounds trustworthy.

The results for the two dimensions will be added up separately and then averaged out to receive a score for each dimension.

APPENDIX H: DEMOGRAPHIC QUESTIONNAIRE

Demographic Information Questionnaire

Below are a series of questions, which will ask you about demographic information. There are no right or wrong answers.

1. What is your gender?

    (1) Male
    (2) Female
    (3) Other

2. What is your age?

    ____

3. Which of the following best describes your ethnic or racial background?

    (1) Caucasian
    (2) Black
    (3) Hispanic
    (4) Asian
    (5) American Indian/ Native American
    (6) Native Hawaiian or other Pacific Islander
    (7) Mixed
    (8) Other

4. Do you know how to fluently speak any other language(s) in addition to English?

    (1) Yes
    (2) No

5. If Yes, please list them?

    ____

APPENDIX I: PERCEIVED VOICE IMPRESSIONS

Perceived Voice Impressions

Instructions: After having listened to the recording, indicate to what degree you agree/disagree with each statement provided.

Strongly Disagree (1); Disagree (2); Somewhat Disagree (3); Neutral (4); Somewhat Agree (5); Agree (6); Strongly Agree (7)

Standard American Accent

1. **This voice sounds like typical American English to me.**
2. This voice sounds like general American English.
3. This is the standard way of speaking American English.

Foreign Accent

1. This voice sounds foreign to me.
2. **This voice seems to have a foreign accent.**
3. This speaker seems to come from a foreign country.

Computer-Synthesized Accent

1. This voice sounds mechanical.
2. **This voice sounds robotic to me.**
3. This sounds like a machine voice.

For the informal pre-testing, these items were shown for each of the respective accent categories and the accents with the highest reliability were used in the actual study.

The bolded items were used in the actual study as a manipulation check to make sure what was supposed to be measured was actually measured.