"Did They Just Say That?!": Investigating the Comprehension of Racialized Adjectives in

Sentential Contexts

Isaias Ceballos III

Advisor: Dr. Fernanda Ferreira

Psychology Department, University of California at Davis

Abstract

Have you ever felt uncomfortable hearing President Obama referred to as "articulate"? This is an example of racially coded language- which can be subtle. This project investigates how racialized adjectives affect language processing speed. We administered a survey in which we ask subjects to list adjectives that they find offensive. Following the survey, we will conduct two rating studies. The first study will ask participants to rate on a Likert scale how offensive they believe an adjective to be, and the second rating study will ask participants to rate adjectives based on how related to race/ethnicity they interpret the adjectives to be. We will use the responses that receive ratings in the middle of the scale to examine how covert linguistic microaggressions affect processing speed in a self-paced reading study- in which participants read sentences word-by-word. The reading study is designed to compare sentences with and without racialized adjectives and where the referent does or does not match the intended covert microaggression (e.g., articulate Barack Obama vs articulate Bill Clinton). We predict that sentences containing a match between the intended referent and the racialized adjective will be processed faster because there is a match of stereotyped expectations that leads to semantic agreement. This work will allow us to identify the effects of racialized language on real-time language processing.

Keywords: racialized, covert, microaggression, self-paced reading, referent

Introduction

As we learn more about the mechanisms of racism, we realize that racism exists even in very covert ways and as we gain more knowledge about the relationship between racism and language, we should question how we, as comprehenders, process racialized language, and how

the use of such language reflects implicit and explicit biases. Racialized language is language that is used to describe members of one racial group more so than another. Unlike racial slurs or other overtly racist language, the racialized meaning can be quite subtle. Previous work suggests that individuals profile other individuals based on their racialized dialect (Massey and Lundy, 2001) and that racialized language likely influences cognitive processes (Wright, 2017). The current study expands on previous work by investigating how specific, covert forms of racialized language influence language processing, potentially providing a sensitive measure of racist attitudes.

Linguistic Profiling and Discrimination

To understand how comprehenders process racialized language, we need to examine how language can provide information about an individual that could be the basis for how they are discriminated against via lexical and grammatical choices. Some studies have shown that discrimination can occur through linguistic profiling (Massey and Lundy, 2001). For instance, Massey and Lundy (2001) examined subtle forms of dialect-based discrimination by conducting a phone audit study in which auditors posed as renters in search of housing with comparable scripts to use when calling about the availability of an apartment or house. The authors found that Black individuals generally experience less access than whites to housing (Massey and Lundy, 2001). This finding demonstrates that listeners generate racialized judgments about speakers with very little information. The only information provided to listeners was the auditors' dialect and voice. This is important towards answering how comprehenders process racialized language because the discrimination in which the rental agency engaged was based on applying stereotyped information about individuals in relation to the dialect that they speak, such that those who spoke Black Accented English (BAE) tended to face the most discrimination

since the BAE dialect is stereotyped to be spoken by low-income black individuals and therefore stigmatized.

Children's Occupational Stereotypes in Language

Furthermore, stereotypes do not just exist for the listener. Stereotypes can manifest in the voice of the speaker, as well. Fortunately, there are studies that examine stereotypicality within language. Stereotypicality has been examined within the vocal patterns of language itself, which provides evidence that there may be internal components of stereotypicality in language (Cartei et al. 2020). In a cross-sectional study, Cartei et al. (2020) examined occupational gender stereotypes in children through a phonetic analysis of their responses in an elicited imitation task. The authors directed children to imitate people who work in specific occupations, such as construction worker or a nurse, while their voices were recorded. Cartei et al. (2020) found that older children's responses showed larger tone differences between gendered occupations (e.g. nurse versus doctor) than younger groups of children. They also found that the level of stereotypicality, or the degree to which the tone demonstrated stereotypes, was significantly related to children's explicit beliefs about gender and occupations. This study demonstrates phonetic evidence for occupational stereotypes and gives clues to how stereotypes are learned in childhood. This study is very relevant to answering how comprehenders process racialized language because it demonstrates that stereotypes can manifest in the way that we talk as early as childhood and that there are internal components in language for stereotypicality.

Microaggressions

Further evidence that stereotypes influence how speakers communicate comes from everyday microaggressions. Microaggressions in everyday language show that stereotypes can influence what speakers choose to say. This research is important because it creates ground-level

frameworks by creating universal, working definitions for a concept that can have many meanings according to different people. Sue et al. (2007) demonstrate that microaggressions exist in subtle and varied ways, which this project aims to focus on. Sue et al. (2007) addresses what microaggressions are and categorizes different forms of microaggressions. The authors conducted a literature review of aversive racism, everyday racism, and narratives from counselors that ranged in ethnic background. They note it is essential to define what a microaggression is because there are inconsistencies in the definitions that are applied. In their literature review, they beforehand defined racial microaggressions as brief, everyday verbal, behavioral, and environmental indignities that convey hostile or negative attitudes to a person or group--regardless of whether the intent is present. They also categorize subforms of microaggressions which include microassaults, microinsults, and microinvalidations. They define microassaults as explicit racial derogations that are violently verbal or nonverbal and meant to harm the intended individual through purposeful discriminatory actions. The authors define microinsults as verbal remarks or comments that convey rudeness and insensitivity to a person's racial heritage. Microinvalidations are defined as verbal comments or behaviors that exclude or negate the thoughts, feelings, and experiences of people of color.

Influence of Race on Language Comprehension

Furthermore, there is research on how race can influence word recognition. Specifically, there is evidence that there are learned associations between race and language for bilingual versus monolingual infants (Singh et al., 2020). Singh et al. (2020) investigated the influence of sensitivity to race on visually-mediated word recognition among monolingual and bilingual infants via the preferential looking paradigm. In this work, they presented two-year-old monolingual and bilingual infants with objects and had speakers of their same race or a different

race pronounce the object label to see if word recognition varied across the two groups of infants based on speaker race and whether the word was pronounced correctly or incorrectly. Singh et al. (2020) found that both monolingual and bilingual infants associated correctly pronounced words with visual targets when the speaker of the word was of the same race. However, when the speaker was of another race, only bilingual infants associated correctly pronounced words with visual targets, while monolingual infants did not fixate on the target regardless of whether the word was pronounced correctly or mispronounced (Singh et al. 2020). There are multiple possible explanations for these results. Regardless of which explanation is correct, this finding shows that when we process language, we take into account the race of the person who is speaking as early as at 2 years of age, which suggests that we use social information—such as the race of the speaker and our stereotypes about people who are racialized in the same way—to process racialized language.

Racialized Language in Society

The study of racialized language is relatively new, and there are few studies that examine how racialized language is processed. Wright (2017), for instance, gathered person-directed adjectives in order to investigate how racialized language occurs in everyday language, and how it is processed by listeners. This study gathered person-directed adjectives through web scraping of written sports journals describing Serena Williams, from which racialized adjectives were then selected. Using the visual world paradigm, participants were presented with a racialized adjective followed by two faces—a black face and a white face—and subjects were instructed to fixate on one of the images to respond. The fixation latencies to the images were then recorded and analyzed. The results demonstrated that subjects tended to make a saccade to the image of a black face faster when there was a racialized adjective that matched stereotypes of black

individuals. Unfortunately, the differences were only marginally significant. Wright (2017) was a novel contribution to the study of raciolinguistics since it provided an original methodology for studying the effects of racialized language on language processing that does not rely on self-report or other conscious measures. The current study applies a similar approach to that of Wright (2017) to the study of sentence processing.

Sentence Processing

This project is inspired by Wright (2017) but is conceptually different in the methodologies used to examine the relationship between language processing and racialized language. This project uses a self-paced reading paradigm to compare the processing speeds of sentences containing racialized adjectives. This paradigm offers other benefits that an eye-tracking experiment cannot provide, as suggested by evidence reported by Ferreira and Henderson (1990).

In Ferreira and Henderson (1990), the authors investigated whether verb information aids in the parsing of temporarily ambiguous sentences through self-paced reading experiments and an eye-tracking experiment. They hypothesized that verb information would aid in the parsing of ambiguous sentences. The results of the eye-tracking experiment and the non-cumulative self-paced reading experiment showed similar findings that supported their hypothesis. Ferreira and Henderson (1990) discuss the benefits of eye-tracking experiments and non-cumulative self-paced reading studies. Specifically, they note that eye-tracking experiments reflect a natural reading situation, while self-paced reading experiments do not. Non-cumulative self-paced reading, however, required participants to depend more on verbal information- which was more profound than in the eye-tracking experiment. Though this study addressed different questions than we hope to answer in this project, it highlights the rationale behind the use of self-paced

reading experiments in the methodology of psycholinguistics. Self-paced reading experiments that are non-cumulative -which are used in this project- are beneficial for the field of psycholinguistics and for this project since they tend to demonstrate stronger effects of particular lexical categories in sentence comprehension than in eye-tracking experiments. In the case of this project, the self-paced reading experiment will demonstrate if the presence of a racialized adjective affects comprehension by comparing the reading times which may show stronger effects than an eye-tracking experiment.

The Current Study

Hence, taken together, the studies reviewed above suggest that comprehenders use stereotypes and stereotyped schemas to process racialized language. The previous studies, however, do leave a knowledge gap as to how exactly comprehenders process racialized language, and further research that uses racialized adjectives in sentences is necessary to truly come to a conclusion. Our project aims to fill in the gaps that remain from previous research. This project is conceptually inspired by Wright (2017) in that we will examine how we, as comprehenders, process racialized adjectives in the context of sentences. However, this project does help fill the gap of knowledge of racialized language comprehension and leads us to answer whether or not stereotypes and stereotyped schemas assist in language processing. This project is innovative in that it collects the racialized adjectives (and control words) directly from subjects. It is also unique in that naive subjects are the judges of whether an adjective is a covert, racialized adjective rather than relying on the decisions of "expert" judges. Rather than focusing on the processing of individual words, we focus on the processing of racialized adjectives in sentences since we typically process language in sentential contexts. The self-paced reading

methodology is more sensitive to the contexts of the racialized adjectives and masks the hypothesis of the project which is why this paradigm was selected over the eye-tracking paradigm proposed by Wright (2017). We hypothesize that sentences containing racialized adjectives and a matched referent will be processed faster (shorter reading times) than sentences of any other condition since there is an agreement of stereotyped knowledge between the referent and the racialized adjective.

Methods

Before the self-paced reading experiment, we conducted three norming studies to gather racialized adjectives and confirm empirically that those adjectives are racialized, but not explicitly racist. : first, a free-response study to gather a corpus of adjectives, second, a rating study to filter for covert adjectives, and third, another rating study to confirm adjectives were racialized.

Demographics

In each of the studies, we collected demographic information from each subject. We asked participants questions regarding their race, ethnicity, gender, sexual orientation, religion, and age. The purpose of gathering demographic information of each subject was so that we could examine the relationship between the adjectives that a subject would list and their identity. The questions were adapted from the California Health Interview survey (CHIS, 2016) that aimed to gather demographic information, but, the questions were designed so that if a specific identity was not included as an option the subject could type their answer in a free-response text box.

Adjective Collection Survey Study

Subjects

Adjectives were provided by 102 undergraduate students at the University of California, Davis

who were recruited through the SONA participation pool and participated in exchange for course

credit. Subjects were at least 18 years old and spoke English as a primary language. In the survey

study, 31 subjects were excluded. Subjects were excluded if they provided responses that could

not be construed as adjectives (e.g., single letters, random key combinations, blank responses) or

if they indicated that they did not answer the survey carefully. The latter criterion was introduced

because this and subsequent studies were conducted online during the pandemic, which makes it

difficult to ensure that subjects take the task seriously (further details are given below). A total of

71 subjects' data were included.

Design

The surveys were divided into two blocks: an experimental block that asked the

participants to list adjectives based on three different prompts and another block that asked

participants to answer questions about their demographic background. Subjects either received a

version of the survey where demographic questions were asked before the experimental block or

after. This was meant to counterbalance the possible influence of the demographic questions on

the target blocks. Moreover, the first prompt asked, "What are some adjectives that people use

to describe you but may bother you and people who are demographically similar to you?" The

aim of this prompt was to gather responses that were explicitly offensive so that it would be

easier to distinguish between overt and covert descriptors. The second prompt asked, " Are there

any adjectives that are used to describe you that make you feel good?" This prompt served as a

distractor from the obviously offensive adjectives that could trigger traumatic experiences for the

subjects. Finally, and critically, the third prompt asked "Are there any adjectives that are less

obviously linked to your identity but may still bother you?" This prompt served to gather target

adjectives- adjectives that were less offensive relative to the first prompt. This survey was

conducted using Qualtrics XM.

Procedure:

Subjects began the survey by consenting to participate in the research with the knowledge

that the survey could trigger traumatic experiences of microaggressions or other interpersonal

harms. Following completion of the consent form, subjects who received survey version 1

completed demographic questions prior to listing adjectives, and subjects who received version 2

completed the survey in the opposite order. At the end of the study, subjects were asked a

question that served as reassurance for our research team that the survey was taken seriously and

without distraction by subjects. Following the completion of the experimental portion of the

survey, participants were told "...if for whatever reason you feel that you did not complete the

survey carefully or accurately, it would be extremely helpful if you could let us know this now.

Your response will in no way affect your compensation". Then they were asked to click on one

of two choices: the first being, "I did complete this survey carefully. Please include my responses

in your analysis" or the second choice "I did not complete this survey carefully. Please exclude

my responses in your analysis". Though it would be quite simple to be dishonest, we made it

transparent that regardless of their response they would receive credit.

Offensiveness Ratings

The offensiveness rating study serves as a filter for the corpus of adjectives obtained in the study

just described so that only adjectives that cannot be construed as slurs are included in the self-

paced reading experiment.

Subjects

Participants were 283 UC Davis students who were older than 18 years old, spoke English as their primary language, and had not participated in the adjective collection survey. Data from 33 of the participants were excluded due to either incompleteness of the tasks or failure to complete the survey tasks carefully, as was determined by catch trials and whether subjects affirmed that they were careful in answering the ratings (see below).

Design

The offensive rating studies consisted of 10 different versions of the list of adjectives that were collected from the previous survey study. The list of adjectives was divided evenly among five studies, however, to accommodate the possible effects of answering demographic questions, we made two versions for each of the 5 studies. The only difference between the two versions was that one version presented demographic questions before the rating scales, while the other presented the questions afterward. This survey was conducted using Qualtrics XM.

Stimuli

A corpus of 361 target adjectives and 100 filler adjectives derived from the first study were rated. Target adjectives came from the third prompt of the survey study that asked participants "Are there any adjectives that are less obviously linked to your identity but may still bother you?" Filler adjectives came from the second prompt that asked participants " Are there any adjectives that are used to describe you that make you feel good?" Adjectives that were included were subject to correction of spelling errors and lowercase errors. Spelling was adjusted to the

standards of the 3rd edition of the Oxford dictionary, so hyphenated descriptors, for example, were accommodated.

Design

The adjectives were rated on a Likert scale ranging from "not at all offensive" to "slightly offensive" to "somewhat offensive" to "offensive" to "extremely offensive". The filler adjectives were adjectives collected in the previous study that subjects said made them feel good. The 361 target adjectives were divided into 5 adjective lists of no more than 80 target adjectives. Each list additionally included 20 filler adjectives. The trial presentation was pseudorandomized such that two fillers couldn't be presented sequentially, and only 20 could be presented per 80 target adjectives. This rating study also contained a reassurance block asking subjects to truthfully state whether they paid close attention to the study and its instructions. This survey was conducted using Qualtrics XM.

Procedure

Subjects were instructed first to either answer questions regarding their demographic background or respond to prompts asking them to rate the displayed adjective on a 5 point scale depending on which version of the study they received. The participant answered the demographic questions by specifying their race, gender, sexual orientation, religious affiliation, and age. Participants answering the experimental block read the prompt " Please indicate how offensive this term is to you." and then participants were presented with a bolded word (e.g., articulate). Participants would have to rate how offensive they believe the adjective to be ranging from "not

at all offensive" to "extremely offensive". The reverse order of this would occur if the participant received the other version of this study.

Racialization Ratings

A second rating study will be conducted to determine the degree to which adjectives rated as less offensive in the previous study were associated with race (e.g., racialized). The methods for the proposed racialized rating study are identical to that of the offensiveness rating study, with the following exceptions. As of note of transparency, this study has yet to be completed.

Subjects

Subjects will be recruited from the SONA participation tool. The pool of participants consists of undergraduate students seeking to gain course credit through participation in research. Participants that are included will be older than 18 years of age and will have English as their primary language. These subjects will be sampled from a separate pool that did not take any of the other studies in this project.

Stimuli

Adjectives that were rated in the middle of the continuum- "somewhat offensive" -in the offensiveness survey will carry over to the racialized rating study. This will filter out adjectives that were rated so high that they could be perceived as slurs and adjectives that were rated so low that they likely wouldn't be perceived as referring to others' racial identities. These remaining adjectives would be included in the rating study in which there would be a Likert scale from "not at all related to race", to "slightly related to race", to "somewhat related to race", to "related to race", and then "extremely related to race". The demographic questions from the previous studies would also be presented.

Design

The prompts in this rating study differ from the prompts in the offensiveness rating study. In the racialized rating study, participants are asked how related to race they believe the presented adjective to be. The list will consist of a much smaller list of adjectives from the previous offensiveness rating study. The surveys will contain an even distribution of 100 rating questions each. In these surveys, 20 percent of these questions will contain filler adjectives. This survey will be conducted using Qualtrics XM.

Procedure

In the racialized rating study subjects will be instructed first to either answer questions regarding their demographic background or then respond to prompts asking them to rate the displayed adjective on a five-point scale depending on which version of the study they received. The participant answered the demographic questions by specifying their race, gender, sexual orientation, religious affiliation, and age. Participants will be instructed to rate adjectives based on how related to racial identity they believe the adjective to be. Specifically, they would be asked, "Please rate the adjectives based on how related to race you believe them to be" while the adjective is presented in bold font. The randomization of adjectives will have the same constraints and the same proportion of filler adjectives to target adjectives as the previous rating study. This rating study will also contain a reassurance question asking subjects if they answered the ratings carefully.

Self-Paced Reading

The ultimate goal of the project is to understand how we process racialized adjectives in the context of sentences. To do so, we will use a self-paced reading paradigm using PennController IBEX Farm (Zehr & Schwarz, 2018) to measure the processing speed of word-by-word reading for sentences containing covert, racialized adjectives.

Subjects

In the proposed self-paced reading experiment, there will be a large sample of 250 participants. The conditions of the various sentences will be randomly assigned to participants. Subjects will be recruited from the SONA participation tool. Participants will be undergraduate students seeking course credit through research participation. The participants will be at least 18 years old and use English as their primary language.

Stimuli

In the self-paced reading study, adjectives that- on average- were rated as " somewhat related to race", "related to race", and "extremely related to race" will be considered racialized adjectives. These racialized adjectives will be placed in sentences that also have a matching referent- or the person that the adjective is referring to. Whether or not the adjective is a match to the referent is determined by whether the presence of both the adjective and referent elicits racist ideas (e.g., a sentence containing both the adjective "articulate" and the referent "Barack Obama"). Stimuli will also include non-target sentences containing non-racialized adjectives. These non-racialized adjectives will be determined by adjectives rated as "not at all related to race" and "slightly related to race" in the racialization rating study. Meanwhile, the referents will come entirely based on celebrity figures that are listed by research assistants. The non-matched referent will be determined by a public figure that does not implicate racist ideas when

the racialized adjective is present. The stimuli will have 4 different conditions that are specified below.

Design

The design of the self-paced reading study is a Latin-square 2 x 2 design. The factors in this design are whether or not a racialized adjective is present in a sentence and whether or not there is a matched referent for that adjective. As a result of this design, there will be four conditions in the self-paced reading study. Sentences in the first condition will contain both a racialized adjective and a matched referent in relationship to how the adjective is used. Sentences of the second condition will contain a non-racialized adjective and will contain the same referent as the first condition. In the third condition, sentences will consist of a racialized adjective and a non-matched referent, while the fourth condition will contain sentences that have a non-racialized adjective and a non-matched referent. The conditions in which there is a racialized adjective and a matched referent are the target sentences that we aim to examine. All other conditions serve as a means of comparison. The number of non-target conditions will make up 20 percent of the presented stimuli. The measured variable of this study will be the reading times of these sentences. This self-paced reading experiment will be conducted using Penn-controller IBEX farm. We will also collect demographic background information from participants using the same questions asked in the previous studies. This will be done so that we can also identify a possible relationship between individual subjects' reading times and their demographic background. This survey of demographics will be created using Qualtrics XM.

Procedure

In the proposed self-paced reading experiment, participants will read sentences one word at a time by pressing the spacebar to reveal the next word. Each letter in the word will have an underline underneath it and the spacing between words will remain as it would in a typical sentence. The participants will continue until they read the randomly assigned list of sentences. The reading times of each of these conditions will be compared to examine if comprehenders process sentences faster when there is a match between a racialized adjective and a matched referent than in other conditions. We will also collect demographic background information from participants using the same questions asked in the previous studies. This will be done so that we can also identify a possible relationship between individual subjects' reading times and their demographic background. The order between the demographic questions and the self-paced reading experiment will depend on which version of the experiment the participant receives to counterbalance a possible influence of asking demographic questions on the experiment at hand.

Results

We were able to conduct and collect data from both the survey study and the offensiveness rating study. However, we have yet to collect data from both the racialization rating study and the self-paced reading study.

Survey Study

Overall, in the survey study that collected the initial corpus of adjectives, there were interesting trends in the adjectives that subjects chose to list. There were a total of 361 adjectives that were among the two versions of the survey study. The adjectives had not been rated on offensiveness yet; however, a majority of descriptive adjectives were slurs or very extreme terms that most people would interpret as being offensive. This was the case for both prompts that

would ask subjects to list adjectives that could be construed as offensive. For the sake of avoiding any triggering language, I will avoid giving examples of slurs we found in our study in this article. One word to note from both of these studies was that no particular appeared most frequently among the two studies, regardless of the order in which the demographic questions were asked. The included responses consisted of unique adjectives for each version.

Offensiveness Rating Study

In the rating study that instructed subjects to rate adjectives based on how offensive they believed the adjective to be, we anticipated that a majority of the adjectives would be polarized so that they were either rated as "extremely offensive" or "not at all offensive". On the Likert scale from "not at all offensive" to "slightly offensive" to "somewhat offensive" to "offensive" and lastly "extremely offensive", we expect a majority of the responses to not be distributed in the middle of the scale. In order to account for this, we would calculate the average responses for each adjective and keep the averaged responses that were either "slightly offensive", "somewhat offensive" or "offensive". The mean number of responses for each version of the offensiveness rating study was 112 since all included participants were required to rate all of the listed adjectives in each version of the rating study. Even though the inclusion criteria for the Likert scale is large, the expectation of average responses being polarized will greatly reduce the corpus of adjectives that remain to be tested in the racialization study. This is due to their being a large presence of terms that can be construed as slurs in the survey study. This polarization would lead to a substantially smaller corpus of adjectives that would need to be filtered in the subsequent rating study.

Relation to Race Rating Study

In the rating study where subjects rate how related to race they believe the adjectives to be, we predict that a majority of the remaining list of adjectives will be rated as very related to race. This is because most of the adjectives were adjectives that can be interpreted to be descriptors of race. However, there are adjectives that may not be racialized but may have intersectional characteristics. For instance, we predict that words like "feisty" will not only be construed as a micro-aggression for women but also for women of color. We anticipate that subjects will still rate these types of descriptions as high in relation to race because there are still allusions to this identity We expect that most ratings will be rated as " somewhat related to race", "related to race", or "extremely related to race" due to the specific prompts in the offensiveness rating study.

Self-Paced Reading Experiment

In the self-paced reading experiment, subjects will read sentences and their reading times will be measured. We anticipate that the results would support the hypothesis such that the condition in which sentences contain a racialized adjective and matched referent in relation to how the adjective is used will be processed the fastest. This is due to their being a stronger adherence to stereotypes in this condition compared to other conditions. However, we do respect that the demographic backgrounds of subjects' can influence the results such that certain subjects may not adhere to stereotypes of other ethnic groups compared to subjects. For instance, subjects of white backgrounds may adhere stronger to stereotypes of other ethnic groups because they do not belong to those ethnic groups. We would conduct a regression analysis for this study in which we also examine the average time taken on the racialized adjective versus the non-racialized adjective.

Discussion

These studies, specifically, answer how we process racialized language by implicating that comprehenders process language that is racialized by using stereotypes and stereotyped schemas. By answering how we, as comprehenders, process racialized language we can address the issue of racism in our society by becoming more aware of the racism that exists in the language we speak.

The survey study primarily served to collect adjectives for the subsequent rating studies, however, the findings in the survey study tell also an interesting story. The current analysis of the responses tells us that there was no underlying descriptor that participants thought was offensive since the adjectives listed were unique responses across each version of the study. Another analysis of this data, in which we investigate the relationship between the listed adjectives and the respective demographic background of the respondents could give insight into which types of adjectives participants list. For instance, there could be a relationship between the racialized adjectives that were listed and the identity of the participant that listed them. We can assume this is the case intuitively, but it would be a meaningful analysis of the data to prove this assumption.

The offensiveness rating study's findings primarily served for filtration of these adjectives for the following studies to be performed, however, it also has telling information. We analyzed the mean number of responses for the sake of ensuring there was no large variation among the different versions of the offensiveness rating study. These findings in sum do not address our hypothesis, but an analysis between the average ratings for each adjective that was tested and the demographic background of the participants could provide insight into how individuals of certain ethnic groups interpret the offensiveness of adjectives.

We were unable to collect data from both the racialization rating study and the self-paced reading study, but we do have expected findings from these studies. In the racialization rating study, we do expect that responses on average will be on the right end of the Likert scale-"somewhat related to race" to "extremely related to race". In the self-paced reading experiment, we expect that sentences containing the racialized adjective and the matched referent will be read faster than any other sentences. An analysis of the average time spent on the racialized and non-racialized adjectives would be an impactful use of the data that was collected, as well, since it could provide insight into how comprehenders parse sentences with racialized adjectives.

We anticipate that the results will be influenced by the demographic background of the subjects for the rating studies and the self-paced reading study. Both of the rating studies are expected to be influenced by the demographic background of subjects since adhering to stereotypes is heavily influenced by the ethnic group that the individual belongs to. The self-paced reading study is also expected to be influenced by the demographic background of the subjects for the same reason.

The recruitment pool were all undergraduates and at first glance, this does pose a limitation on the external validity of the study since all participants were undergraduate college students. However, it can be argued that this is a general issue with SONA recruitment and is not necessarily a limitation in the study. If the results support the hypothesis under the limitations of the participant pool, it only demonstrates the robustness of the study since the results would hold even more significance under a larger pool of participants that are representative of the nation's racial demographic background.

Though this is a study examining the racialized adjectives, we do acknowledge the intersectionality of the adjectives that were included in the original survey study. Some of the

descriptors that were listed by participants involved gendered slurs and micro-aggressions. Further research would be necessary to demonstrate conclusive results and would be needed to investigate the micro-aggressions that spanned other dimensions of the subjects' identities such as gender, sexual orientation, and religion.

Conclusion

From these studies, we know that discrimination on the basis of language has become a socially palatable substitute for discrimination against speakers of the language, such as experiences of more discrimination towards speakers of certain dialects compared to others. As equally important, these studies demonstrate that otherwise innocuous language can become racialized to take on covertly racist meaning.

The literature has yet to address the specific components of comprehension of racialized language and has only touched on a few dimensions of racism. This project focuses on the covert rather than the many forms of racism that exist in our society. In this respect, this project could aid in filling in the aforementioned gaps of information. In the future, research could be done to examine if prosodic features influence the comprehension of racialized adjectives. Specifically, to examine if the tone has an influence on comprehension via ERP studies on auditory stimuli containing racialized adjectives.

References

California Health Interview Survey. CHIS 2016 Adult Public Use Files. [Computer file]. UCLA Center for Health Policy Research, Los Angeles, CA. January 2019.

Cartei, V., Oakhill, J., Garnham, A., Banerjee, R., & Reby, D. (2020). "This Is What a Mechanic Sounds Like": Children's Vocal Control Reveals Implicit Occupational Stereotypes. Psychological science, 31(8), 957-967.

Henderson, J. M., & Ferreira, F. (1990). Effects of foveal processing difficulty on the perceptual span in reading: implications for attention and eye movement control. Journal of Experimental Psychology: Learning, Memory, and Cognition, 16(3), 417.

Massey, D. S., & Lundy, G. (2001). Use of Black English and racial discrimination in urban housing markets: New methods and findings. Urban affairs review, 36(4), 452-469.

Mercer, S. H., Zeigler-Hill, V., Wallace, M., & Hayes, D. M. (2011). Development and initial validation of the Inventory of Microaggressions Against Black Individuals. Journal of counseling psychology, 58(4), 457.

Singh, L., Tan, A. R., Lee, K., & Quinn, P. C. (2020). Sensitivity to race in language comprehension in monolingual and bilingual infants. Journal of Experimental Child Psychology, 199, 104933.

Sue, D. W., Capodilupo, C. M., Torino, G. C., Bucceri, J. M., Holder, A., Nadal, K. L., & Esquilin, M. (2007). Racial microaggressions in everyday life: implications for clinical practice. American psychologist, 62(4), 271.

Wright, K. E. (2017). The Reflection and Reification of Racialized Language in Popular Media.

Zehr, J., & Schwarz, F. (2018). PennController for Internet Based Experiments (IBEX). https://doi.org/10.17605/OSF.IO/MD832