

Nicole Holliday* and Marie Tano

“It’s a Whole Vibe”: testing evaluations of grammatical and ungrammatical AAE on Twitter

<https://doi.org/10.1515/lingvan-2020-0095>

Received June 12, 2020; accepted January 18, 2021; published online May 11, 2021

Abstract: This study focuses on listener perceptions of African American English (AAE) on Twitter, examining both grammatical and ungrammatical usages, as well as how these perceptions may be affected by the race of the speaker and rater. We conducted an experimental survey designed to address the following questions: 1. Does avatar race affect perception of the grammaticality of AAE? 2. Are differences between grammatical and ungrammatical AAE discernible to naive raters of different races? 3. How are various social attributes evaluated for avatars of different races and different linguistic varieties? Results indicate that participants generally do not downgrade avatars who use grammatical AAE on ratings of grammaticality or personal characteristics. However, participants of all races disprefer ungrammatical uses of AAE, with black raters being especially sensitive to ungrammatical AAE. These findings have implications for sociolinguistics in that they demonstrate that participants across racial backgrounds may differentiate grammatical versus ungrammatical AAE online, and that contrary to expectations based on previous literature, AAE is not universally downgraded in these contexts. However, results also indicate that the use of AAE still negatively impacts listeners’ perceptions of speakers as educated, demonstrating that some widespread biases against AAE-speakers persist in an online context.

Keywords: African American English; language ideologies; online language; perception; sociolinguistics

1 Introduction: African American English online

African American English (AAE) has been well-described in the sociolinguistic literature, with a number of formative studies focusing specifically on the ways in which perceptions of the variety influence the experiences of black speakers. In recent years, with the widespread expansion of online communication, the diffusion and status of AAE online may have begun to shift. In his formative 2015 work, “Online Imagined Black English”, Manuel Arturo Abreu articulates a phenomenon in which AAE has become part of the fabric of online communication itself. This phenomenon however, is complicated by AAE’s deep indexical links to black culture, as well the negative stigma that also frequently gets attached to black speakers, regardless of language variety (Baugh 2015). While AAE is still clearly stigmatized in the great majority of professional and institutional contexts (Baugh 2015; Rahman 2008; Rickford and King 2016; *inter alia*), it has long been known to carry some level of covert prestige that speakers can exploit to index authenticity, toughness, or coolness (Cutler 1999; Labov 1972; *inter alia*). As Abreu notes: “in institutional contexts, African American English faces heavy stigma, while in other contexts, this warmth, authenticity, and seductive danger may be communicative and pragmatic boons. In this sense, the same stereotyped indexical meanings that white supremacist culture imposes on Black English endow it with covert prestige in non-institutional contexts”. Abreu later argues that it

*Corresponding author: Nicole Holliday, Linguistics, University of Pennsylvania, Philadelphia, PA, USA, E-mail: nholl@sas.upenn.edu. <https://orcid.org/0000-0001-8921-2509>

Marie Tano, Linguistics and Cognitive Science, Pomona College, Claremont, CA, USA. <https://orcid.org/0000-0002-4470-5702>

is precisely AAE's ability to index such positive yet counter-mainstream social characteristics, as well as the variety's persistent ideological links to youth and urbanity that make it an appealing tool for online communication, especially in fora such as Facebook and Twitter. In particular, Twitter has proven to be a useful platform for examining both racialized language use and ideologies online, given its higher engagement among young people, African Americans, and urban residents (Grieve et al. 2018).

Though recent work has begun to examine the role of racialized language online, most studies have not yet focused on how participants evaluate the language of racialized personae online. The current study builds on the body of work that focuses on how listeners evaluate ethnolinguistic variation in an online context, referencing the findings of earlier works such as Kang and Rubin (2009) and D'Onofrio (2019). In particular, research by Kang and Rubin (2009) demonstrates that participants' perceptual judgments are mediated by social stereotypes and expectations related to language and race. For African American speakers, given widespread negative ideologies about AAE, interpretations of their language use may be influenced by such stereotypes. However, given AAE's aforementioned ideological links to youth and urbanity, the direction of evaluations of white and black personae who use AAE may not be straightforwardly predictable in an online context. Of particular interest to the current study are the ways in which AAE may function as a mechanism for the construction of an identity that calls on ideologies about specific types of (de)racialized personae, as well as how such personae are differentially evaluated in an online context (Bucholtz and Lopez 2011; Cutler 1999; Eberhardt et al. 2015; *inter alia*).

The current study aims to examine the ways in which avatars of different races who use three different linguistic styles: Mainstream U.S. English (MUSE), Grammatical AAE, and Ungrammatical AAE,¹ are evaluated by black and white English speakers in the U.S. Additionally, it tests how these participants perceive personality traits related to competence and warmth when they are presented with tweets from avatars of different races who employ the linguistic styles of interest. Finally, the study contextualizes these grammatical and personal evaluations within broader sociolinguistic literature on perceptions of AAE, and discusses the ways in which online expectations and norms may be shifting some long-held language ideologies while maintaining others.

2 Methods and analysis procedures

2.1 Data collection

The current study employs a survey and guise-evaluation methodology to examine how participants perceive white and black avatars who use three different linguistic styles: Grammatical AAE, Ungrammatical AAE, and Mainstream U.S. English (MUSE). Data was collected through a survey conducted via Qualtrics. Participants were recruited through email and posts on the second author's social media accounts (Instagram, Twitter) during the Fall of 2019. The data analyzed comes from 115 participants, the survey had a mean completion time of approximately 20 min per participant, and participants received no financial compensation for their participation. Following the guise-evaluation section of the study (see Section 2.2), participants were asked a series of multiple-choice demographic questions including region, race, gender, sexuality, age, occupation/student status, etc. All participants self-reported as L1 speakers of American English, and were currently living in the U.S. The sample included participants from a variety of ethnic as well as regional backgrounds. Table 1 presents racial demographics of the study's participants.

¹ This is a style sometimes referred to as Mock AAE, characterized by an imitation of features that may seem to outsiders like AAE, but that would be ungrammatical for most native speakers of the variety (Bucholtz and Lopez 2011). As it is not a rule-governed variety of English spoken by any community, but it is a recognizable style, we have categorized it as such here.

Table 1: Participants by race.

Race	
Black/African American	30% ($N = 34$)
White (not Hispanic or Latino)	26% ($N = 30$)
Hispanic or Latino	18% ($N = 21$)
Asian	17% ($N = 20$)
Other race	9% ($N = 10$)

The sample included a sizeable representation of participants who identified as Black/African American, White (Not Hispanic or Latino), Hispanic or Latino, and Asian. Though we collected this finer-grained participant race data, the only significant differences observed emerged between ratings by black and white participants; subsequently, we binned the other racial categories (Hispanic/Latino, Asian, Other Race) together in the presentation of the results. Additionally, it is important to note that 88% ($N = 101$) of the survey respondents were students at highly selective small liberal arts colleges on the West Coast, while 12% ($N = 14$) were graduate and undergraduate students at various east coast universities, and that 97% ($N = 111$) of the participants reported their age as 18–24 years. Therefore, the results are not necessarily representative of the larger population, though they do provide a snapshot of how a young, racially diverse, college-educated sample may evaluate the varieties of interest.

2.2 Materials

The survey was specifically designed to evaluate differences in perceptions of MUSE and both grammatical and ungrammatical AAE usage on Twitter (GAAE and UAAE), both in terms of grammaticality and personal characteristics associated with use of the varieties. Participants were presented with a series of three tweets at a time from the same invented user profile. While the avatars (tweeters) were invented, the texts themselves were adapted from real tweets. Only the profile’s photo has been changed to a stock photo (to control for race), and the username has been redacted in order to protect user privacy. We constructed six avatar profiles, each corresponding to a specific race/linguistic style combination. The stimuli thus consisted of two profiles that used GAAE, two that used UAAE, and two that used MUSE. In each style condition, one avatar was black and one was white.² Participants thus viewed a total of six different manufactured avatar profiles, each with three tweets of the same linguistic style attributed to them. Additionally, each tweet that came from avatars who used either UAAE or GAAE employed one of the following morphosyntactic features: invariant habitual *be*, negation using *ain’t*,³ leveling of “doesn’t” to “don’t”, and negation. All features were selected due to their descriptions as both high frequency and emblematic of the variety in previous studies (Rickford 1999; Sidnell 2002; Van Hofewegen and Wolfram 2010; Wolfram 2004). In order to control for lexical effects, the MUSE avatars also employed some unique lexical items that occur frequently on social media, but that may not carry a clear association with AAE grammar itself, such as *geeked*, *yerr*, *period*, and *vibe*.

This study’s design employs digitally altered versions of tweets from real, anonymized users that contained features of either grammatical AAE, ungrammatical AAE, or MUSE in order to provide an authentic reflection of how these varieties are used on social media. Through an iterative process of searching Twitter for the grammatical features of interest; we selected the tweets that appeared with the relevant features, then

² While three of the avatars had female user photos and three had male user photos, none of the model results indicated significant differences in judgments by gender, so this will not be discussed in further detail. Future work should investigate the effects of gender and other avatar demographic traits in such studies.

³ Though *ain’t* as a negation marker is used in many varieties of American English, evaluations of tweets containing *ain’t* did not differ from those containing other features of GAAE, for evaluations of grammaticality or personal characteristics.

anonymized them. The only exclusion criteria for tweets was potential for semantic interference that may influence judgements; care was taken to avoid content and language that may trigger strong evaluations on dimensions of interest, such as profanity (Holliday and Villarreal 2020). Stock photos were obtained by using Google’s advanced search function and looking for the relevant combinations of “[race][gender]” i.e. “white man” and “black woman”. This method has the advantage of ensuring that the photos selected were ideologically visually representative of the relevant social categories (D’Onofrio 2019; Kay et al. 2015). An example tweet from each style is shown in Figures 1a–c.



Figure 1a: Example tweet from white GAAE avatar.



Figure 1b: Example tweet from the white UAAE.



Figure 1c: Example tweet from the black MUSE avatar.

After being presented with a set of three tweets from each of the six avatars, participants were asked to rate the grammaticality of each avatar’s tweets in a binary fashion (grammatical vs. ungrammatical). Subsequently, they were asked to rate the avatars on personal traits, using a 10-point Likert scale (1 being “not at all”, 10 being “definitely”) based on a series of characteristics derived from Fiske et al. (2002) and related to perceptions of a speaker’s competence and warmth. Within these dimensions, the survey examined the participant’s perceptions of the speakers via evaluations of their intelligence (“educated”), assumed personality (“funny”, “friendly”) and sense of decorum (“rudeness”).

3 Results

We conducted a number of regression models (R Core Team 2020) and visualizations in R using the ggplot package (Wickham 2016), in order to address the primary research questions. Overall, results indicate significant differences between grammaticality ratings for the three varieties tested, with UAAE universally rated negatively, and with few differences in ratings between GAAE and MUSE. Results also indicate significant differences in the personality characteristic ratings of the avatars of different race/linguistic style statuses

affected by participant race, such that black participants are somewhat more critical of UAAE than white ones. For each of these models, we observed no significant differences between white participants and those of other, non-black races, so the latter will be omitted from the presentation of these results, though we will discuss differences between black and white participants where applicable. The results related to grammaticality will be presented first, followed by the personality trait ratings and a discussion of the larger patterns of interest.

3.1 Grammaticality

First, we turn to the question of whether there are differences between how the participants evaluate avatars who use the three linguistic styles of interest: MUSE, GAAE, and UAAE. We conducted a binomial mixed effects logistic regression model, where the outcome variable was grammaticality rating (grammatical or not) with the independent variables were Avatar Race (Black or White), Participant Race (White, Black, Other) and Linguistic Style (GAAE, UAAE, and MUSE), with crossed random effects for Avatar and Participant. Overall results indicate that the GAAE-using avatars were rated as more grammatical than both the UAAE and the MUSE-using avatars by all participants ($p < 0.01$), though the difference between GAAE and MUSE was much less pronounced than the difference between GAAE and UAAE. Figure 2 visualizes the overall results, and Table 2 shows the full model output that tested for differences in grammaticality judgment by avatar race, avatar linguistic style, and participant race.

The model shows a main effect such that the white avatars were overall rated as less grammatical than the black avatars. However, this effect was counteracted in the UAAE tweets, which show that the white UAAE

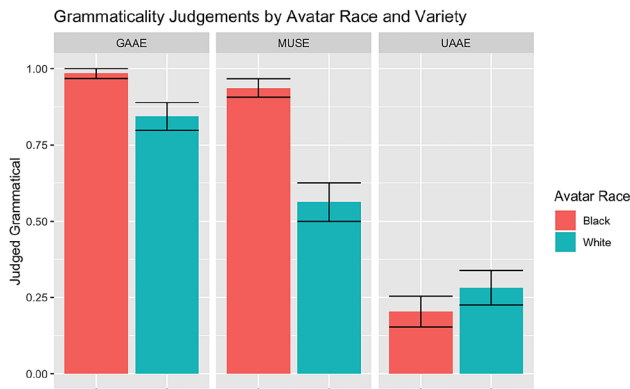


Figure 2: Grammaticality ratings by linguistic style and avatar race with error bars for standard error.

Table 2: Results of regression model testing grammaticality rating by avatar grammar, linguistic style, and participant race, plus the interaction of avatar grammar and avatar race.

	Model
(Intercept)	5.85 (1.18)***
Avatar Race = White	-2.82(1.11)***
Avatar Grammar = UAAE	-7.49 (1.25)***
Avatar Grammar = MUSE	-1.56(1.17)*
Participant Race = Black	-0.88 (0.56)
AvatarUAAE*AvatarRaceWhite	3.49 (1.24)**
AvatarMUSE*AvatarRaceWhite	-0.55 (1.28)
AIC	310.69
BIC	346.25
Num. Obs	384

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

avatars are actually downgraded less than the black UAAE avatars. Given the well-documented negative societal biases against even grammatical AAE and black speakers in general, as well as the fact that MUSE remains the dominant language of institutional structures in the U.S., these results may be unsurprising (Baugh 2015; Wiley and Lukes 1996). However, even when white avatars use a style that is largely judged as ungrammatical, the penalty they pay for this use is lower than that of the black avatars, despite the fact that the black avatars are rated more highly in general.

3.2 Personal characteristics

With respect to how avatars are evaluated on a number of personality traits related to competence and warmth (“educated”, “friendly”, “funny”, and “rude”), we observe an overall pattern in which the avatars who use GAAE are again evaluated the most favorably, and the UAAE avatars as rated least favorably, patterning with the results for grammaticality. In order to test for differences in the ratings for personal characteristics, we conducted linear regression models for each characteristic followed by ANOVAs for pairwise comparisons, to examine differences for each trait by linguistic style, as well as differences in the ratings that might be conditioned by participant or avatar race. Figure 3 shows the results for each linguistic style/personality combination, with the translucent bars representing white participants and the opaque ones representing black participants, with error bars representing the standard error. The full results of the regression model can be found in Appendix B (Figure 4).



Figure 3: Personal characteristics ratings by participant race and avatar linguistic style. Error bars indicate the standard error for each comparison.

Overall, we observe main effects for the models such that generally, participants evaluated the black avatars more favorably than the white ones, and the GAAE and MUSE avatars more favorably than UAAE avatars. This pattern further supports the results obtained for the grammaticality judgments, although there are some differences for the individual characteristics of interest, especially by participant race and avatar race, which are discussed in turn. The results presented in the visualizations come from the models for personal characteristic rating by style and participant race, while the discussion adds in results for avatar race which are not presented in the visualizations due to the nature of the 4-way interaction of the variables.

3.2.1 Personal characteristics: educated

Overall, there were no differences between the GAAE and MUSE avatars in terms of evaluations of sounding “educated”, though results indicate that UAAE avatars were rated significantly worse than MUSE ones

($R^2 = 0.26$, $F(2,1140) = 129.7$, $p < 0.001$). ANOVA results also indicate significant differences in ratings of the varieties for “educated” by participant race ($F(2,1140) = 15.2$, $p < 0.001$), such that the GAAE avatars are rated significantly more “educated” by black participants than by white ones ($F(2,1140) = 15.2$, $p < 0.001$). Figure 4 demonstrates these results.

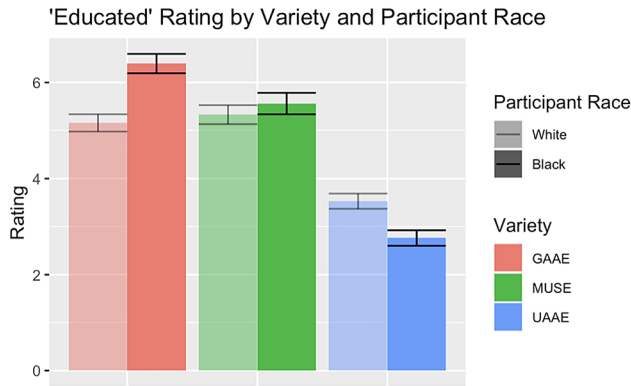


Figure 4: Ratings for trait “educated” by participant race and linguistic style. Error bars represent standard error.

This pattern is of interest because although black avatars are rated overall more favorably for all characteristics across the experiment, GAAE use does still seem to mitigate participants’ perception of them as “educated” when compared to MUSE, especially for white participants. These ratings may indicate that despite the general pattern of GAAE being judged by the participants as “grammatical”, competence-linked traits like education are still rated lower for GAAE avatars than MUSE ones, in line with expectations about the stigmatized nature of AAE (Baugh 2015; Rahman 2008; Rickford and King 2016; *inter alia*). Additionally, the white UAAE avatar was negatively evaluated by all participants, while the black UAAE avatar was rated lower by black participants than by white ones, demonstrating potential differences between black and white participants in sensitivity to ungrammatical AAE and its interaction with avatar race ($F(2, 1,140) = 3.4$, $p < 0.05$) (Figures 5 to 7).

3.2.2 Personal characteristics: friendly

Similar to the results for participant ratings for “educated”, the regression and ANOVA results for judgements on the characteristic “friendly” also differ somewhat by avatar grammar and participant race. Again, we observe no significant differences between ratings for the GAAE and MUSE avatars for this characteristic ($R^2 = 0.20$, $F(2,1140) = 0.21$, $p > 0.05$). However, main effects of the regression model indicate that all participants significantly downgraded the UAAE avatars with respect to “friendliness” ($F(2,1140) = 58.0$, $p < 0.001$). White avatars were also rated as less friendly overall than black ones ($F(2,1140) = 103.6$, $p < 0.001$). Additionally, black participants especially downgraded the black avatar who used UAAE, just as they did for the trait “educated” ($F(2,1140) = 2.8$, $p < 0.05$). Figure 5 demonstrates these results.

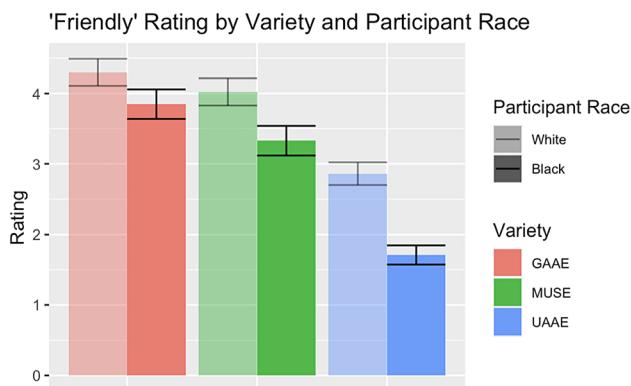


Figure 5: Ratings for trait “friendly” by participant race and linguistic style. Error bars represent standard error.

3.2.3 Personal characteristics: funny

The regression and ANOVA results also indicate significant differences in ratings of the avatars for the trait “funny” by participant race, avatar race, and linguistic style. White avatars were evaluated as less “funny” than black avatars ($R^2 = 0.12$, $F(2,1140) = 72.7$, $p < 0.001$). However, black participants actually rated the white UAAE avatars better than white participants, paralleling the same type of penalty for black UAAE avatars we observed for “educated” and “friendly” ($F(2,1140) = 3.2$, $p < 0.05$). Again, this result demonstrates a potential for greater sensitivity to grammaticality effects on the part of black participants in the perception of this trait, as well as a possible reluctance on the part of white participants to downgrade black avatars on any of these traits. Figure 6 demonstrates these results.

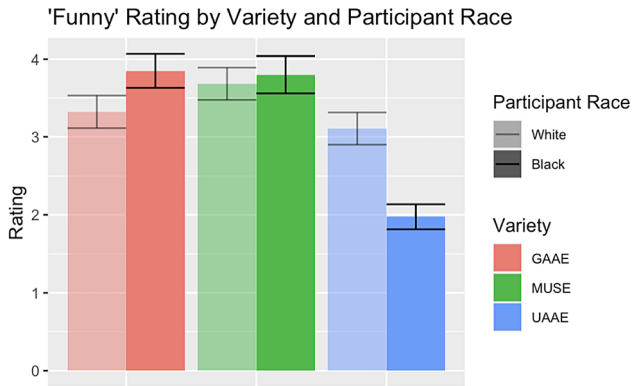


Figure 6: Ratings for trait “funny” by participant race, avatar race and linguistic style. Error bars represent standard error.

3.2.4 Personal characteristics: rude

Finally, with respect to ratings for the characteristic “rude”, the regression model does not indicate differences in evaluations by avatar or participant race. However, the ANOVA results weakly suggest potential differences between UAAE avatars and the other varieties, such that UAAE avatars are evaluated as more “rude”, indicating that linguistic style may be more predictive than race or other factors for the evaluation of “rudeness”. Figure 7 demonstrates these results.

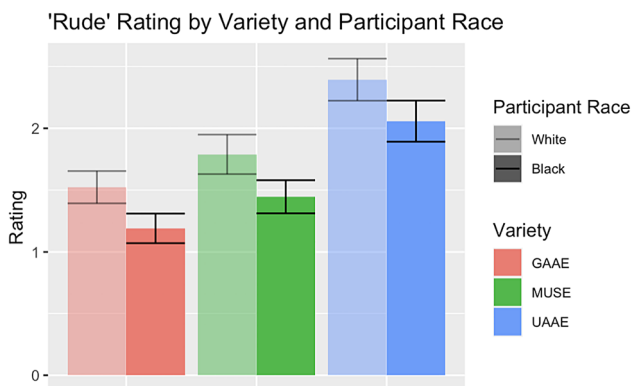


Figure 7: Ratings for trait “rude” by participant race and linguistic style. Error bars represent standard error.

4 Discussion

The results of this experiment, which was designed to test evaluations of grammaticality and personal characteristics of white and black avatars who use different racialized varieties of English, have demonstrated that

contrary to some expectations, overall, participants positively evaluated both black and white grammatical AAE speakers and negatively evaluated ungrammatical AAE speakers. However, we also observe some differences in the ratings of avatars of different races who employed the three linguistic styles of interest, especially for personal characteristics related to competence and warmth, with some variation by participant race.

These results may seem contrary to expectations, given the social stigma frequently attached to AAE, though we believe that these differences may be attributed to the unique participant population, as well as the fact that the data were presented as tweets, reflecting changing norms and expectations about online language (Squires 2010). The participant pool for this study was primarily undergraduate students from small, western colleges where many students may commonly express socially conscious and liberal views. For instance, several of the participants commented that while the examples of UAEE seemed “ridiculous,” the tweets were generally “very like what [their] twitter feed normally looks like.” Other participants reported that the study made them aware of their own biases, in that they realized how they would “subconsciously judge people’s character and intelligence level” when using AAE, and how their “perceptions of people changes based on language use.” Other participants expressed feeling “uncomfortable with the use of AAE” by non-Black people, given the historical use of AAE “as a joke” that “reinforces racist ideas about intelligence.” As a result, these findings may reflect the norms of a particular type of young, highly educated American who may have different expectations for online language than they do for other types of public language use. Our results indicate a possible normalization of AAE on social media, in the fact that both white and black participants who may have varying levels of experience with AAE evaluate grammatical uses of AAE in similar ways. However, the differences between white and black participants’ evaluations of the interaction of avatar race and linguistic style show that black participants typically downgrade ungrammatical AAE more when it is attributed to black avatars. In this way, it seems like black participants are penalizing black avatars for UAEE use, possibly due to the fact that they expect more grammatical uses of AAE from black avatars than from white ones. While there appears to be a greater acceptance of AAE online overall, the difference between grammatical and ungrammatical AAE still may not be entirely apparent to white audiences, who may be more likely to group grammatical and ungrammatical uses together, while black audiences differentiate uses and expect more grammatical AAE from black avatars than from white ones.

Also of interest is the fact that the differences in positive evaluations of grammaticality and personal traits were similar for MUSE and GAEE with respect to the traits of “educated”, “friendly”, “funny”, and “rude”. Indeed, this result indicates not only a greater acceptance of GAEE online, but also generally positive evaluations of a version of MUSE that employs lexical features vaguely associated with black speakers online. This supports the claims of researchers such as Grieve et al. (2018: 314), who have argued that AAE is the “main source of [American] lexical innovation”, and asserted that many of the words and phrases innovated by African-American individuals on Twitter influence the use of mainstream varieties online. Furthermore, even the white avatars who used GAEE in the current study were not generally downgraded on grammaticality or personal characteristics, illustrating that a certain imagined deracialized linguistic style may have also come to represent the generic, young, mainstream, American online (Abreu 2015; Squires 2010). However, this type of deracialization and subsequent integration of AAE into the mainstream does not necessarily challenge the dominance of traditional standard language ideologies. As noted by scholars such as Jane Hill, this style of ideological linguistic deracialization of AAE “add[s] value to an ‘American’ identity” that is predominantly white and degrades and marginalizes the Black community further” (Hill 2008: 160). Though AAE use may be considered more acceptable online than in other types of mainstream contexts, the linguistic style’s acceptance as limited to specific online spaces may serve to maintain a traditional linguistic hierarchy which reinforces AAE’s position as linguistically inferior and only appropriate in limited contexts (Abreu 2015).

Appendix A: Stimuli tweets

1. Selected tweets for white UAAE avatar



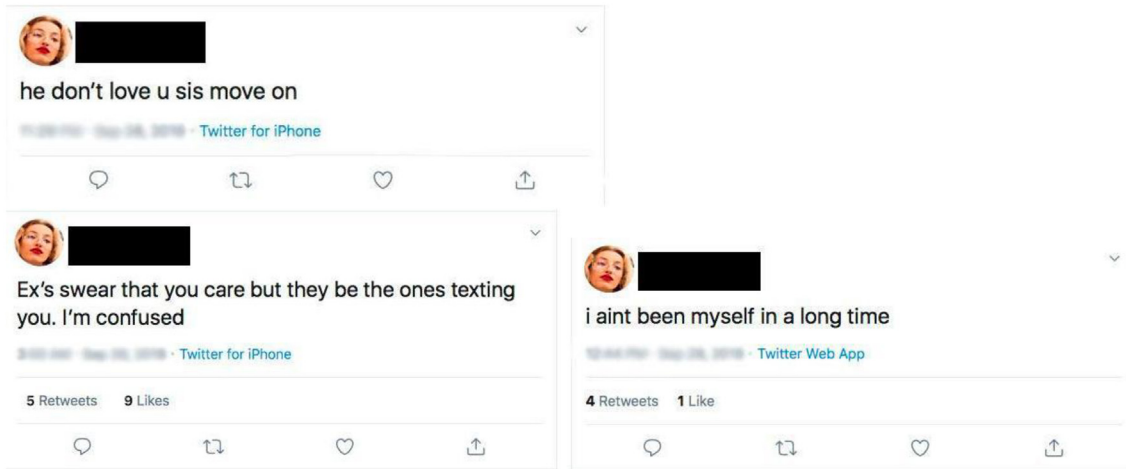
2. Selected tweets for black GAAE avatar



3. Selected tweets for black MUSE avatar



4. Selected tweets for white GAEE avatar



5. Selected tweets for white MUSE avatar



6. Selected tweets for black UAAE avatar



Appendix B: Statistical summary

1a. Personality trait ratings regression model.

	Educated	Friendly	Funny	Rude
(Intercept)	6.13 (0.27)***	4.84 (0.26)***	4.39 (0.29)***	1.52 (0.21)***
Participant Race = Black	0.55 (0.36)	-0.12 (0.35)	0.86 (0.40)*	-0.32 (0.29)
Avatar grammar = GAEE	-0.40 (0.37)	0.21 (0.36)	-0.57 (0.42)	-0.13 (0.30)
Avatar grammar = UAEE	-2.12 (0.37)***	-1.66 (0.36)***	-0.73 (0.42)	0.57 (0.30)
Avatar Race = White	-1.61 (0.37)***	-1.64 (0.36)***	-1.41 (0.42)***	0.53 (0.30)
ParticipantRace = Black*Grammar = GAEE	1.22 (0.51)*	-0.10 (0.50)	-0.14 (0.57)	-0.20 (0.42)
ParticipantRace = Black*Grammar = UAEE	-1.69 (0.51)**	-1.29 (0.50)*	-2.27 (0.57)***	0.02 (0.42)
ParticipantRace = Black*AvatarRace = White	-0.64 (0.51)	-1.15 (0.50)*	-1.48 (0.57)**	-0.05 (0.42)
Grammar = GAEE*AvatarRace = White	0.46 (0.53)	0.13 (0.52)	0.41 (0.59)	-0.27 (0.43)
Grammar = UAEE*AvatarRace = White	0.64 (0.53)	0.99 (0.52)	0.31 (0.59)	0.08 (0.43)
ParticipantRace = Black*	-0.44 (0.73)	0.69 (0.71)	1.10 (0.81)	0.42 (0.59)
Grammar = GAEE*Avatar_Race = White				
ParticipantRace = Black*	1.38 (0.73)	1.66 (0.71)*	2.04 (0.81)*	-0.03 (0.59)
Grammar = UAEE*Avatar_Race = White				
R ²	0.26	0.20	0.12	0.05
Adj. R ²	0.26	0.19	0.11	0.04
Num. obs.	1,152	1,152	1,152	1,152
RMSE	2.51	2.44	2.79	2.03

$p < 0.001$ ***, $p < 0.01$ ** , $p < 0.05$ *.

1b. Personality trait model ANOVAs.

Term	df	Educated	Friendly	Funny	Rude
Participant race	1	2.5	28.1***	1.0	7.9**
Avatar variety	2	129.7***	58.0***	22.7***	18.5***
Avatar race	1	91.6***	103.6***	72.7***	18.3***
Participant race* Avatar variety	2	15.2***	2.0	9.2***	0.0
Participant race* Avatar race	1	1.2	1.6	1.7	0.1
Avatar Variety*Avatar race	2	8.3***	15.1***	6.4***	0.1
Participant race* Avatar Variety*Avatar race	2	3.4*	2.8*	3.2*	0.4
Residuals	1,140				

$p < 0.001$ ***, $p < 0.01$ ** , $p < 0.05$ *.

References

- Abreu, Manuel Arturo. 2015. *Online imagined Black English*. London: Arachne. Available at: http://arachne.cc/issues/01/online-imagined_manuel-arturo-abreu.html.
- Baugh, John. 2015. Use and misuse of speech diagnostics for African American students. *International Multidisciplinary Research Journal* 9(4). 291–307.
- Bucholtz, Mary & Qiuana Lopez. 2011. Performing blackness, forming whiteness: Linguistic minstrelsy in Hollywood film. *Journal of Sociolinguistics* 15. 680–706.
- Cutler, Cecilia A. 1999. Yorkville Crossing: White teens, hip hop and African American English. *Journal of Sociolinguistics* 3(4). 428–442.
- D’Onofrio, Annette. 2019. Complicating categories: Personae mediate racialized expectations of non-native speech. *Journal of Sociolinguistics* 23(4). 346–366.

- Eberhardt, Maeve & Kara Freeman. 2015. 'First things first, I'm the realest': Linguistic appropriation, white privilege, and the hip-hop persona of Iggy Azalea. *Journal of Sociolinguistics* 19. 303–327.
- Fiske, Susan T., Amy J. C. Cuddy, Peter Glick & Jun Xu. 2002. A model of (often mixed) stereotype content: competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology* 82(6). 878–902.
- Grieve, Jack, Andrea Nini & Diansheng Guo. 2018. Mapping lexical innovation on American social media. *Journal of English Linguistics* 46(4). 293–319.
- Hill, Jane. 2008. *The everyday language of white racism*. Hoboken, NJ: Wiley-Blackwell.
- Holliday, Nicole & Dan Villarreal. 2020. Intonational variation and incrementality in listener judgments of ethnicity. *Laboratory Phonology* 11(1). 1–21.
- Kang, Okim & Donald L. Rubin. 2009. Reverse linguistic stereotyping: Measuring the effect of listener expectations on speech evaluation. *Journal of Language and Social Psychology* 28(4). 441–456.
- Kay, Matthew, Cynthia Matuszek & Sean A. Munson. 2015. Unequal representation and gender stereotypes in image search results for occupations. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 3819–3828.
- Labov, William. 1972. *Language in the inner city: Studies in the Black English vernacular*. Philadelphia, PA: University of Pennsylvania Press.
- R Core Team. 2020. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Available at: <https://www.R-project.org/>.
- Rahman, Jacquelyn. 2008. Middle-class African Americans: reactions and attitudes toward African Americans English. *American Speech* 83(2). 141–176.
- Rickford, J.R. 1999. *African American vernacular English: Features, evolution, educational implications*. New Jersey: Wiley-Blackwell.
- Rickford, John R. & Sharese King. 2016. Language and linguistics on trial: Hearing Rachel Jeantel (and other vernacular speakers) in the courtroom and beyond. *Language* 92(4). 948–988.
- Sidnell, Jack. 2002. Outline of AAVE Grammar. Available at: https://cdt.org/wp-content/uploads/2017/11/Outline_of_AAVE_grammar__Jack_Sidnell_2002_1_Afr.pdf.
- Squires, Lauren. 2010. Enregistering internet language. *Language in Society* 39(4). 457–492.
- Van Hofwegen, Janneke & Walt Wolfram. 2010. Coming of age in African American English: A longitudinal study 1. *Journal of Sociolinguistics* 14(4). 427–455.
- Wickham, Hadley. 2016. *ggplot2: Elegant Graphics for Data Analysis*. <https://ggplot2.tidyverse.org> (accessed 5 April 2020).
- Wiley, Terrence G. & Marguerite Lukes. 1996. English-only and standard English ideologies in the U.S. *Tesol Quarterly* 30(3). 511–535.
- Wolfram, Walt. 2004. The grammar of Urban African American Vernacular English. In Bernd Kortman & Edgar W. Schneider (eds.), *A handbook of varieties of English*, vol. 2, 111–132. Berlin: Mouton de Gruyter.