

Worse for the wear: Effects of raciolinguistic ideologies, gender ideologies, and clothing on ESL pronunciation perception

Ruthanne Hughes*

Abstract. Accents are products of perception as well as production; it is crucial that accentedness research address teacher discrimination rather than focusing on so-called student "deficiencies." Raciolinguistic ideologies and gender ideologies are factors that can affect teacher perception, interacting together in nuanced, non-additive ways. This case study employing matched-guise methodology investigates a White ESL teacher's differential ratings of pronunciation based on students' race, gender, and cultural clothing. Results show that clothing helps index figures of personhood that are more than the sum of race and gender ideologies and emphasize that even progressive and linguistically-trained teachers may still engage in biased rating practices.

Keywords. raciolinguistic ideologies; gender ideologies; clothing; perception; English as a second language

1. Introduction. Second language acquisition studies investigating language learners' accents have traditionally focused on the production of accent and on accent reduction through pedagogical techniques (e.g., Hişmanoğlu, 2006; Ketabi, 2015; McNulty & Lazarevic, 2012). However, accent is a construct of perception, as well as production. Speech perception is a subjective and imprecise phenomenon, affected by factors as culturally significant as race and gender and as incidental as subconscious triggers (cf. stuffed animal placement in Hay & Drager, 2010). Learners of English are often judged on accent, not only in daily interactions with native speakers of English, but also in the classroom. Such linguistic evaluations tend to privilege the perspective of an imagined White¹, monolingual, middle-class listener, or the "white listening subject" (Flores & Rosa, 2015). Previous studies have shown discriminatory effects towards international teaching assistants, but teachers' perception of students is largely unstudied (Kang, 2010, 2012; Kang, Rubin, & Lindemann, 2015; Rubin, 2012; Rubin & Smith, 1990; Staples, Kang, & Wittner, 2014). This case study employs a matched-guise design to investigate how one English as a Second Language (ESL) teacher perceived and rated pronunciation based on nonlinguistic factors, including race, gender, and heritage clothing.

2. Raciolinguistic ideologies. Listeners rely on expectations in every aspect of listening, from phonology (the McGurk effect, e.g. McGurk & MacDonald, 1976) to syntax (garden path sentences, e.g. Sanz, Laka, & Tanenhaus, 2013) to semantics (homophones, e.g. Hawkins, Reicher, Rogers, & Peterson, 1976). Part of these expectations are rooted in simple pattern and

* Special thanks to Nina Moreno for her help with this draft and to Stewart Hughes, Anusha Ánand, Archie Crowley, Drew Crosby, and Shana Scucchi for workshopping this presentation. Author: Ruthanne Hughes, University of South Carolina (wengerj@email.sc.edu).

¹ There is debate about whether *White* should be capitalized or left lowercase. With respect to both sides of the discussion, I have chosen to capitalize *White* here, with the intention to work against implications of lowercase *white* that might serve to normalize Whiteness as the unmarked standard. The use of capitalized *White* in this work further serves to bring Whiteness into focus as an object of study and to interrogate its influence upon social relationships. I leave quotations to contain their original capitalization.

grammatical recognition, but others are born from ideological constructions and valuations. These ideological valuations move beyond cultural preferences for style and content to also include judgments of speakers. Flores and Rosa (2015) coined the term *raciolinguistic ideologies* to describe how expectations about speakers, or the bodies of speakers, relate to ideologies of standardized language and perception of speech. They define *raciolinguistic ideologies* as “conflat[ing] certain racialized bodies with linguistic deficiency unrelated to any objective linguistic practices” (Flores & Rosa, 2015: 150).

Within the classroom, the White listening subject (Flores & Rosa, 2015) may perceive language as non-standard, even when they in fact heard no non-standard language. Additionally, it is important to consider that the White listening subject is not necessarily a White person, nor even an individual (Rosa & Flores, 2017). Rosa & Flores (2017) argue that “racially hegemonic perceptions can be enacted [...] not simply by white individuals but rather by whiteness as an historical and contemporary subject position that can be situationally inhabited both by individuals recognized as white and nonwhite” (Rosa & Flores, 2017: 628). People of many positionalities can embody the White listening subject because the term describes the act of listening with the White subject position in focus. Hegemonic systems derive their power from the actions of both the oppressors and the oppressed, self-replicating through internalized oppression that resists an examination of the default paradigms of what is accepted as normal.

Institutions, such as language schools, infused with *raciolinguistic ideologies* may also act as White listening subjects, perceiving and judging student speech according to racialized standards irrespective of the attitudes of individual teachers. Even when individual teachers hold reflexive attitudes and work against racism in aspects of their private lives, they may take on the role of White listening subject in the school setting in response to *raciolinguistic ideologies* circulating in the broad discourse. An example of this can be seen in Flores, Lewis & Phuong (2018), which proposes a *raciolinguistic chronotope* perspective using data from an ethnographic study conducted over several years in a bilingual charter school. The K-8 school comprised both native English and native Spanish speakers, with roughly 85% of the student population identifying as Latinx. Half of the school day was conducted in Spanish and half in English. The authors argued for “*raciolinguistic chronotopes* (Rosa, 2016) that co-construct race and language in ways that produce particular relationships between the past, present, and future, as a point of entry for understanding the formation of this institutional listening subject position” (Flores et al., 2018: 16). A *raciolinguistic chronotope* about anxiety regarding the future of Spanish in the United States can be seen in interview data with one of the teachers at the bilingual charter school. The teacher stated the students used Spanish sounds when speaking English and discussed feeling pressure as their English teacher, arguing “if they are not performing in my class, you know potentially 20 years down the road, English being the language of America, what is that gonna look like? (Teacher Interview, 6/11/15)” (Flores et al., 2018: 19). This teacher expressed a sense of obligation toward listening to her students as a future member of society, anticipating the judgments they might cast on her students. At other points in interviews, this teacher “*expresse[d] ambivalence about the fairness of broader society*” (Flores et al., 2018: 19). However, she still embodied a White listening subject in her perceptions and assessment of her students, drawing on *raciolinguistic chronotopes* and institutional norms to justify her judgements. This teacher was not an anomaly but represented *chronotopes* and ideologies that circulate in this school. Even if individual teachers do not wish to judge a student or might not judge the student’s language if they encountered it outside of the classroom, within the institution, even an institution that values bilingualism, teachers may take on the role of the

White listening subject that will judge students in the future, believing it to be their job to prepare their students for this eventuality (Flores et al., 2018).

In affecting perception, these ideologies have serious and far-reaching effects, particularly in situations of power relations within institutions. Deviancy from standard language may be perceived even where none exists, affecting how racialized speakers are assessed and instructed. This study takes raciolinguistic ideologies as a framework through which to examine teacher perception of students.

3. Figures of personhood. This study uses raciolinguistic ideologies as a framework, but it also investigates the link between gender and clothing and processes of racialization. The additional framework of figures of personhood is useful here, as it allows the conceptualization of multiple axes of social meaning and how they may be linked to language. Compared to a traditional approach in sociolinguistics that has focused on ethnolects or regional dialects, which highlight a single characteristic (race and region, respectively), Agha's framework of enregisterment along with Eckert's account of style seem to handle this multiplicity of social dimensions elegantly. According to Agha, registers are recognizable sets of signs, linguistic and otherwise, that are linked to recognizable figures of personhood, or "indexical images of speaker-actor" (Agha, 2005: 39). The sets of signs that comprise registers include multiple modalities: "language is part of a broader semiotic system that includes such things as clothing, territory, musical taste, activities, and stances" (Eckert, 2004: 45). The images of speaker-actor may be defined along any number of social dimensions, for example, the "uptight White teacher" or the "sassy Black drag queen."

Important for Agha as well was the notion of enregisterment, which recognizes that registers are not stable facts about the world but outcomes of sociohistorical processes. These processes are driven by both metalinguistic discourse as well as the actual performance of language. Metalinguistic labels link speech repertoires and "enactable pragmatic effects" (Agha, 2004: 23) including images of speakers that may encode information about gender, class, and race. These "cultural models of speech" (Agha, 2004: 23) connect speech repertoires to categorizations of speakers and expectations about those speakers' behaviors. Eckert's notion of style also captures this, on a more individual scale. She argued that "the use of a stylistic resources affects not only the receiving style but the resource itself, feeding back into discourses of race, ethnicity, class, etc." (Eckert, 2004: 42). The process of enregisterment means that registers and styles both constitute and are constituted by the circulating discourses.

Importantly, figures of personhood are ideological phenomena, not biographical facts of a speaker's identity, such that when a speaker uses a particular accent that invokes a particular figure of personhood, that figure may or may not parallel a speaker's identity. This can be seen through practices referred to as "crossing," where individuals may use language that "feels anomalously 'other'" (Rampton, 1995: 54) in order to create a link between the speaker's Self and the figure of the Other invoked. From another angle, the figure of personhood as ideology rather than biology can also be seen in cases of speaker misidentification. An individual's accent may be linked to a particular figure of personhood that does not match the speaker's identity. For example, when listening to a Korean native speaker speaking English as a second language, listeners may believe the speaker to be Chinese or Japanese, or even Indian or Latinx (Lindemann, 2003). This may be because the accent heard invoked a figure of personhood corresponding to these racial or national identities, or it might be that the figure of personhood activated does not specify race or nationality, perhaps instead relying on a broad nonnative

categorization. Either way, however, the image may participate as its own figure in discourse and ideological circulation, impacting the way the speaker is perceived.

In addition, while sociolinguists have sometimes assumed that language primarily serves as clues about a speaker's race (Baugh, 2015), racial signs of the body can shape how language is heard (Flores & Rosa, 2015). In the same way that registers can enact figures of personhood (Agha, 2005), figures of personhood can also enact registers, contributing to the perception of nonnative accentedness regardless of auditory input (Rubin, 1992). If "competence in Standard Language [...] become[s] emblematic of additional attributes," (Agha, 2004: 37), acting as a stand-in for class or race, so can class, race or additional attributes become emblematic of Standard Language, or the lack of Standard Language.

4. Perception of race. The presence and impact of language ideologies related to race can be seen in studies examining the impact of perception of race on individuals' judgments of speakers. Alim (2007) observed an elementary school teacher identifying language practices of African American students (using the forms *he was* and *she was*, which are grammatical in both Standardized American English and African American Language) as a "major problem." This is notable because those same practices would have been regarded as "standard language" if performed by White students. Similarly, experimental studies have found that the perceived race of the speaker corresponded with lower ratings and lower comprehension when raters believed they were hearing an East Asian person versus a White person (Kang & Rubin, 2009; Rubin, 1992; Rubin, Ainsworth, Cho, Turk, & Winn, 1999). These studies employed matched-guise designs where raters saw a photo of an instructor who was either East Asian or White but heard the voice of a native English speaker in both guises. When viewing the East Asian guises, raters provided significantly worse judgments and performed worse on comprehension tasks. Gender may also exacerbate stigmatized racial identities. Ramjattan (2017) argued that gender and race interacted in the perception of non-White English teachers so that gender compounded stigmatized racial identities. Qualitative interviews conducted among ten non-White English language teachers revealed that students engaged in racist nativist microaggressions and perceived non-White teachers nonnative, regardless of their nationality or native language. Gender interacted with these racialized identities, with non-White women receiving negative reactions for both their race and their gender.

Raciolinguistic ideologies are significant forces in communities that can impact the instruction and assessment of racialized students. However, while these ideologies can be seen to circulate in communities through ethnographic observation, ideologies are not discrete entities whose effects can be directly quantified. A related concept, reverse linguistic stereotyping, does provide the opportunity to quantifiably operationalize of the racist behavior that is both rooted in and reifying raciolinguistic ideologies. Reverse linguistic stereotyping is the converse of the linguistic stereotyping hypothesis, which states that "even brief samples of speech varieties associated with low-prestige groups can cue negative attributions regarding individual speakers" (Kang & Rubin, 2009: 441). Reverse linguistic stereotyping, therefore, means that perception of low-prestige group membership can cue distorted perceptions of that speaker's speech. This is expressed in Kang & Rubin's (2009) statement that "listeners often hear what they expect to hear, rather than accurately perceive nonnative student speech" (Kang & Rubin, 2009: 451).

Kang and Rubin (Kang, 2008, 2012; Kang & Rubin, 2009, 2014; Kang et al., 2015; Rubin, 2012; Staples et al., 2014) measure a proclivity to engage in reverse linguistic stereotyping by measuring listener language attitudes towards speakers believed to be Euro-American native speakers in contrast to attitudes towards speakers believed to be "foreign" nonnative speakers.

However, for both guises, participants hear the same voice. The difference in the attitudes for these two guises, measured through Zahn and Hopper's (1985) Speech Evaluation Instrument, yields a score showing the participant's proclivity towards reverse linguistic stereotyping. Kang, in particular, has examined reverse linguistic stereotyping in the classroom, with a focus on identifying rater characteristics related to a proclivity towards reverse linguistic stereotyping. These individual differences include the rater's native English speaker status (Kang, 2008, 2012; Kang & Rubin, 2009), language background (Kang & Rubin, 2009; Kang, Vo, & Moran, 2016), linguistic knowledge and training (Kang, 2008, 2012), teaching experience (Kang, 2008, 2012; Kang & Rubin, 2009), and contact with nonnative speakers (Kang, 2008, 2012; Kang & Rubin, 2009; Staples et al., 2014).

Reverse linguistic stereotyping can take many forms. In a study of 74 campaign speeches delivered by White Republican and Democrat presidential candidates, Dupree & Fiske (2019) found that White liberals self-presented less competence when addressing racial minority interlocutors as opposed to White interlocutors. Meanwhile, no such effect was found for conservatives. This suggests something interesting about identity: people with a more overt bias against racism displayed competence-downshift as a reaction to low-competence stereotypes of minorities. If competence-downshift affects ESL teachers, it could be expected that they may actually lower expectations of pronunciation for groups they expect to have worse pronunciation, leading to artificially high ratings for those individuals.

5. Perception of gender. Other ideologies relevant to the construction of figures of personhood may be those of gender (O'Loughlin, 2007; Ramjattan, 2017; Strand, 1999; Thompson, 1991). In fact, the terminology of the listening subject was coined in regard to ideologies of gender and language. Inoue (2003) argued that the listening subject was instrumental in the creation of a category of Japanese woman who speaks Japanese women's language. Inoue describes how, during the late 19th and early 20th centuries, intellectual men wrote about the reported speech of school-girls, criticizing their speech as a corrupt form of language. Despite the small number of girls and women speaking in the reported ways, this ideology of schoolgirl language became so common that it eventually came to be synonymous with Japanese women's language and shifted in connotation, coming to be held up as a sign of desirable Japanese womanhood. The listening subject in this case held great power to shape not only the way Japanese women were perceived, but the ideologies surrounding women's language and the identities of Japanese women in modernity.

Gender ideologies are complex and circulate in both macro national discourses and micro daily interactions. On a smaller level than that of Japanese national discourse, ideologies of gender in education are common (Eckert & McConnell-Ginet, 2013). K-12 school practices in the United States reinforce a gendered order with language like "boys and girls" or practices like gender-segregated lines, classes, or physical activities. The gender balance of adults working in education also reinforces ideas about women's and men's work and characteristics, with administrators being predominantly men, but teachers being predominantly women. Further, the idea that women are more nurturing and men are more skilled is echoed in the fact that men are more likely to be teachers of older students. This is particularly visible at the college and university level.

More specifically to the language learning context, gender has been viewed as correlated with language ability. A common folk theory of language says that women are better language learners. And indeed, some research has provided arguments for this. For example, some studies have shown women to receive better ratings of accentedness than men (Thompson, 1991). It is

important to note, though, that few studies have teased out whether this apparent difference in language skill is due to objective linguistic competence or to rater bias. However, several matched-guise studies have concluded that the perception of some aspects of language, specifically the perception of phoneme boundaries, are impacted by perception of gender (Johnson & Strand, 1998; Strand, 1999; Strand & Johnson, 1996). Ideologies of gender require further study, particularly in light of investigation into raciolinguistic ideologies.

6. Other indexical factors. As large and salient as the concepts of race and gender are, they are not solely responsible for construction of identity. Figures of personhood are comprised of multiple indexical resources. This study also addresses clothing as an indexical factor. Numerous ethnographic studies have highlighted ways in which clothing indexes aspects of identity (Chun, 2011; Ramjattan, 2017; Zentella, 2016), but it has not previously been introduced as a variable within perceptual experiments. Clothing may work to index additional factors unrelated to gender or race, or it may also serve to accentuate those features. From a perceptual standpoint, clothing may act to intensify stereotypes about certain figures of personhood.

7. Methodology. Research questions for this study included: How did the teacher rate pronunciation differently based on students' perceived 1) race, 2) gender, and 3) clothing? In order to address these questions, this study employed a matched-guise design. Audio stimuli included four White American young adult voices, including two men and two women. All were speakers of Standardized American English, as determined by the researcher. Two were from South Carolina, one was from Ohio, and one was from Maryland. Linguistic stimuli included 24 sentences containing a word with /ʒ/, balanced for whether the target phoneme appeared word-medially and word-finally and whether the target word was placed in sentence-initial, sentence-medial, or sentence-final position.

Visual stimuli included twelve photos. Three race conditions were represented: Arabic, Chinese, and White. For each condition, there were two photos of men and two photos of women. For each sub-condition, one person was wearing heritage clothing and the other was wearing nonheritage clothing. In the Chinese condition, heritage clothing was represented by the woman wearing a *qipao* and the man wearing a *changshan*. In the Arabic condition, the woman was wearing a *hijab* and the man a *kufiya*. In the White condition, heritage clothing was represented by the person wearing clothing displaying a United States flag. Prior to the start of the experiment, photos were submitted to a native-viewer survey to ensure that photos were perceived as the race intended. Photos that were most consistently perceived as intended were chosen for the study.

The participant in this case study was a young, politically progressive White woman with linguistic training, teaching ESL at a university level English program at a university in the American South. The procedure of the experiment included three steps. First, the participant saw a picture of a student and heard an audio clip of a sentence being read. It is important to note that the participant heard native speakers but, following the intention of the matched-guise design, believed she was listening to ESL students. Second, the participant was asked to rate the nativelikeness of the pronunciation of the target word on a seven-point Likert scale. Finally, the participant completed a background demographic survey.

8. Analysis. Raw data was managed in Excel and visualized through pivot tables showing the interactions between the average ratings related to three variables: race, gender, and clothing. As this case study included results from a single teacher, descriptive statistics, including averages and ranges, were generated.

9. Results. When average ratings by race, by gender, and by clothing type are examined individually, they do not reveal any differences in ratings. Arabic, Chinese, and White photos each received an average rating of 6.25 (Figure 1). Men received an average rating of 6.17 and women an average of 6.33 (Figure 2). Nonheritage clothing received an average of 6 and heritage clothing an average of 6.5 (Figure 3).

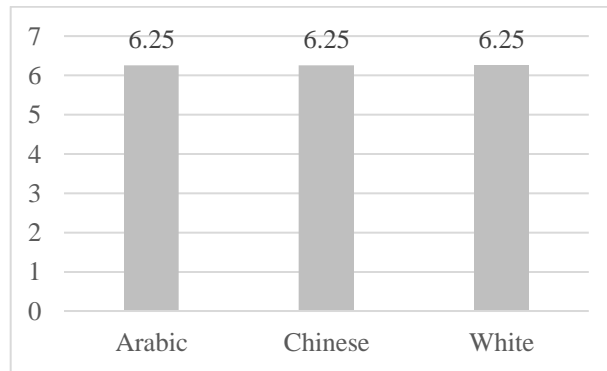


Figure 1. Average ratings by race

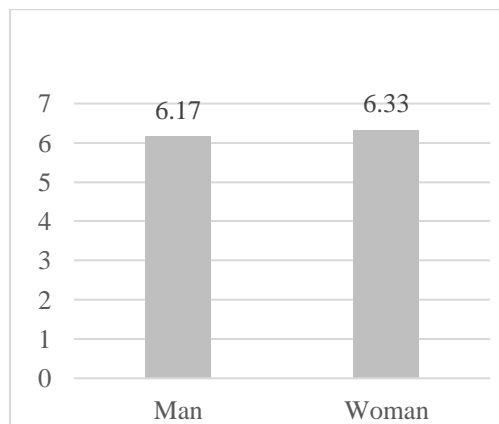


Figure 2. Average ratings by gender



Figure 3. Average ratings by clothing type

When the interaction of race and gender is examined, similar non-results are seen. Figure 4 shows that Arabic men and Arabic women both received an average rating of 6.25. Chinese men

and women also received the same rating, 6.25. White men did receive a slightly different average, of 6, with White women receiving an average rating of 6.5.

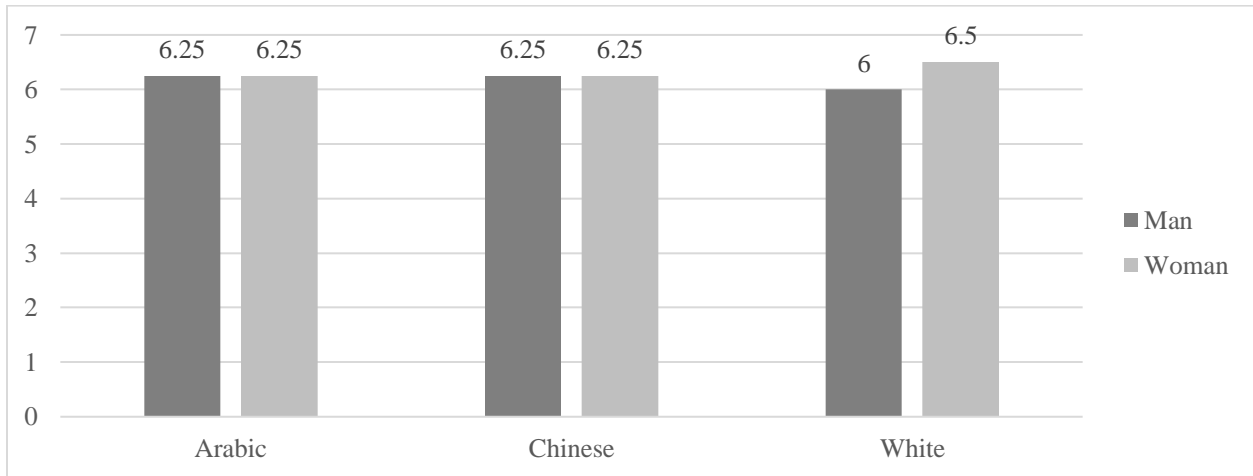


Figure 4. Average ratings by interaction of race and gender

When the interaction of gender and clothing is examined, slightly stronger results emerge. As seen in Figure 5, men wearing nonheritage clothing received an average rating of 5.67 while women wearing nonheritage clothing received an average of 6.33. Men wearing heritage clothing received an average rating of 6.67 while women in heritage clothing received a rating of 6.33. In sum, when wearing nonheritage clothing, women were rated higher than men. When wearing heritage clothing, women were rated lower than men.

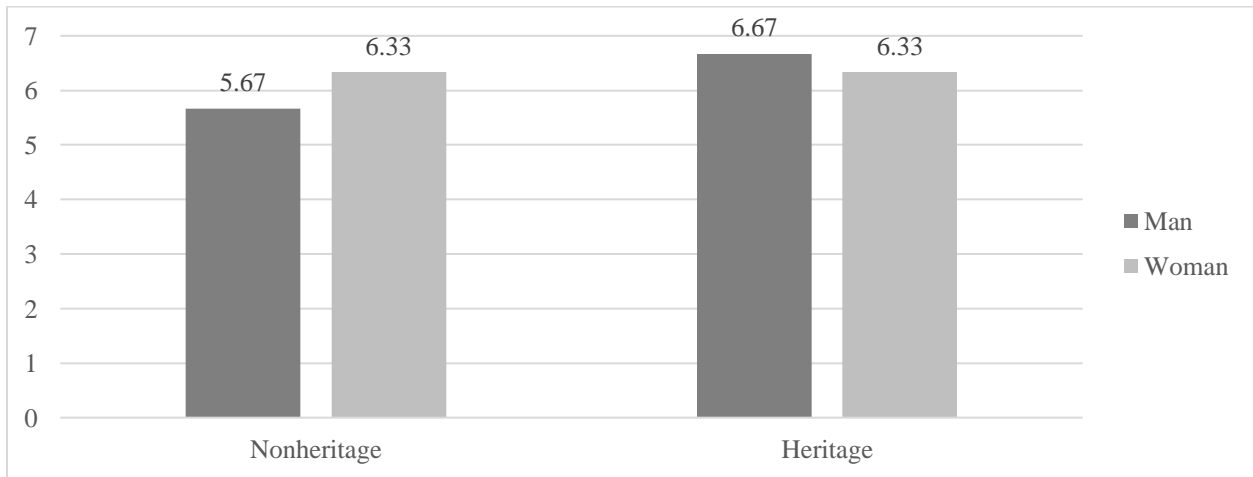


Figure 5. Average ratings by interaction of gender and clothing

The interaction of race and clothing also reveals more differences in ratings. Figure 6 shows that Arabic students received an average rating of 6.75 when wearing nonheritage clothing, but 5.75 when wearing heritage clothing. Chinese students received a rating of 5.75 when wearing nonheritage clothing and 6.75 when wearing heritage clothing. White students received a rating of 5.5 when wearing nonheritage clothing and 7 when wearing heritage clothing. While the nonheritage clothing received higher ratings than the heritage clothing for the Arabic guise, the opposite is true for the other two guises. Both Chinese and White ratings were higher for heritage

clothing than nonheritage clothing, with the White students in heritage clothing receiving the highest possible rating.

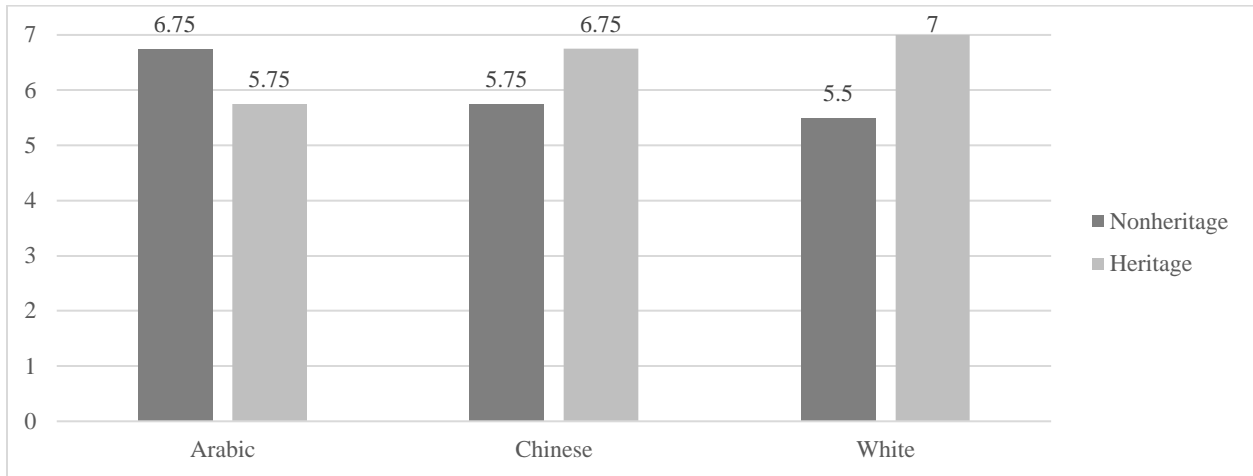


Figure 6. Average ratings by interaction of race and clothing

This phenomenon with Chinese and White guises patterning together, and opposite to the pattern of the Arabic guise, is even more apparent when all three variables are considered together, shown in Figure 7. For the Arabic guise, nonheritage clothing was correlated with higher ratings for both genders. However, for both the Chinese and White guises, heritage clothing was correlated with higher ratings for both genders.

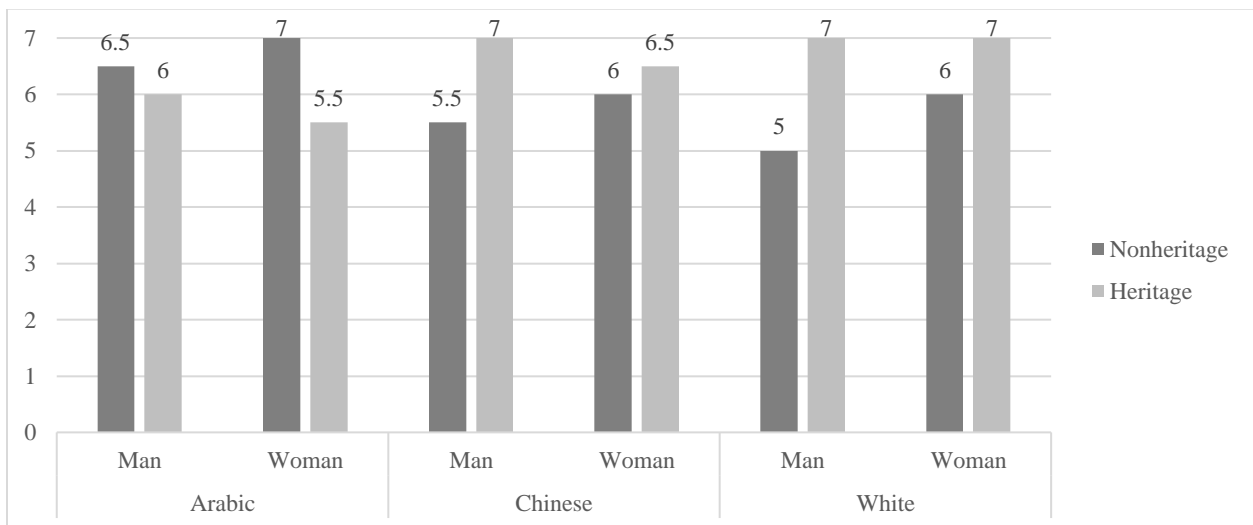


Figure 7. Average ratings by interaction of race, gender, and clothing

This racialized pattern can also be seen in Table 1, which highlights the comparisons between men and women within each race condition and clothing type. For both Arabic heritage clothing and Arabic nonheritage clothing, men scored better than women. For the Chinese and White guises, however, women scored better than men for both heritage and nonheritage clothing. Comparing men and women based on the clothing type shows that women were rated worse than men for the Arabic guise, but better than men for the Chinese and White guises.

	Arabic		Chinese		White	
	Heritage clothing	Nonheritage clothing	Heritage clothing	Nonheritage clothing	Heritage clothing	Nonheritage clothing
Woman	6	5.5	7	6.5	7	7
Man	6.5	7	5.5	6	5	6

Table 1. Comparison of gender, within race and clothing conditions

Similarly, Table 2 highlights the comparison of clothing within each race and gender category and shows the same pattern. Heritage clothing was correlated with worse ratings for the Arabic guise but was correlated with better ratings for the Chinese and White guises.

	Arabic		Chinese		White	
	Man	Woman	Man	Woman	Man	Woman
Heritage	6	5.5	7	6.5	7	7
Nonheritage	6.5	7	5.5	6	5	6

Table 2. Comparison of clothing, within race and gender conditions

On the level of individual photographs, three individual figures received a 5, the lowest score given. Those three figures included the Arabic woman in heritage clothing, the White man in nonheritage clothing, and the Chinese man with nonheritage clothing. The Arabic woman's rating follows from the interactions between gender, race, and clothing discussed above. However, the low results for the White and Chinese men in nonheritage clothing is surprising. These low ratings for the nonheritage guise may be explained by an expectation downshift for racialized students. If the teacher lowered her expectations for racialized students that she expected to perform worse, she may have inflated ratings for those students, leading to harsher ratings for students for whom there were higher expectations. Additionally, the American flag may have been particularly significant for ratings in the White guise. Due to the lack of clear heritage clothing in American culture, the flag was selected to represent traditional American identity. However, this decision may have introduced additional variables. White students wearing the flag received the highest possible ratings. White students wearing nonheritage clothing received some of the lowest ratings, however, particularly for the men. While the nonheritage category ratings may be explained by an expectation downshift, the heritage category may also have been inflated by effects of the flag itself. On the one hand, the flag may have activated an expectation that White American students are native English speakers and therefore perfectly natively like. On the other, the flag may have activated other stereotypes or expectations based on patriotic ideologies. Patriotic ideologies are particularly strong in the United States, and there is evidence to suggest that even minor indexes of nationality are able to sway the perception of speech (Hay & Drager, 2010). Future research should explore other avenues of measuring heritage clothing in the White guise that do not introduce these complications.

10. Discussion. Results of this study align with previous work conducted in both the field of second language acquisition and the field of sociolinguistics. The teacher in this case study did

participate in reverse linguistic stereotyping (Kang & Rubin, 2009; Rubin, 1992), with ratings showing effects of nonlinguistic factors, including gender and clothing in addition to race. Her ratings also displayed competence downshift effects (Dupree & Fiske, 2019) and indicated that even a progressive, well-intentioned, and well-trained teacher can participate in biased rating behavior. Although this study did not explicitly examine broader discourses or ideologies within which this teacher was situated, this reverse linguistic stereotyping can be understood as linked to raciolinguistic ideologies (Flores & Rosa, 2015) surrounding expectations for how certain figures of personhood (Agha, 2004) will speak. These results indicate that these figures of personhood are complex, involving more than the salient, macro factors of race and gender. In particular, clothing was key in mediating figures of personhood. Future research should further develop the understanding of the role of this variable and of other semiotic signs.

11. Limitations and future directions. Although this study is a case study with limited generalizability, it raises interesting implications for future research. Further studies should first address several limitations of the present study, including greater control of voice and photo conditions. The voices in the present study were not controlled for pitch, speech rate, or similarity. Further research should introduce greater controls and native-listener tasks in order to norm the voices prior to the experiment. Photo conditions in the present study may also have been too variable. Future studies should control for the amount of clothing visible in each photo. Qualitative observation should also be employed in order to more fully understand what is indexed by clothing and to determine if it may be a compounded variable, perhaps indexing other variables such as class. Further, all photos should be subjected to a native-viewer task prior to the start of the experiment in order to ascertain that they are perceived as representing all conditions expected.

Future directions of this research will also include the expansion and diversification of the participant pool, as well as the inclusion of speakers of other varieties of English in audio stimuli. Future work will combine qualitative and quantitative methods in order to investigate ideologies circulating in teaching institutions and the relationship between these ideologies and rating behavior.

12. Conclusions. When race, gender, and heritage clothing were considered independently, results showed no differences in ratings. However, differences emerged in the interaction of the three variables, with clothing mediating interactions. To understand these results, we must move from an examination of individual variables to a framework of figures of personhood. Figures of personhood are complex, with interactions between multiple variables. The interaction between race and gender in these results was mediated by clothing, such that Arabic student were hampered by heritage clothing, as compared to the White and Chinese guises, and Arabic women in particular were hampered by gender and clothing, as compared to Chinese and White women. Clothing as a variable indicates more meaningful social dynamics at work in teachers' perception of students, beyond race and gender alone. Race, gender, and clothing interact in nuanced, non-additive ways, indexing a holistic figure, with clothing helping to index figures of personhood that are more than the sum of race and gender ideologies.

For the sake of students, especially those who are already in a marginalized position within the institution of education, the effects of nonlinguistic factors on teachers' rating behaviors must be further investigated as teachers and researchers alike seek to mitigate the harmful effects of racial bias. The interactions of race, gender, and clothing in the figures of personhood in this study highlight the need for intersectional approaches to discrimination. Further, these results emphasize that even progressive and linguistically trained teachers may still engage in biased

rating practices. By shifting the focus of accent research from the speaker to the White listening subject, this research holds implications for moving towards systemic change of accent discrimination, rather than a focus on individual remediation.

References

- Agha, Asif. 2004. Registers of language. In Alessandro Duranti (ed.), *A companion to linguistic anthropology*. 23–45. Malden, MA: Blackwell.
<https://doi.org/10.1002/9780470996522.ch2>.
- Agha, Asif. 2005. Voice, footing, enregisterment. *Journal of Linguistic Anthropology* 15(1).
<https://doi.org/10.1525/jlin.2005.15.1.38>.
- Alim, H. Samy. 2007. Critical hip-hop language pedagogies: Combat, consciousness, and the cultural politics of communication. *Journal of Language, Identity, and Education* 6(2). 161–167. <https://www.joycerain.com/uploads/2/3/2/0/23207256/criticalhiphoplanguage.pdf>.
- Baugh, John. 2015. SWB (Speaking while Black). In Lanehart, Sonja (ed.), *The Oxford handbook of African American language*. 755–769. Oxford: Oxford University Press.
<https://dx.doi.org/10.1093/oxfordhb/9780199795390.013.68>.
- Chun, Elaine. 2011. Reading race beyond black and white. *Discourse and Society* 22(4), 403–421. <https://doi.org/10.1177/0957926510395833>.
- Dupree, Cydney H. & Susan Fiske. 2019. Self-presentation in interracial settings: The competence downshift by white liberals. *Journal of Personality and Social Psychology* 117(3). 579–604. <https://doi.org/10.31234/OSF.IO/PV2AB>.
- Eckert, Penelope. 2004. The meaning of style. In Chiang, Wai Fong, Elaine Chun, Lara Mahalingappa, & Siri Mehus (eds.), *Proceedings of the Eleventh Annual Symposium About Language and Society—Austin, Texas Linguistic Forum [SALSA]*. 41–53.
<https://web.stanford.edu/~eckert/PDF/salsa2003.pdf>.
- Eckert, Penelope & Sally McConnell-Ginet. 2013. *Language and Gender*, 2nd edn. New York: Cambridge University Press. <https://doi.org/10.1017/CBO9781139245883>.
- Flores, Nelson, Mark C. Lewis & Jennifer Phuong. 2018. Raciolinguistic chronotopes and the education of Latinx students: Resistance and anxiety in a bilingual school. *Language & Communication*, 62. 15–25. <https://doi.org/10.1016/j.langcom.2018.06.002>.
- Flores, Nelson & Jonathan D. Rosa. 2015. Undoing appropriateness: Raciolinguistic ideologies and language diversity in education. *Harvard Educational Review*, 85(2). 149–171.
<https://doi.org/10.17763/0017-8055.85.2.149>.
- Hawkins, Harold L., Gerlad M. Reicher, Miriam Rogers & Lesli Peterson. 1976. Flexible coding in word recognition. *Journal of Experimental Psychology: Human Perception and Performance* 2(3). 380–385. <https://doi.org/10.1037/0096-1523.2.3.380>.
- Hay, Jennifer & Katie Drager. 2010. Stuffed toys and speech perception. *Linguistics* 48(4). 865–892. <https://doi.org/10.1515/LING.2010.027>.
- Hişmanoğlu, Murat. 2006. Current perspectives on pronunciation learning and teaching. *Journal of Language and Linguistic Studies* 2(1).
<https://dergipark.org.tr/en/pub/jlls/issue/9923/122840>.
- Inoue, Miyako. 2003. The listening subject of Japanese modernity and his auditory double: Citing, sighting, and siting the modern Japanese woman. *Cultural Anthropology* 18(2). 156–193. <https://doi.org/10.1525/can.2003.18.2.156>.
- Johnson, Keith & Elizabeth A. Strand. 1998. Speech perception without normalization or higher order invariants. In “Paradigms and Problems in Phonetic Research” workshop at the Annual Meeting of the Linguistic Society of America, New York City.

- Kang, Okim. 2008. Ratings of L2 oral performance in English: Relative impact of rater characteristics and acoustic measures of accentedness. *Spain Fellow Working Papers in Second or Foreign Language Assessment* 6. 181–205. <https://doi.org/10.1080/15434303.2011.642631>.
- Kang, Okim. 2010. Relative salience of suprasegmental features on judgments of L2 comprehensibility and accentedness. *System* 38(2). 301–315. <https://doi.org/10.1016/j.system.2010.01.005>.
- Kang, Okim. 2012. Impact of rater characteristics and prosodic features of speaker accentedness on ratings of international teaching assistants' oral performance. *Language Assessment Quarterly* 9(3). 249–269. <https://doi.org/10.1080/15434303.2011.642631>.
- 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. <https://doi.org/10.1177/0261927X09341950>.
- Kang, Okim & Donald L. Rubin. 2014. Listener expectations, reverse linguistic stereotyping, and individual background factors in social judgments and oral performance assessment. In John M. Levis & Alene Moyer (eds.), *Social Dynamics in Second Language Accent* 10. Berlin: De Gruyter Mouton.
- Kang, Okim, Donald L. Rubin & Stephanie Lindemann. 2015. Mitigating U.S. undergraduates' attitudes toward international teaching assistants. *TESOL Quarterly* 49(4). 681–706. <https://doi.org/10.1002/tesq.192>
- Kang, Okim, Son Ca Thanh Vo & Meghan Moran. 2016. Perceptual judgments of accented speech by listeners from different first language backgrounds. *The Electronic Journal for English as a Second Language* 20(1). 1–24. <http://www.tesl-ej.org/pdf/ej77/a1.pdf>.
- Ketabi, Saeed. 2015. Pronunciation teaching: Past and present. *International Journal of Applied Linguistics & English Literature* 4(5). 2200–3592. <https://doi.org/10.7575/aiac.ijalel.v.4n.5p.182>
- Lindemann, Stephanie. 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>.
- McGurk, Harry & John MacDonald. 1976. Hearing lips and seeing voices. *Nature* 264. 746–748. <https://doi.org/10.1038/260170a0>.
- McNulty, Anastassia, & Bojan Lazarevic, B. 2012. Best practices in using video technology. *Teaching English with Technology* 12(3). 49–61. <https://eric.ed.gov/?id=EJ1144964>.
- O'Loughlin, Kieran. 2007. An investigation into the role of gender in the IELTS oral interview. In Linda Taylor & Peter Falvey (eds.), *IELTS Collected Papers: Research in Speaking and Writing Assessment*. Cambridge: Cambridge University Press. <https://www.ielts.org/teaching-and-research/research-reports/volume-03-report-1>.
- Ramjattan, Vijay. 2017. Racist nativist microaggressions and the professional resistance of racialized English language teachers in Toronto. *Race Ethnicity and Education* 1–17. <https://doi.org/10.1080/13613324.2017.1377171>.
- Rampton, Ben. 1995. Language crossing and the problematisation of ethnicity and socialisation. *Pragmatics* 5(4). 485–513. <https://doi.org/10.1075/prag.5.4.04ram>.
- Rosa, Jonathan D. & Nelson Flores. 2017. Unsettling race and language: Toward a raciolinguistic perspective. *Language in Society* 46. 621–647. <https://doi.org/10.1017/S0047404517000562>.

- Rubin, Donald L. 1992. Nonlanguage factors affecting undergraduates' judgments of nonnative English-speaking teaching assistants. *Research in Higher Education* 33(4). 511–531. <https://doi.org/10.1007/BF00973770>.
- Rubin, Donald L. 2012. The power of prejudice in accent perception: Reverse linguistic stereotyping and its impact on listener judgements and decisions. In John Levis & Kimberly LeVelle (eds.), *Proceedings of the 3rd Pronunciation in Second Language Learning and Teaching Conference*. 11–17. Ames, IA: Iowa State University. <https://doi.org/10.13140/RG.2.1.1465.4485>.
- Rubin, Donald L., Stuart Ainsworth, Eunsook Cho, Don Turk & Laura Winn. 1999. Are Greek letter social organizations a factor in undergraduates' perceptions of international instructors? *International Journal of Intercultural Relations* 23(1). 1–12. [https://doi.org/10.1016/S0147-1767\(98\)00023-6](https://doi.org/10.1016/S0147-1767(98)00023-6).
- Rubin, Donald L. & Kim A. Smith. 1990. Effects of accent, ethnicity, and lecture topic on undergraduates' perceptions of nonnative English-speaking teaching assistants. *International Journal of Intercultural Relations* 14. 337–353. [https://doi.org/10.1016/0147-1767\(90\)90019-S](https://doi.org/10.1016/0147-1767(90)90019-S).
- Sanz, Montserrat, Itziar Laka, & Michael K. Tanenhaus (eds.). 2013. *Language down the garden path: The cognitive and biological basis for linguistic structures*. Oxford University Press.
- Staples, Shelley, Okim Kang, & Elizabeth Wittner. 2014. Considering interlocutors in university discourse communities: Impacting U.S. undergraduates' perceptions of ITAs through a structured contact program. *English for Specific Purposes* 35(1). 54–65. <https://doi.org/10.1016/j.esp.2014.02.002>.
- Strand, Elizabeth A. 1999. Uncovering the role of gender stereotypes in speech perception. *Journal of Language and Social Psychology* 18(1). 86–100. <https://doi.org/10.1177/0261927X99018001006>.
- Strand, Elizabeth A. & Keith Johnson. 1996. Gradient and visual speaker normalization in the perception of fricatives. In *Natural Language Processing and Speech Technology: Results of the 3rd KONVENS [Conference on Natural Language Processing] Conference*. 14–26.
- Thompson, Irene. 1991. Foreign accents revisited: The English pronunciation of Russian immigrants. *Language Learning* 41(2). 177–204. <https://doi.org/10.1111/j.1467-1770.1991.tb00683.x>.
- Zahn, Christopher J. & Robert Hopper. 1985. Measuring language attitudes: The speech evaluation instrument. *Journal of Language and Social Psychology* 4(2). 113–123. <https://doi.org/10.1177/0261927X8500400203>.
- Zentella, Ana Celia. 2016. “Socials,” “Poch@s,” “Normals” y los demás: School networks and linguistic capital of high school students on the Tijuana-San Diego border. In H. Samy Alim, John R. Rickford, & Arnetha F. Ball (eds.), *Raciolinguistics: How language shapes our ideas about race*. 327–346. New York: Oxford.