Proceedings of the LSA

# Sounding Black, looking White: Perceptions of racial identity in pop music

Abby Killam*

**Abstract**. This study investigates how listeners perceive racialized linguistic cues in pop songs, specifically examining how Black and White singers are evaluated when using features of African American English (AAE). To create the stimuli, four professional American singers sung two pop songs with paired lyrics, one in an AAE guise and one in a Mainstream U.S. English (MUSE) guise. Using data from focus groups ($N = 36$) and an online survey ($N = 157$), participants identified the perceived racial identity of each singer across both auditory-only and audiovisual conditions, specified whether they heard differences between AAE- and MUSE-guise lyrics, and evaluated each singer. Results show that listeners were less accurate in identifying the White singers' racial identity and that the AAE versions were generally preferred across all singers, suggesting that AAE is largely normalized within the pop genre. Visual cues had minimal effects on singer evaluations, indicating that White artists using AAE were not necessarily viewed less favorably when their racial identity was made explicit. These findings highlight the role of genre and raciolinguistic ideologies in shaping listener expectations when listening to music and contribute to broader discussions on how Black linguistic practices are systematically commodified in the American music industry.

**Keywords**. African American English; appropriation; perception; music; raciolinguistic ideologies

**1. Introduction**. The pop genre, among many others, has been built on ritualized appropriation and commercialization of Black cultural and musical forms dating all the way back to the beginnings of the music industry in the United States (Alim et al. 2008). From swing, jazz, and the blues, to the whitewashing of early rhythm-and-blues into 'rock'n'roll' and the more recent 'hiphopification' of pop, the music industry has systematically extracted Black musical innovation, recast it with White artists, and resold it as something 'new' or 'more palatable' for mass consumption (Gamble 2024). By stripping Black art of its deeper cultural and political significance and rebranding it instead as something 'American,' this process produces a form of "denigrated Blackness" (Brown & Kopano 2014: 254) that becomes readily subject to appropriation by White artists and more mainstream genres like pop.

Prior sociolinguistic studies have documented cases in which racialized varieties, like African American English (AAE), have been taken up by White performers to appropriate racialized identities through acts of *linguistic minstrelsy* (Bucholtz & Lopez 2011), temporarily masking their real-life social identities by donning culturally-marked styles for commercial gain. For example, Eberhardt and Freeman (2015) demonstrated how White Australian rapper Iggy Azalea backgrounds her Whiteness through 'hyperperformance' of AAE features to construct a sonically Black persona. Similarly, in her analysis of the music video "7 rings," Cherid (2021) showed how White American pop artist Ariana Grande engaged in 'blackfishing,' a practice in which non-Black individuals adopt aesthetic markers of Black culture, such as specific makeup styles,

* I am profoundly grateful to Jen Nycz, Nadja Tadic, Minnie Quartey, Erin Tirpak, Claire Henderson, Rachel Milito, members of Georgetown Variation Lab, and every single one of my participants for making this work possible.
Author: Abby Killam, Georgetown University (adj52@georgetown.edu).

hairstyles, and fashion, in the attempt to 'appear Black.' Cherid (2021) also emphasizes that this practice typically operates across modalities, encompassing both linguistic (auditory) and aesthetic (visual) forms of stylization. These currents of industry have profound effects on American audiences, the omnivorous consumers of such cultural products whose tastes are curated and circulated by the music industry. However, there remains a critical gap in understanding how *audiences* perceive and evaluate these racialized linguistic cues while engaging with such performances. This exploratory study directly addresses this gap by investigating how listeners perceive the use of AAE when sung by both Black and White artists, thus shedding light on the sociolinguistic boundaries of appropriation in the contemporary American music industry.

This study has two primary goals. First, to examine how American audiences perceive and evaluate racialized linguistic cues (i.e., features of AAE) when listening to music. Second, to explore how an artist's phenotypic appearance (e.g., a White face) influences listener evaluations across auditory-only and audiovisual conditions, assessing whether expectancy violations arise from a perceived mismatch between linguistic and visual cues based on the listener's raciolinguistic ideologies. To address these questions, this paper presents two complementary studies: a series of focus groups (Study 1) and a perception survey experiment (Study 2). In both studies, listeners were asked to parse and evaluate both auditory-only and then audiovisual stimuli to determine how racial identity is interpreted in music when potentially ideologically conflicting cues are present (i.e., a White singer performing in an AAE guise). I begin by reviewing prior research on auditory and visual perceptions of race, as well as how genre functions as an expectation-priming context for language use. I then present Study 1, which uses focus group discussions and a semi-matched guise experiment to explore how listeners interpret singers' racial identities in the context of performance. Next, I present Study 2, a matched-guise survey experiment designed to assess whether evaluations of singers change over auditory-only and audiovisual conditions. Across both studies, the results suggest that audiences rely on broader socioindexical knowledge and genre as a structural framework for perception when interpreting the social meaning of linguistic variation in performative contexts. Overall, participants generally exhibited higher accuracy in classifying Black singers and expressed an overall preference for the use of AAE across all singers, suggesting that AAE is largely accepted or normalized within the pop genre. The presence of visual cues had very minimal impact on singer evaluations, indicating that White artists performing in the AAE guise were not viewed less favorably by listeners when their racial identity was made visually explicit. The paper concludes by examining the broader implications of these findings in light of the American music industry.

**2. Literature review**. Linguistic expectations concerning speaker identity are shaped by both *auditory* (e.g., Niedzielski, 1999) and *visual cues* (e.g., Hay & Drager 2010). These multimodal perceptions are mediated by an individual's socioindexical knowledge, that is, their culturally-specific understanding of how linguistic varieties and features are associated with enregistered ways of speaking. However, these perceptual processes become much more complex when individuals encounter ideologically dissonant audiovisual cues, such as when a White singer adopts Black speech features in performance. Such mismatches can produce heightened physiological and cognitive responses in the brain known as *expectancy violations,* which require listeners to spend significant neurological effort to reconcile the conflicting stimuli with preexisting social expectations (Hansen et al. 2017). Investigating such moments of dissonance therefore provides insight into how sociocultural expectations both shape and constrain perceptual processes.

Within the United States, race has often been treated as something that is self-evident, a kind of social identity that can be both 'seen' and 'heard.' However, and in contrast to this folk belief,

this paper defines *race* as a social construct, more accurately understood as a social practice, expressed through language, cultural rituals, symbols, ideologies, and the communities that individuals regularly engage with (Obasogie 2010). This perspective aligns with Hebdige's (1984) concept of *bricolage*, in which individual resources such as cultural symbols and linguistic variables combine to form a more complex, socially-interpretable entity, and Eckert's (2008) notion of the *indexical field*, which models how social meaning is constructed and projected through acts of identity. Within this framework (Eckert 2008: 454), sociolinguistic variables carry multiple and contextually dependent meanings that reflect and embed ideology into language, and these meanings are differently interpreted depending on the experiences of language users and their competence within shared sociocultural frameworks. As these social meanings are (re)interpreted by both users and non-users, they become ideologized and embedded into broader cultural belief systems, and, over time, these ideologized meanings solidify into cultural assumptions about the linguistic variable itself, and, by extension, about the linguistic behavior deemed 'natural' for the assumed speakers employing those variables.

2.1. 'HEARING' RACE USING AUDITORY CUES. A large body of sociolinguistic literature affirms that listeners are remarkably adept at classifying voices by racial identity. For example, Purnell et al. (1999) found that listeners were 70% accurate in their judgments of an ethnically-associated dialect (i.e., AAE, Mainstream U.S. English (MUSE), Chicano English) after hearing less than a second of speech. In a review of perception studies that examined the abilities of listeners to distinguish between African American and White American voices, Thomas and Reaser (2004) found that most listeners across diverse regions and ethnic backgrounds are skillful at recognizing 'prototypical' AAE features, but they struggle to identify African American speakers who use fewer of these features. However, there is no clear consensus on which auditory cues are most salient or minimally sufficient for accurate ethnic judgments. Instead, listeners appear to rely on a mixture of cues, including the segmental quality, prosody, and voice quality of an utterance, perceiving variations in fundamental frequency, jitter, shimmer, harmonics-to-noise ratio, and vowel duration and backness (Walton & Orlikoff 1994).

2.2. 'SEEING' RACE USING VISUAL CUES. A significant body of research from psychology and neuroscience demonstrates that humans are also highly adept at categorizing race through visual cues. While visual cues can originate from any part of the body, this study specifically examines *faces* as the primary source of visual stimuli. At a preconscious level, Stein et al. (2014) found that participants detected own-race faces more quickly and easily than other-race faces, a perceptual advantage attributed to the preconscious activation of socially learned knowledge structures such as stereotypes. Similarly, Craig et al. (2022) found that judgments of gendered and racial identities were more influenced by preexisting stereotypes and associations than by specific facial features. Thus, the perception of racial identity does not merely involve visual characteristics of faces, but also crucially hinges on the implicit biases and stereotypes held by the observer, as perhaps most clearly shown by Obasogie's (2010) research working with the racial perceptions of blind participants, who, despite their lack of sight, still conceptualized race in visual terms, which additionally shaped the way they thought about and experienced interpersonal racial differences. This finding further suggests that race is less about what one physically sees than how one is socialized to 'see.'

2.3. THE IMPORTANCE OF MUSICAL GENRE. Genre functions as an expectation-priming context for language use, fundamentally and radically influencing the way that music is both performed by artists and perceived by audiences (Squires 2019; Gibson 2024). Early sociolinguistic work on

language in music noted that singers often diverge from their real-life regional variety in song. For example, Trudgill (1997) found that British singers regularly employed distinctly 'American' phonological features, which he interpreted as a form of transnational appropriation, or an 'act of conflicting identity.' However, Gibson (2011) argued that such stylization reflects not the singer's attempt to 'sound American,' per se, but rather their orientation towards a supralocal register or genre-specific style norm, which is referred to in later work as *Pop Song English* (PSE) (Gibson 2024), thus their use of stylized language is intended to index the pop genre itself rather than a particular regional or social identity.

This observation raises the question of whether racialized linguistic features in pop are heard by audience members as 'acts of conflicting identity' when produced by White artists (Trudgill 1997), or, instead, simply normalized genre conventions that have undergone processes of *indexical bleaching* (Squires 2019). Squires (2020: 1) defined indexical bleaching as the "semiotic process whereby the indexical links between a linguistic form and its social meanings are weakened or altogether lost, while the form's linguistic meaning and general grammatical and/or referential functions remain intact." If AAE features have undergone such bleaching in the pop genre, then it could be reasonably assumed that this was driven by repetition through imitation, or the sustained 'borrowing' (i.e., appropriation, or seizure) of Black aesthetics, culture, and language by White artists. As these elements are repeatedly absorbed into pop performances, they gradually become dissociated from their original sociocultural contexts and more broadly recognized as 'mainstream.' Given the prevalence of cultural 'borrowing' in the music industry, these processes of indexical bleaching may render the boundaries of linguistic appropriation more ambiguous for listeners.

**3. Audiovisual stimuli**. To create the auditory stimuli for this study, I recruited four professional American singers aged 25 to 30: a White male singer, a White female singer, a Black female singer, and a Black male singer. Crucially, the Black singers were bidialectal speakers of AAE and MUSE, whereas the White singers were monodialectal speakers of only MUSE. Each singer was instructed to perform two original mini-songs specifically designed for this study by the author using music creation tool Suno.com. The lyrics were directly sourced from Justin Bieber's discography[1]. To explore the number and type of AAE features that may influence perception, the songs were constructed with paired lyrics, where the first lyric in each pair included 1–3 AAE features (AAE: "*my mama <u>don't</u> like you"*), followed by a MUSE version of the same variable(s) (MUSE: "*my mama <u>doesn't</u> like you*"). For this study, I selected the nine most frequently occurring AAE features found in the discographies of four White pop artists (Justin Bieber, Ariana Grande, Post Malone, Meghan Trainor). These features include: copula deletion, interdental fricative fortition (e.g., *that you,* <dætʃu>), fortition of /z/ before nasals (i.e., *business*, <bɪdnɪs>), future marker *gon',* object-to-demonstrative pronoun variation (e.g., *them lips*/*those lips*), subject-to-demonstrative pronoun variation (i.e., *they mouths/their mouths*), variation in third person subject-verb agreement (e.g., *she don't like you*), remote past BEEN (e.g., *we been so close*), and habitual BE (e.g., *I be goin' through changes*). After recording these songs, a professional song mixer blended the vocals with the backtracks, ensuring an industry-like sound.

[1] This study builds on prior work examining Justin Bieber, a White Canadian artist who incorporates AAE features into his music. That analysis compared how Black and White artists used AAE across their discographies and how those patterns reflect industry norms around racialized language use.

The visual stimuli were drawn from the Chicago Face Database (Ma et al. 2015). Since this study focuses on approximating pop singers, the chosen models were rated to look highly attractive (≥4.0/5.0 or higher) and between the ages of 25–30. To ensure that the selected models were physically comparable, I then selected faces with similar facial metrics, employing the Euclidean distance metric to identify the most similar faces across different racial and gender categories: Black Female, Black Male, White Female, and White Male.

## 4. Study 1: Focus groups.

4.1. METHODS AND PROCEDURE. Participants for the focus groups were U.S. adults, recruited through snowball sampling and personal networks. In total, 23 focus groups were conducted between September 2024 and February 2025, with each group consisting of 1–3 participants and lasting approximately 45-minutes. The mean age of participants was 26.7 years (SD = 5.5), with the majority of participants being female, White, highly educated, and from the Midwest.

The focus groups were conducted over Zoom and divided into four stages, each targeting specific aspects of auditory and visual perception to elicit participants' impressions and evaluations of the singers. In the first stage, participants listened to each of the four singers perform two short songs in either an AAE or MUSE guise. After each performance, they responded to a series of open-ended questions to gauge their impressions of the singer. In the second stage, participants were presented with back-to-back paired clips, in which the same lyric was performed in both the AAE and MUSE guises. After these clips were played, participants were asked if they heard any differences between the two clips, and, if so, which version they thought best suited the singer. This second stage aimed to (1) determine whether participants actually *recognized* the use of AAE in pop music and, if so, (2) assess whether they displayed preferences for which guise seemed more fitting or more authentic for a given singer, particularly in relation to their perceived racial identity. In the third stage, participants were then shown the purported faces of the four singers and were asked whether the visual representations matched their expectations concerning the singers' racial identities based on the voices they had heard. Finally, participants filled out a demographic questionnaire.

4.2. RESULTS.

4.2.1 RACIAL IDENTITY CLASSIFICATION. In focus group sessions, participants demonstrated consistently high accuracy when identifying the racial identity of the Black singers, correctly identifying both the Black male (BM) and the Black female (BF) singer 86.1% of the time. In contrast, accuracy rates for identifying White singers were considerably lower: the White Female (WF) singer was accurately identified only 66.7% of the time, while the White male (WM) singer was accurately identified just 52.7% of the time. To assess whether participants' demographic or social factors (e.g., age, gender, education, ethnicity, region, engagement with social issues) influenced racial identification accuracy for White singers, a logistic regression was conducted. However, none of these factors significantly predicted racial identification accuracy.

Next, guise (AAE or MUSE) was examined as a factor influencing racial identity guesses, revealing distinct patterns of misclassification (see Figure 1). White singers' racial identities were less reliably categorized, particularly in the MUSE guise. Most notably, accuracy for WM dropped significantly to 33.3% when performing in the MUSE guise. A chi-square test ($\chi 2(7) = 25.98$, $p = 0.0005$) and a subsequent logistic regression analysis confirmed a statistically significant relationship between singer identity, guise, and accuracy. Across all singers, racial identity classification was more accurate in the AAE guise ($\beta=1.4164$, $p = 0.046$), with BM ($\beta=2.2938$, $p = 0.004$) and BF ($\beta=2.9054$, $p = 0.001$) significantly more likely to be correctly classified. These

findings suggest that AAE serves as a strong indexical marker of Blackness, particularly reinforcing the racial identity of the Black singers for listeners in the context of music. Additionally, this relatively stable identification rate for the Black singers across guises likely also suggests that listeners may have relied on additional racialized vocal cues in the speech signal beyond morphosyntactic and phonetic variation (e.g., variation in fundamental frequency, jitter, shimmer, HNR; see Thomas & Reason 2004). In contrast, the MUSE guise provided fewer overt indexical cues for the White singers, which may have rendered their racial identities more ambiguous for listeners, leading to higher rates of misclassification.

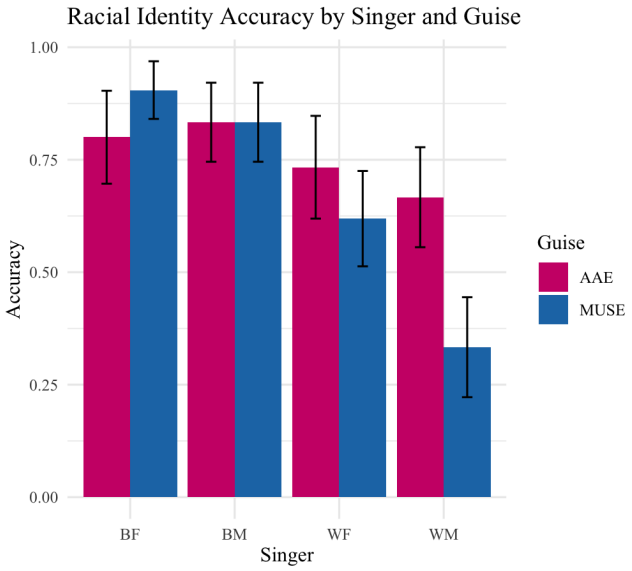


Figure 1. Racial identity guess accuracy by singer and guise (focus group)

4.2.2 NUMBER AND TYPE OF AAE FEATURE. This study also examined whether the *number* of AAE features in a lyric or the *type* of feature itself influenced listeners' detection of variation in pop lyrics. To test whether the sheer number of features presented to listeners in a single lyric affected rates of detection, a chi-square test ($\chi 2 = 0.53$, $p = 0.467$) and a logistic regression ($p = 0.388$) were conducted. Results showed no significant effect of feature density. However, these comparisons were limited by an imbalance in the stimuli, as only one contained a combination of three AAE features (copula omission, fortition, *gon'*). As such, further research is needed to more systematically assess the role of feature density in perceptual salience. Next, I examined how frequently each type of AAE feature was detected and the mean play number at which it was noticed. The results shown in Table 1 indicate notable variation in the perception of different features. Among the six features analyzed, fortition and 3rd person singular present-tense variation were the most frequently detected, with both exceeding a 65% detection rate. In contrast, habitual BE was the least detected, identified in only 11.5% of instances. These results suggest that certain AAE features are more salient in music, while others go by largely unnoticed.

To model these patterns, I conducted a mixed-effects logistic regression predicting whether participants reported hearing a linguistic feature or not, with linguistic variable type and order of presentation as predictors. Repeated exposure to a particular feature significantly increased the likelihood of detection (Play 2: Log Odds = 2.2120, $p = 0.001$, Play 3: Log Odds = 2.5934, $p < 0.001$, Play 4: Log Odds = 3.3503, $p < 0.001$), suggesting a priming effect in which multiple presentations of the same feature heightened participants' awareness of variation. Consistent

| AAE Feature | *N* Presented | *N* Heard | Proportion Heard | Mean Play Number |
|---|---|---|---|---|
| Fortition (*interdental, /z/ before nasals)* | 64 | 42 | 0.656 | 1.29 |
| 3rd Person Conjugation Variation | 52 | 34 | 0.654 | 1.26 |
| Copula Omission | 56 | 20 | 0.357 | 1.80 |
| Future Marker *Gon'* | 56 | 17 | 0.304 | 1.53 |
| Demonstrative Variation (*they, them*) | 36 | 9 | 0.250 | 1.78 |
| Habitual BE | 52 | 6 | 0.115 | 2.00 |

Table 1. Linguistic features presented and heard during focus groups

with the descriptive results, habitual BE (Factor Weight = 0.0030, Proportion Heard = 0.115, $p < 0.001$) and demonstrative pronoun variation (Factor Weight = 0.0010, Proportion Heard = 0.250, $p < 0.001$) were the least detectable features. Notably, despite being highly marked and stigmatized in most MUSE speaking contexts, these features were rarely recognized, perhaps suggesting that these features have undergone indexical bleaching in the pop genre, and thus blend in seamlessly with the music for listeners.

Copula omission (Factor Weight = 0.0150, Proportion Heard = 0.357, $p < 0.001$) and future marker *gon'* (Factor Weight = 0.0359, Proportion Heard = 0.304, $p < 0.001$) were detected more moderately and variably by participants. I argue that these features may be relatively unmarked for listeners due to the role of syllable economy in music. Since *gon'* is a shorter and a more rhythmically adaptable alternative to *gonna,* and copula omission simply removes a semantically empty verb, the AAE guise lyrics may have been perceived as more fluid and natural in musical performance. This interpretation aligns with several focus group comments, where participants described the MUSE versions of the lyrics as 'clunky' or awkward compared to their AAE counterparts. These insights suggest that such phonetic reductions may not only be expected but even preferred by audiences in musical contexts.

Finally, 3rd Person present-tense verbal variation (Factor Weight = 0.6255, Proportion Heard = 0.654, $p = 0.651$) and fortition (Proportion Heard = 0.656) were the most frequently detected AAE features, suggesting either that these features stood out due to their marked or stigmatized status compared to their MUSE counterparts or because they violated listener expectations of how lyrics should be stylized within the pop genre. A closer analysis of fortition revealed a notable difference in detection rates depending on the phonological environment. While word-initial fortition of an interdental fricative (e.g., *that you,* [dætʃu]) was less remarked upon, the fortition of /z/ before nasals in the lexical item (i.e., *business*, [bɪdnɪs]) was highly noticeable. Many listeners did not recognize it as a legitimate word to be used in the pop genre by White singers and reacted strongly, perceiving it as inauthentic and forced:

> **P14:** "When she [WF singer] was like, *business* [bɪd.nɪs], like, I hated that. So I'm like, your voice does not match this at all… you're trying to use Black slang without being Black. Like, that's what I thought when I heard that. And then the next block, when she said *business* [bɪz.nɪs], like, even that sounded weird, because I'm like, well, now the word sounds too long. There's no winning."
> **P20:** "I noticed, like the- the first clip the singer said, or saying *business* [bɪz.nɪs]. And then after that, she said, *business* [bɪz.nɪs]… it did feel like a little bit more jarring to hear her say *business* [bɪd.nɪs], just because, like, I don't really think that fits with her. Like, I

feel like that's more like, like, more like, I guess, African American kind of English, or like, yeah, like dialect, yeah."

These frequent comments suggest that the fortition of /z/ in *business* triggered expectancy violations, thus reflecting a disruption of norms of language use in the pop genre and the activation of raciolinguistic ideologies, as participants explicitly linked this feature to the singer's racial identity and evaluated its appropriateness in context.

Regarding 3rd person singular variation, many commenters evaluated this feature metalinguistically, framing it not as acceptable stylistic choice but as an outright grammatical error. Unlike other AAE features that were occasionally accepted as informal or genre-appropriate, the singers' production of *was/were* variation (i.e., "Thought you was lookin' through me") was overwhelmingly labeled as 'incorrect' rather than merely nonstandard. Participants described these forms using terms such as 'incorrect,' 'cutting,' 'Black slang', 'street style,' and 'rap-ish.' Some of these comments reflect the enregisterment of this feature as being specifically associated with Black speakers, however other comments did not explicitly acknowledge its racialized indexicality, instead focusing solely on its nonstandardness, thus engaging in and reifying covert racialized language ideologies. This perception of inferiority weaponizes grammatical differences between AAE and MUSE by reinforcing racialized linguistic hierarchies through notions of standardness, even when race is not specifically mentioned.

## 5. Study 2: Matched-guise survey.

5.1. METHODS AND DESIGN. As this study specifically examines the perceptions of American audiences, a total of 157 U.S. born and raised adult participants were recruited from Prolific and compensated $3.25 to complete a 15-minute survey ($12/hour rate). All participants were between the ages of 18-40, reported English as their first language, had a 99% approval rating on Prolific, and did not have any vision or hearing impairments. Survey participants tended to be older than those in the focus groups, averaging at 31 years old. The majority of participants were White, heterosexual, and female, with the largest regional representation from the South.

The Qualtrics survey consisted of three stages designed to evaluate how auditory and visual cues shape perceptions of racial identity, two of which are described for the purposes of this study. In the introduction of the survey, participants were encouraged to wear headphones during the survey. In the first section, participants listened to each singer (BM, BF, WM, WF) perform one lyric each in one of the two guises. Then, participants rated each singer on a 6-point semantic differential scale across six dimensions: confidence, experience, smoothness, authenticity, sexiness, and trendiness. The selected attributes were selected based on a qualitative analysis of the focus group comments, which identified the most salient factors participants considered when evaluating the singers. This approach diverges from traditional sociophonetics matched-guise studies, which often rely on dimensions such as intelligence and formality. While such attributes are frequently tied to competence and solidarity in linguistic perception studies, they were found to be less relevant in the context of pop music. To ensure that participants did not default to neutral responses, I opted for closed, forced choice questions, which prevent respondents from expressing ambiguity in their evaluations.

Participants were also asked to indicate the perceived race of the singer through a multiple-choice question with the options: White, Mixed, Black, and Another Race. If participants selected 'Mixed' or 'Another Race,' a free-response box appeared, prompting them to specify what racial background(s) they imagined for the singer. Additionally, all participants, regardless of their initial choice, were required to indicate their confidence in their racial assessment of the

singer using a sliding scale, expressed as a percentage. Finally, an optional text box was provided where participants could note additional observations about the singers.

In the second stage, participants were presented with audiovisual stimuli, pairing each singer's audio recording with a corresponding 'headshot' of the singer. Participants then re-rated the singers using the same 6-point semantic differential scales from the first stage in order to determine whether the added visual information influenced their initial perceptions. An optional free-response box was included to allow participants to elaborate on these evaluations. Independent t-tests and linear regressions were used to determine whether there were significant differences between the ratings of the singers across the auditory and audiovisual conditions.

To systematically examine the perception of AAE features in pop music, I designed 78 distinct survey versions that systematically balanced linguistic variables and guises sung by each singer. These surveys explored the same linguistic variables as the focus groups as well as combinations of these features in a single lyric to test whether feature density heightened participants' awareness of linguistic variation. Effects of feature density were analyzed using linear regression, while a mixed-effects logistic regression examined the roles of guise, feature type, and repeated exposure to specific features across plays in shaping perceptual judgments.

### 5.2. Results.

5.2.1 Racial identity classification. Compared to the focus groups, participants' ability to correctly identify the singers' racial identities declined substantially in the survey (Table 2). In the focus groups, accuracy was relatively high for the Black singers (BM and BF: 86.1%) but notably lower for the White singers (WM: 52.7%, WF: 66.7%). In the survey, accuracy dropped across *all four singers* (BM: 63.1%, BF: 54.1%, WM: 64.3%, WF: 58.0%). Interestingly, while WM had the lowest accuracy in the focus groups, he had the highest accuracy in the survey (64.3%), albeit at low levels. In terms of self-reported confidence ratings, participants' overall confidence hovered around 50%, with slightly higher confidence for the Black singers than the White singers. These low confidence ratings may also be to participants' discomfort with making racial judgments, as evident by some of the comments left by participants (e.g., P134: "Gawd I feel uncomfortable saying that bc I literally have no idea but y'all asked").

| **Black Male** | | | | **Black Female** | | | |
|---|---|---|---|---|---|---|---|
| **Guise** | ***N*** | **% Correct** | **Average Confidence** | **Guise** | ***N*** | **% Correct** | **Average Confidence** |
| **AAE** | 78 | 67.9% | 56.6% | **AAE** | 78 | 52.6% | 53.8% |
| **MUSE** | 79 | 58.4% | 54.4% | **MUSE** | 79 | 55.7% | 54.3% |
| **White Male** | | | | **White Female** | | | |
| **Guise** | ***N*** | **% Correct** | **Average Confidence** | **Guise** | ***N*** | **% Correct** | **Average Confidence** |
| **AAE** | 79 | 58.2% | 49.4% | **AAE** | 79 | 53.2% | 50.2% |
| **MUSE** | 78 | 70.5% | 54.1% | **MUSE** | 78 | 62.8% | 50.5% |

Table 2. Racial identity guess accuracy by singer, guise, & confidence ratings (survey)

The lower accuracy observed in the survey compared to the focus groups likely reflects differences in the amount of linguistic input across the two methods. In the focus groups, participants heard two full mini-songs performed by each singer in either their AAE or MUSE guise, thus providing a more substantial basis from which to assess the singer's racial identity. In contrast, the survey provided listeners with only a single lyric (1-3 seconds long). This pattern

aligns with Purnell et al. (1999), who tested how the length of linguistic input affected accuracy in racial identification. Their study found that when participants heard only a single word stimulus ("Hello"), their accuracy in correctly identifying a speaker's racial identity was around 70%. However, when exposed to full-sentence stimuli, accuracy improved to 77% or higher across all guises, which similarly suggests that shorter linguistic input increases misclassification errors.

5.2.2 NUMBER AND TYPE OF AAE FEATURE. To assess whether specific AAE features influenced racial classification, I conducted logistic regressions for each singer, testing both the *cumulative* effect of multiple AAE features within a lyric and the *individual* effect of specific features. The results indicate that the presence of AAE features in song lyrics influenced racial classification accuracy differently across singers. No significant effects emerged for WF. For the Black singers, the more AAE features a lyric contained, the more likely participants were to correctly identify them as Black (BM: Log Odds: 1.815, Odds Ratio: 6.14, $p = 0.0182$; BF: Log Odds: 1.4403, Odds Ratio: 4.22, $p = 0.0187$). In contrast, for the WM, the presence of more AAE features in a lyric reduced classification accuracy, increasing the likelihood that he was misidentified as Black (Log Odds: -1.1125, Odds Ratio: 0.33, $p = 0.0719$). Each additional AAE feature in a lyric decreased the likelihood of correctly identifying WM as White by approximately 67%, suggesting that listeners relied on these racialized cues to construct an image of a Black singer.

The type of AAE feature itself also influenced classification accuracy. For BM, copula omission (Log Odds: -1.853, $p = 0.0112$) and demonstrative variation (Log Odds: -1.592, $p = 0.0428$) puzzlingly reduced the likelihood of correct classification. No individual AAE feature significantly influenced accuracy for BF or WF. However, for WM, certain AAE features increased the likelihood of correct racial identification, specifically, *gon'* (Log Odds = 1.75, Odds Ratio = 5.74, $p = 0.006$), demonstrative pronoun variation (Log Odds = 1.75, Odds Ratio = 5.78, $p = 0.012$), and remote past BEEN (Log Odds = 1.54, Odds Ratio = 4.68, $p = 0.025$). However, others, including 3rd person verbal variation, habitual BE, fortition, and copula omission, contributed to greater misclassification. These findings are explored further in the discussion.

5.2.3 AUDITORY AND AUDIOVISUAL EVALUATIONS. Another aim of this study was to investigate how evaluation condition (auditory vs. audiovisual) and guise (MUSE vs. AAE) influenced listener ratings of singer attributes, with a particular focus on whether or not the singer's revealed racial identity (Black or White) in the audiovisual condition significantly influenced these ratings. Overall, the inclusion of visual cues had minimal effect on singer ratings. However, small but significant differences emerged related to race and gender: Black singers received slightly higher ratings (M = 0.1188) when visual information was present compared to White singers (M = -0.006) ($t(42.67) = 3.79$, $p < 0.001$), and male singers (M = 4.30) were rated higher than female singers (M = 3.96) ($t(46) = -2.92$, $p = 0.005$).

Next, I examined how specific attributes varied in their ratings across singers. The results indicate that certain traits, notably "Confident" and "Experienced," received consistently high ratings across all singers, whereas attributes such as "Trendy" and "Sexy" generally received lower ratings. Statistical comparisons between the auditory and audiovisual conditions revealed few significant differences. Only some condition- and guise-specific effects emerged (see Figure 2; e.g., increases for BF on "Experienced" ($p = 6.73e\text{-}05$) and "Smooth" ($p = 0.001$), and a decrease for WM on "Trendy" ($p = 0.05$)). WF showed no significant differences between conditions. A linear regression confirmed these findings, with singer identity, condition, and linguistic guise as significant predictors of ratings ($R^2 = 0.888$, $p < 0.001$). BF received the highest overall ratings, followed by BM ($\beta = -0.10$, $p = 0.038$), WF ($\beta = -0.38$, $p < 0.001$), and WM ($\beta = -0.93$, $p < 0.001$). In terms of the condition, the addition of visual cues yielded a non-significant

increase in ratings (β = 0.056, $p$ = 0.093). Lastly, performances in AAE were rated significantly higher than those in MUSE (β = 0.172, $p$ < 0.001), suggesting an overarching preference for hearing AAE features in pop music contexts, regardless of the singer's racial identity.

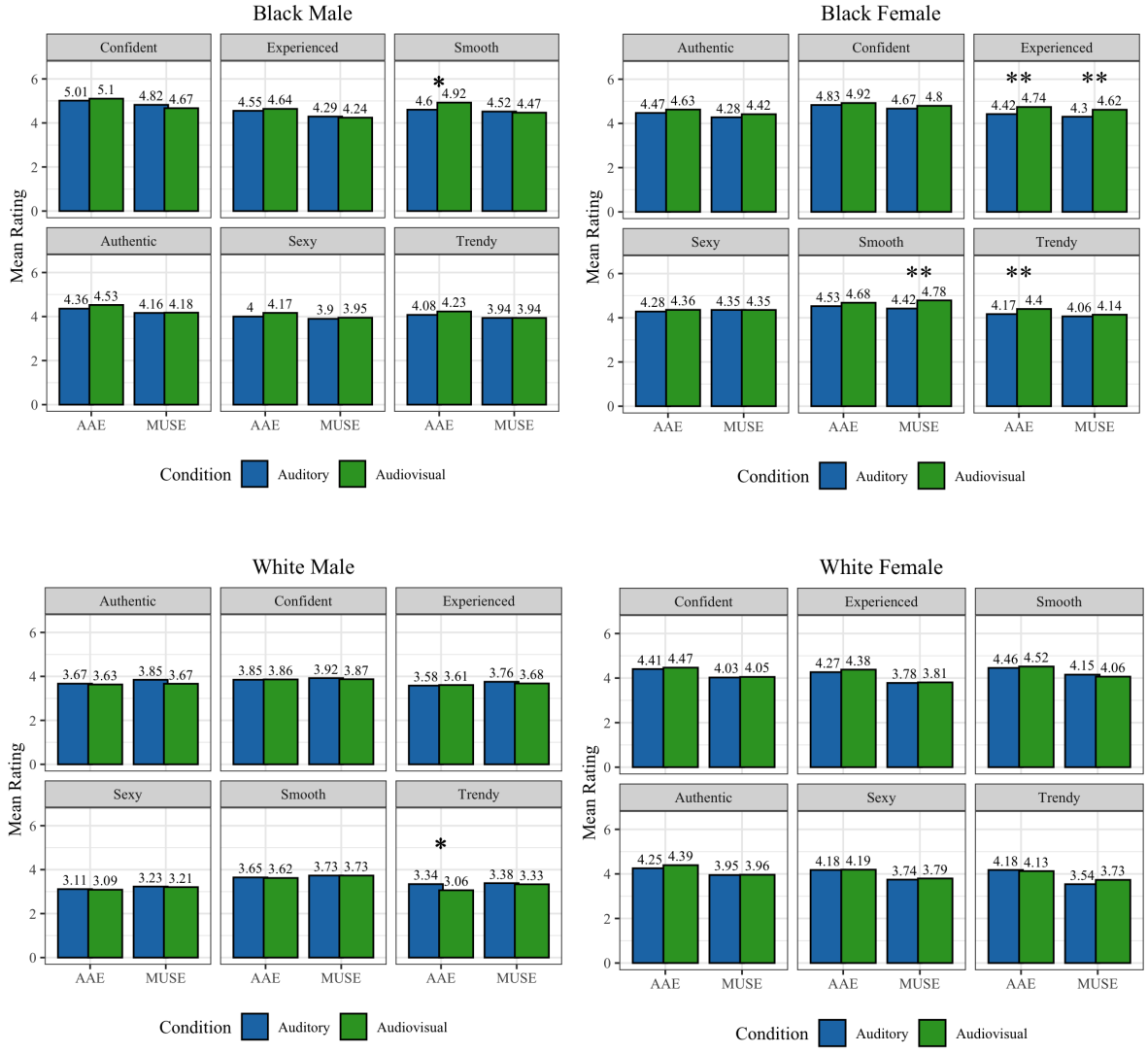


Figure 2. Mean ratings of singer attributes across conditions and by guise (survey)[2]

**6. Discussion.** This study had two primary aims. First, to examine how American audiences perceive and evaluate racialized linguistic cues when listening to music. Second, to explore how an artist's phenotypic appearance influences listener evaluations across auditory-only and audiovisual conditions, assessing whether expectancy violations arise from a mismatch between visual and linguistic cues based on the listener's preexisting social knowledge. The comparison of evaluations across auditory and audiovisual conditions revealed only a handful of marginally significant effects, suggesting that, in general, the addition of visual information did not substantially alter overall perceptions of the artists. Notably, for the White singers performing in the AAE guise, there were no significantly negative consequences in how they were evaluated across

[2] Attributes showing significant differences are indicated by asterisks (*p < .05; ** p < .01).

conditions, indicating that their use of AAE may not have triggered cognitive dissonance in this musical context. To better understand why added visual information may not have significantly impacted these evaluations, we turn to the first key question of this study: how did participants classify the racial identity of each singer in both the focus groups and the survey, and how did guise and specific AAE features contribute to this classification task?

As shown in the results, listeners were generally more accurate in identifying Black singers than White singers in the focus groups, while accuracy dropped across all singers in the survey context, likely due to the poverty of linguistic input. Furthermore, guise played an important role in racial identification accuracy, as, overall, singers were significantly more likely to be correctly identified when performing in an AAE guise. This pattern likely stems from the fact that the Black singers were bidialectal, making their performances in both AAE and MUSE likely sound more 'natural' and 'authentic' to listeners. In contrast, the White singers, as monodialectal speakers of MUSE, may have sounded feigned or forced to listeners when performing in the AAE guide. Focus group participants frequently remarked on these unnatural performances:

> **P6**: "Some of the words don't sound natural to him, because I don't think he uses it."
> **P7**: "It is kind of like, weird to hear, like, a White singer, like, like use slang- like a lot."
> **P14**: "So I'm like, "your voice does not match this at all." Like, "you can't like, try- me, I'm hearing that." And I'm like, "you're trying to use Black slang without being Black." Like, that's what I thought when I heard that."

Such reactions highlight listeners' sensitivity to phonetic inconsistencies associated with a lack of fluency in AAE, as well as their attunement to the ideological implications of such unlicensed forms of linguistic crossing in the lyrics. Thus, rather than creating ambiguity in racial perception—where listeners might have evaluated the White singers as 'sounding Black' due to their use of AAE—the AAE guise instead at times heightened awareness of the singers' Whiteness. The expectancy violations triggered by this inconsistent linguistic performance may have drawn attention to the perceived overt racial incongruence, thus making their Whiteness more salient rather than masking it and leading to higher accuracy in racial classification.

However, interestingly, the survey results reveal a different pattern for WM in the MUSE guise, where the absence of distinct racial markers likely led to greater misclassification. This finding aligns with Trechter and Bucholtz's (2001) assertion that MUSE is often perceived as 'neutral' or unmarked given its proximity to the unmarked White identity, thus making it difficult for listeners to assign a clear racial identity. This interpretation is consistent with previous perceptual studies (e.g., Thomas & Reaser 2004), who found that listeners struggled to identify the racial identity of speakers who produced fewer prototypical features, instead having to rely on subtle acoustic cues embedded in the speech. Thus, while the AAE guise served as a more robust socioindexical cue strongly associated with Blackness, MUSE did not equally index Whiteness to the same degree, resulting in greater racial ambiguity and lower classification accuracy for White singers, and particularly WM, in conditions with limited auditory input.

Participants' accuracy and evaluations of the singers also varied considerably depending on the number and type of AAE features present. For WM in the auditory-only portion of the survey, an increased presence of AAE features generally led listeners to incorrectly identify him as Black, indicating that multiple co-occurring AAE features significantly shifted listeners' racial perceptions. This finding sheds light on why, at times, some White singers can perceptually 'pass' as Black in musical contexts, where visual cues are not as readily available (e.g., Iggy Azalea, see Eberhardt & Freeman 2015). In such cases, the reliance on racialized linguistic markers alone can obscure a singer's true racial identity. Additionally, certain AAE features,

such as 3rd person present-tense verbal variation, habitual *BE*, fortition, and copula omission, led to a higher misclassification rate for WM. While further research is needed to fully understand the nuances in how these features are processed and interpreted by listeners, these findings may suggest that some AAE features index Blackness more strongly than others.

Yet, importantly, not all AAE features triggered the same racialized associations. For example, the fortition of /z/ before nasals (i.e., *business,* <bɪdnɪs>) was nearly always strikingly marked for listeners, who strongly disfavored its use by White singers as it was interpreted as particularly unnatural. This expectancy violation thus led to a more accurate classification of WM as White through their activation of raciolinguistic ideologies. Similarly, other AAE features, such as future marker *gon',* demonstrative variation, and remote past BEEN, were also consistently perceived as racially incongruous when produced by WM, again increasing racial classification accuracy. Such reactions suggest that certain features may function as even *stronger* indexes, not merely as linguistic markers of Black identity, but possibly serving as highly salient cues stereotypically associated with Black speech. This finding notably highlights listeners' attunement to mismatches between linguistic features associated with one racial group and the singer's perceived racial identity.

However, overarchingly, listeners found the White singers' production of AAE features to be permissible or even *preferred* within this musical context, as all singers were statistically rated higher while performing in the AAE guise. This effect suggests that certain features may be less indexical of race when embedded in a pop song. For example, some listeners described AAE features such the future marked *gon'* and copula omission as more rhythmically fluid or stylistically natural in the context of the music (e.g., P13: "The first version [AAE guise], I feel like, was so much better than the other ones. However, it was mainly because it's like… it just flows better ...I strongly preferred the first version [AAE] as opposed to the second [MUSE].").

These mixed findings further suggest that perceptions in this sociolinguistic context are highly listener-dependent. In other words, while some listeners dispreferred certain AAE features when sung by a White singer, others implicitly accepted them without strong racial associations, likely interpreting them primarily as stylistic choices or phonetic reductions rather than inherently racialized markers. This interpretation also aligns with the findings comparing auditory and audiovisual evaluations of each artist, in which the White artists performing in the AAE guise were not rated significantly lower when their racial identity was made visually apparent. Such indiscriminate acceptance likely reflects how pop music serves as an active site of indexical bleaching, representing an institutional context wherein racial associations of specific AAE features are neutralized by White performances. The perception that AAE sounded normal or 'natural' across all singers, regardless of race, was a recurring theme in the focus groups:

> **P6**: "I think it sounded more natural for her both ways."
> **P5**: "For her, it was harder to tell the distinctions. Like it just seemed very natural for her… if this was all together in one song. I thought it would be purposeful."

These responses, among others, suggest that, for some listeners, *genre* likely functions as the primary perceptual filter, reducing all linguistic variation to simply 'pop lyrics' and overriding racialized ideologies tied to enregistered language varieties and speakers' racial identities, thus agreeing with the findings of Gibson (2024) and Squires (2019).

Finally, this fluidity of AAE in pop music is best contextualized and understood within the broader commercial framework of the American music industry, which has historically absorbed, repackaged, and mainstreamed Black cultural and linguistic forms while simultaneously reinforcing racialized boundaries outside of performance. As pop music—a genre designed and

marketed for mass appeal, borne from the incorporation of various non-White styles and genres, notably Hip Hop—remains deeply embedded in American consumer culture, AAE thus is reproduced and absorbed in a highly paradoxical space. It is, at times, lauded for its authenticity, rendered invisible in musical contexts, or stigmatized for its implicit or explicit association with Blackness. All of these interpretations and evaluations are thus ultimately dependent on the expectations and ideologies of the listener and consumer of such mainstream cultural products.

**7. Conclusion.** This study advances our understanding of how American audiences perceive racial identity and linguistic authenticity in pop music. The findings reveal that evaluations of racialized linguistic features like AAE are highly context- and listener-dependent rather than uniform. Based on the results of the focus groups and surveys, listeners consistently preferred the use of AAE within the pop music context, and even when explicit visual cues of racial identity were made relevant, listeners generally did not negatively evaluate White singers performing in the AAE guise, suggesting that in pop music, certain AAE features are stylistically preferred over their MUSE counterparts due to their normalized use in the genre. However, this perception was by no means shared by all listeners, as some participants formed their evaluations of the singers primarily in regards to overarching racialized linguistic ideologies, viewing the appropriation of Black speech features as inherently problematic, regardless of genre conventions.

These findings add to the broader discussion surrounding the commodification of Black art in the American music industry, where Black linguistic and cultural practices are systematically repackaged through White performers. And over time, these relentless processes of repetition dilute, or 'bleach,' the original racial indexicality of certain AAE features, making them thus more susceptible to further exploitation and more readily accepted in performative contexts, like pop music. Through sustained exposure to this type of subtle, industry-driven appropriation, American audiences are conditioned to expect to hear AAE in mainstream music, further normalizing its use by both Black and White singers. Ultimately, this research highlights the crucial role of genre and racial ideologies in shaping listener expectations and perceptions while listening to music, and suggests that linguistic appropriation in music remains a contested site of ideological negotiation, where artist authenticity, genre-appropriate conventions, and the commercially-driven imperatives of the music industry all coalesce, each fundamentally shaping the way that audiences consume, interpret, and anticipate musical performances.

## References


Alim, H. Samy, Awad Ibrahim, & Alastair Pennycook (eds.). 2008. *Global Linguistic Flows: Hip Hop Cultures, Youth Identities, and the Politics of Language*. New York: Routledge.

Brown, Tamara L. & Baruti N. Kopano (eds.). 2014. *Soul thieves: The appropriation and misrepresentation of African American popular culture*. New York: Palgrave Macmillan.

Bucholtz, Mary, & Qiuana Lopez. 2011. Performing blackness, forming whiteness: Linguistic minstrelsy in Hollywood film. *Journal of Sociolinguistics* 15(5), 680–706. https://doi.org/10.1111/j.1467-9841.2011.00513.x

Cherid, Maha I. 2021. "Ain't got enough money to pay me respect": Blackfishing, cultural appropriation, and the commodification of Blackness. *Cultural Studies, Critical Methodologies* 21(5). 359–364. https://doi.org/10.1177/15327086211029357

Craig, Belinda M., Nigel Chen & Ottmar V. Lipp. 2022. Featural vs. Holistic processing and visual sampling in the influence of social category cues on emotion recognition. *Cognition and Emotion* 36(5). 855–875. https://doi.org/10.1080/02699931.2022.2057442

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(3). 303–327. https://doi.org/10.1111/josl.12128
Eckert, Penelope. 2008. Variation and the indexical field. *Journal of Sociolinguistics* 12(4). 453–476. https://doi.org/10.1111/j.1467-9841.2008.00374.x
Gamble, Steven (2024). *Digital Flows: Online Hip Hop Music and Culture*. Oxford: Oxford University Press.
Gibson, Andy. 2011. Flight of the Conchords: Recontextualizing the voices of popular culture. *Journal of Sociolinguistics* 15(5). 603–626. https://doi.org/10.1111/j.1467-9841.2011.00515.x
Gibson, Andy. 2024. Pop Song English as a supralocal norm. *Language in Society* 53(3). 471–498. https://doi.org/10.1017/S0047404523000131
Hansen, Karolina, Melanie C. Steffens, Tamara Rakić & Holger Wiese. 2017. Competent and warm? How mismatching appearance and accent influence first impressions. *Social Cognitive Affective Neuroscience* 12(3). 507–515. https://doi.org/10.1093/scan/nsw148
Hay, Jennifer, & Katie Drager. 2010. Stuffed toys and speech perception. *Linguistics 48*(4). 865–892. https://doi.org/10.1515/ling.2010.027
Hebdige, Dick. 1979. *Subculture: The meaning of style*. New York: Methuen.
Ma, Debbie S., Joshue Correll & Bernd Wittenbrink. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods* 47(4). 1122–1135. https://doi.org/10.3758/s13428-014-0532-5
Niedzielski, Nancy. 1999. The Effect of Social Information on the Perception of Sociolinguistic Variables. *Journal of Language and Social Psychology* 18(1). 62–85. https://doi.org/10.1177/0261927X99018001005
Obasogie, Osagie K. 2010. Do blind people see race? Social, legal, and theoretical considerations. *Law & Society Review* 44(3–4). 585–616. https://doi.org/10.1111/j.1540-5893.2010.00417.x
Purnell, Thomas, William Idsardi & John Baugh. 1999. Perceptual and phonetic experiments on American English dialect identification. *Journal of Language and Social Psychology* 18(1). 10–30. https://doi.org/10.1177/0261927X99018001002
Squires, Lauren. 2019. Genre and linguistic expectation shift: Evidence from pop song lyrics. *Language in Society* 48(1). 1–30. https://doi.org/10.1017/S0047404518001112
Squires, Lauren. 2020. Indexical beaching. *The International Encyclopedia of Linguistic Anthropology.* 1–4. https://doi.org/10.1002/9781118786093.iela0455
Stein, Timo, Albert End & Philipp Sterzer. 2014. Own-race and own-age biases facilitate visual awareness of faces under interocular suppression. *Frontiers in Human Neuroscience 8*. 582.
Thomas, Erik R., & Jeffrey Reaser. 2004. Delimiting perceptual cues used for the ethnic labeling of African American and European American voices. *Journal of Sociolinguistics* 8(1). 54–87. https://doi.org/10.1111/j.1467-9841.2004.00251.x
Trudgill, Peter. 1997. Acts of conflicting identity: The sociolinguistics of British pop-song pronunciation. In Nikolas Coupland & Adam Jaworski (eds.), *Sociolinguistics,* 251–265. London: Palgrave.
Walton, Julie H., & Robert F. Orlikoff. 1994. Speaker race identification from acoustic cues in the vocal signal. *Journal of Speech and Hearing Research* 37(4). 738–745. https://doi.org/10.1044/jshr.3704.738