THE INFLUENCE OF REGIONAL DIALECT VARIATION ON RACE CATEGORIZATION

Tessa Bent and Yolanda Holt

Indiana University; East Carolina University
tbent@indiana.edu; holty@ecu.edu

ABSTRACT

Socio-ethnic and regional group membership influence segmental and suprasegmental talker characteristics. Listeners can use acoustic-phonetic variation to identify a talker's socio-ethnic and regional dialect; often with above chance accuracy. However, the interaction between socio-ethnic identity and regional dialect in speech perception has rarely been evaluated. This work examined how regional dialect variation influences socio-ethnic perceptions of race within the United States. Black and White talkers from two distant North Carolina communities produced /hVd/ words. Listeners unfamiliar with the dialect completed a two-alternative forced-choice race identification task. An interaction was observed between talker socio-ethnic racial identity and regional dialect. Listeners' identification of White talkers was well above chance for both communities. Listeners' identification of Eastern North Carolina Black talkers, participants of the African American Vowel Shift, was above chance; but identification of Black Western North Carolina talkers, participants of the Southern Vowel Shift, was at chance.

Keywords: Speech perception, regional dialect, race.

1. INTRODUCTION

Both segmental and suprasegmental characteristics of a talker's speech can be impacted by a range of socio-indexical factors, including the speaker's sex, gender, sexual orientation, age, regional dialect, socio-ethnic racial identity and peer group, among others [1-7]. Using speaker dependent acoustic-phonetic variation, listeners can categorize talkers by group membership with above chance accuracy [7-13]. However, the same segmental and suprasegmental elements can convey information about multiple socio-indexical categories. Thus, the interaction of these acoustic cues may influence listener perceptions of talker identity in complex ways [14]. We investigate how the interaction between two socio-indexical variables – socio-ethnic racial identity and regional dialect – influence listeners' race categorization accuracy.

In the United States, legal segregation and other social separations of Black from White Americans have resulted in distinct English dialects for the two socio-ethnic groups [15]. Listeners are sensitive to the associations between dialect features and a talker's socio-ethnic identity [16, 17]; therefore, within the U.S., listeners can identify the race of Black and White American talkers with above chance accuracy from stimuli that vary in length from passages [18] to sentences [19] to a single vowel [20]. Accuracy rates are lower, but still relatively robust, with synthetically-manipulated speech including speech played backward, time compressed, low- or high-pass filtered, or monotonized [10, 19, 21]. Although accuracy rates in these studies are above chance, most investigations show substantial differences between talkers in listeners' abilities to accurately identify talker race [e.g., 18], pointing to the fact that associations between talker race and acoustic-phonetic features are socially acquired. Furthermore, the socio-phonetic variation present in the speech of socio-ethnically defined groups, such as Black and White Americans, are not exclusive to the U.S. or to English [22, 23]. Through linguistic profiling, these systematic, perceptually indexed socio-phonetic variants can have educational, social, political, and economic consequences, even when the variation is not within the talkers' conscious manipulation [24].

Although Black English was originally identified as a singular pan American dialect, recent socio-linguistic analyses of Black American speech has highlighted the regional variation within this socio-ethnic dialect [25, 26]. The history of subsuming regional variation in Black American English under a nationally defined socio-ethnic racial category erased a socio-indexical component of these talkers' identities. There is substantial evidence that the acoustic-phonetic characteristics of both African American English and White American English are influenced by regional variation [5, 25, 27, 28]. Although variation can be present at any level of sound structure, many studies of regional variants within African American English have focused on vowel differences [27, 29, 30] as well as some work on /ɹ/ variants [25]. Thus, vowels can simultaneously convey information about socio-ethnic racial and regional affiliations. However, nearly all investigations of socio-ethnic racial identification in the North American context use speech gathered from talkers in a single geographic community or do not

specify the community and regional dialect of the talkers.

Two studies that have investigated the influence of regional variation on socio-ethnic racial identification suggest that a community's regional dialect variation impacts listeners' socio-ethnic racial categorization accuracy [10, 18]. For example, in [10] speech selections from interviews with Black and White talkers from two communities in North Carolina were presented to listeners, most of who were unfamiliar with the dialects, in a socio-ethnic racial categorization task. Speech samples included dialect features known to vary between the regional community dialects or across the socio-ethnic racial categories of Black and White talkers. The stimuli were unmodified, low-pass filtered, or monotonized and then presented to listeners to assess the contributions of vowel quality and prosodic variation on listeners' race categorization. Results suggested that Black talkers who do not use acoustic-phonetic characteristics typically associated with African American English are more often misidentified as White by listeners unfamiliar with the regional community dialect. The results from the synthetically-manipulated conditions indicated that both vowel quality and intonation are important cues for talker socio-ethnic racial identification. Although these studies highlight the interaction between regional and socio-ethnic dialect variation on listener perception, these studies [10, 18] had two limitations, which we address here. First, both studies used conversational speech samples, which may provide listeners with multiple segmental and suprasegmental cues to a talker's socio-ethnic racial identity, although the synthetic manipulations in [10] provided insight into the cues used by listeners at a broad level (i.e., intonation, segmental information). Specific information on how vowel productions influenced socio-ethnic racial categorization was not described. Another limiting factor is that the talkers from Hyde County, included in both studies, represent an isolated dialect community [31]

Here, we continue the investigation of how regional community dialect variation within North Carolina influences listeners' socio-ethnic racial categorization but shift the focus away from the isolated Hyde County dialect to two other North Carolina regions. The vowel characteristics of talkers from these two regions are influenced by community and socio-ethnic racial affiliated vowel shifts [27]. We assess whether these vowel characteristics influence listeners' socio-ethnic racial categorization by constraining the stimuli to /hVd/ words. Although cues beyond vowel differences may be present in these productions, the range of possible acoustic-phonetic cues available to listeners is limited. We hypothesize that talkers' adherence to two vowel shifts – the Southern Vowel Shift (SVS) and the African American Vowel Shift (AAS) [27, 28] – will impact listeners' race categorization accuracy. The SVS includes a realignment of the front tense vowels FLEECE and FACE with their lax counterparts KIT and DRESS. This realignment is preceded by PRICE vowel laxing and gliding in towards TRAP rather than up to KIT. A concomitant change is occurring in the purported AAS with raising of the lax front vowels in KIT, DRESS, and TRAP, coincident with fronting of the LOT vowel. A separate change, fronting of the GOOSE and GOAT vowels, is widespread in English dialects but not universally used by Black Americans [7, 28].

2. METHOD

2.1. Stimuli

There were 240 stimuli from 24 talkers' productions of 10 /hVd/ words: heed, hid, head, had, hayed, hide, hood, howed, hoyed, and who'd. Twelve talkers were from Eastern North Carolina (Pitt County representing the Coastal Plain dialect) and twelve were from Western North Carolina (Iredell County representing the Piedmont dialect), two geographically distant regions separated by 370 km. Within each dialect grouping, half of the talkers self-identified as Black and half as White. Each race and regional dialect group had equal numbers of male and female talkers.

2.2. Listeners

Listeners (n = 44) were current Indiana residents, 18 males and 26 females, with an average age of 21 years (range = 18 - 30). Listeners identified primarily as not Hispanic or Latino (n = 39) with three identifying as Hispanic or Latino and two choosing not to respond. Participants indicated their race as White (n = 31), Black or African American (n = 7), multiple races (n = 3), Asian American (n = 2), or other (n = 1). Most listeners indicated their U.S. regional dialect as either Midland (n = 23) or Northern (n = 13). Listeners rated their exposure to U.S. dialects on a scale from 1 - 5 (1 = no exposure or only brief casual exposure; 5 = daily at home exposure). Listeners rated the ambient dialect (Midland American English) as the one to which they were most frequently exposed (M = 4.5; range = 3 - 5). Many listeners also had some exposure to the Southern dialect (M = 2.7; range = 1 - 5).

2.3. Procedure

Testing was conducted in a sound-attenuated booth. Listeners were presented with one word per trial and

indicated the talker's race as Black or White by clicking on the corresponding button on the computer screen. Stimuli were presented in two randomized blocks of 240 trials each. Listeners heard each stimulus twice during the experiment, once in each block. Each word was presented binaurally over Sennheiser HD208 Pro headphones at approximately 68 dB SPL. Stimulus presentation and response recording was automatically controlled by a program written in PsychoPy [32] on a Mac Mini.

2.4. Analysis

Listeners' responses were averaged across words and across the three talkers representing each dialect region, self-reported race, and self-reported sex group (e.g., Eastern North Carolina Black women). Listener responses that corresponded to the talker's self-identified race (i.e., talker's race regardless of region) were scored as correct. Percent correct scores were converted to rationalized arcsine units (RAU) to facilitate meaningful comparisons across the entire range of the scale [33].

3. RESULTS

Race categorization accuracy scores in RAU were analysed with a repeated-measures ANOVA in which there were three within-subjects variables: race (Black, White), dialect (Eastern North Carolina, Western North Carolina), and sex (female, male). All main effects were significant. Listeners were more accurate at identifying the race of talkers from Eastern North Carolina than Western North Carolina (86 vs. 70 RAU) [$F(1,43) = 192.67$, $p < 0.001$, $\eta_p^2 = 0.818$], White talkers than Black talkers (89 vs. 68 RAU) [$F(1,43) = 40.63$, $p < 0.001$, $\eta_p^2 = 0.486$], and women than men (81 vs. 76 RAU) [$F(1,43) = 14.64$, $p < 0.001$, $\eta_p^2 = 0.254$]. All two-way interactions were also significant. The interaction between talker race and dialect arose because the race of White talkers from both dialect regions was categorized with high accuracy (87 and 90 RAU). However, race categorization accuracy of Black talkers from Eastern North Carolina was much higher than Black talkers from Western North Carolina (83 vs. 53 RAU) [$F(1,43) = 97.31$, $p < 0.001$, $\eta_p^2 = .694$] (Figure 1). The interaction between talker sex and dialect arose because listeners showed similar performance for male and female talkers from Western North Carolina (71 vs. 70 RAU) but were more accurate in identifying the race of women than men for the Eastern North Carolina talkers (91 vs. 81 RAU) [$F(1,43) = 36.46$, $p < 0.001$, $\eta_p^2 = 0.459$]. The final two-way interaction between talker race and sex arose because there was no difference in race categorization accuracy between Black men and women (68 RAU for both), but White women's race was categorized more accurately than men's (93 vs. 84 RAU) [$F(1,43) = 5.17$, $p = 0.028$, $\eta_p^2 = 0.107$]. The three-way interaction was not significant.

Figure 1: Race categorization accuracy in RAU for talkers from Eastern (left) and Western North Carolina (right) with the Black talkers indicated by the solid line with filled circles and the White talkers by the dotted line with the open squares. Error bars represent standard deviation.

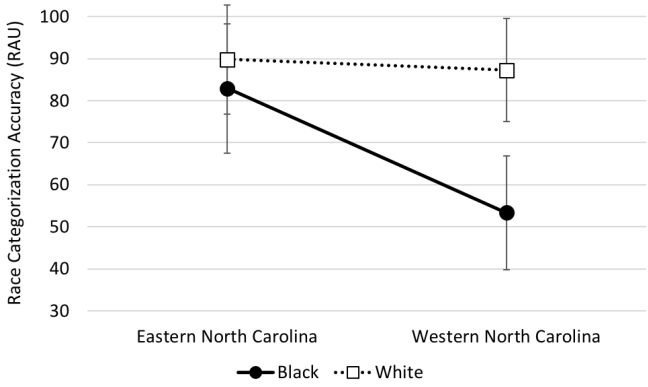
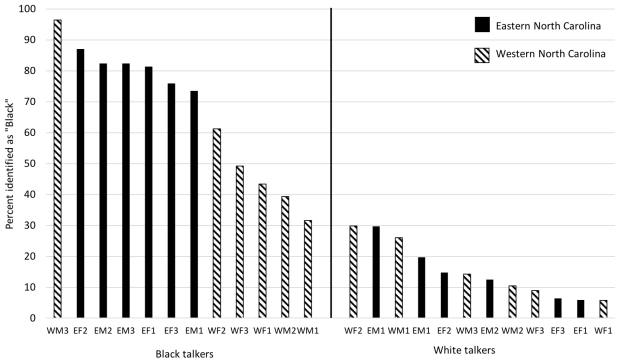


Figure 2: Percent of stimuli categorized as "Black" for the 24 talkers. Black talkers are on the left and White talkers are on the right. Eastern North Carolina talkers are shown with black bars and Western North Carolina talkers with striped bars. Talker codes represent dialect (E for Eastern and W for Western), sex (M for male and F for female), and a unique number.



Scores for the 24 talkers were converted to the percent of trials in which listeners categorized the talker as Black (Figure 2). As with the group results, there is a clear influence of dialect on race categorization of the Black talkers. All Black talkers from Eastern North Carolina are identified as Black more frequently than Black talkers from Western North Carolina, except for one Western North Carolina talker who was most often categorized as Black overall. Additionally, there was no overlap in race categorization scores between the Black and White talkers. The highest rate

of categorization as Black for a White talker was 30%. In contrast, some Black talkers, particularly those from Western North Carolina, were identified more often as White than Black. Although there are clear influences of regional variation on race identification, there are also substantial individual differences across talkers.

4. DISCUSSION

Listeners' race categorization accuracy was above chance for the word-length stimuli used here (76% correct overall). However, accuracy rates varied across groups in several ways. First, listeners were more accurate with White talkers than Black talkers. Furthermore, race categorization accuracy was influenced by talker dialect, such that the race of White talkers from both dialect regions was identified with high accuracy, but the race of Black talkers was identified with higher accuracy on average for talkers from Eastern North Carolina than Western North Carolina. Race categorization accuracy for Black talkers from Western North Carolina was near chance on average. Previous work has suggested that regional dialect variation can influence race categorization accuracy [10, 18], but both previous studies employed conversational speech samples that were more varied and longer than the stimuli used in the current study. Because our stimuli were isolated /hVd/ words, the acoustic-phonetic cues available to listeners during their categorization judgements were highly constrained. Due to the nature of these stimuli, one central cue available to listeners was spectral vowel differences. Therefore, talkers' adherence to two vowel shifts may have influenced listeners' categorization patterns. As shown in [27], Black talkers from Western North Carolina frequently produce a vowel shift that is associated with Southern White English (i.e., reversal of the DRESS and FACE vowels) and show less adherence to the African American Shift (i.e., of the front lax vowels, they only raise the TRAP vowel). In contrast, Black talkers from Eastern North Carolina do not participate in the Southern Vowel Shift [28] and show greater participation in the African American Shift. Therefore, the race of Black speakers who exhibit more phonetic variants associated with White American English and fewer associated with African American English may be more difficult to identify. Similar to previous studies, listeners are less accurate at identifying the race of Black talkers who produce phonetic and prosodic variants that are typical of White American English [10, 34, 35], although here the extent of adherence to the racially-affiliated variants is likely conditioned by regional dialect rather than solely by idiolect.

We have highlighted the influence of the regional and socio-cultural vowel shifts – the Southern Vowel Shift and the theorized African American Shift [7] – to explain the interaction between talker race and regional dialect observed here; however, it is possible that other cues may have shaped listeners' responses. For example, for the Western North Carolina male talker who was identified most frequently as Black, voice quality may have played a role. This talker had a lower fundamental frequency and a substantially creakier voice quality than other talkers in the study. Differences in voice quality (e.g., f0, jitter, shimmer) between Black and White talkers have been reported previously, including lower f0 for Black male talkers than White male talkers [20, 36]. Thus, although using /hVd/ words limited the information available to listeners, other acoustic-phonetic cues are present even in such short, highly constrained stimuli. In addition to differences in voice quality, variation in vowel duration and stop consonant features (e.g., whether the stop was released or had voicing during closure) could also play a role. For example, previous work found differential use of duration in vowels by Black and White talkers [2], a cue that may have influenced listeners' categorization judgements.

Naïve listeners were used to identify and delimit the specific cues they associated with Blackness or Southerness. Without substantial exposure to the two dialect communities, the listeners are relying on their stereotypes of Black and White talkers from the South to assign talkers to socio-ethnic categories. Their difficulty in accurately classifying Black speakers who use speech patterns more closely aligned with White Southern speech indicates that participation in the Southern Vowel Shift is not associated with African American English in the perceptual schema of Midwestern listeners.

Future analyses will assess how the spectral and temporal characteristics of individual talker's vowel productions as well as voice quality and available cues in the final stop consonant impacted race categorization. We also will test listeners who vary in their exposure to the specific dialects and assess the impact of listener race. We anticipate that listener experience with the dialects, particularly exposure to talkers from Western North Carolina, will increase race categorization accuracy, as listeners who have greater exposure to speakers from relevant socio-cultural groups are frequently more accurate in race/ethnicity categorization tasks [17, 35].

5. ACKNOWELDGEMENTS

We thank Jodene Jackson, Danielle Pacholski, and Jesse Sherman for their assistance with data collection.

6. REFERENCES

[1] Hay, J., Drager, K. 2007. Sociophonetics. *Annual Review of Anthropology* 36, 89-103.

[2] Holt, Y.F., Jacewicz, E., Fox, R. 2016. Temporal variation in African American English: The distinctive use of vowel duration. *Journal of Phonetics and Audiology* 2, 1-8.

[3] Clopper, C.G., Smiljanic, R. 2011. Effects of gender and regional dialect on prosodic patterns in American English. *J. Phon* 39, 237-245.

[4] Hunter, E.J., Ferguson, S.H., Newman, C.A. 2016. Listener estimations of talker age: A meta-analysis of the literature. *Logopedics Phoniatrics Vocology* 41, 101-105.

[5] Clopper, C.G., Pisoni, D.B., de Jong, K. 2005. Acoustic characteristics of the vowel systems of six regional varieties of American English. *J. Acoust. Soc. Am.* 118, 1661-1676.

[6] Thomas, E.R., Carter, P.M. 2006. Prosodic rhythm and African American english. *English World-Wide* 27, 331-355.

[7] Thomas, E.R. 2007. Phonological and phonetic characteristics of African American vernacular English. *Lang. and Ling. Compass* 1, 450-475.

[8] Bent, T., Holt, R.F. 2017. Representation of speech variability. *Wiley Interdisciplinary Reviews: Cognitive Science* 8, e1434.

[9] Munson, B., McDonald, E.C., Deboe, N.L., White, A.R. 2006. The acoustic and perceptual bases of judgments of women and men's sexual orientation from read speech. *J. Phon.* 34, 202-240.

[10] Thomas, E.R., Reaser, J. 2004. Delimiting perceptual cues used for the ethnic labeling of African American and European American voices. *Journal of Sociolinguistics* 8, 54-87.

[11] Gelfer, M.P., Mikos, V.A. 2005. The relative contributions of speaking fundamental frequency and formant frequencies to gender identification based on isolated vowels. *Journal of Voice* 19, 544-554.

[12] Ptacek, P.H., Sander, E.K. 1966. Age recognition from voice. *Journal of Speech and Hearing Research* 9, 273-277.

[13] Lass, N.J., Hughes, K.R., Bowyer, M.D., Waters, L.T., Bourne, V.T. 1976. Speaker sex identification from voiced, whispered, and filtered isolated vowels. *J. Acoust. Soc. Am.* 59, 675-678.

[14] Campbell-Kibler, K. 2011. Intersecting variables and perceived sexual orientation in men. *American Speech* 86, 52-68.

[15] Labov, W. 1970. The logic of nonstandard English, In. *Language and poverty*. Elsevier, 153-189.

[16] Newman, M., Wu, A. 2011. "Do you sound Asian when you speak English?" Racial identification and voice in Chinese and Korean Americans' English. *American Speech* 86, 152-178.

[17] Wong, P., Babel, M. 2017. Perceptual identification of talker ethnicity in Vancouver English. *Journal of Sociolinguistics* 21, 603-628.

[18] Wolfram, W. 2000. On the construction of vernacular dialect norms. *Chicago Linguistic Society*. Chicago, IL, 335-358.

[19] Lass, N.J., Mertz, P.J., Kimmel, K.L. 1978. Effect of temporal speech alterations on speaker race and sex identifications. *Lang Speech* 21, 279-290.

[20] Walton, J.H., Orlikoff, R.F. 1994. Speaker race identification from acoustic cues in the vocal signal. *Journal of Speech, Language, and Hearing Research* 37, 738-745.

[21] Lass, N.J., Almerino, C.A., Jordan, L.F., Walsh, J.M. 1980. The effect of filtered speech on speaker race and sex identifications. *J. Phon.* 8, 101-112.

[22] Stuart-Smith, J., Timmins, C., Tweedie, F. 2007. 'Talkin' Jockney'? Variation and change in Glaswegian accent. *Journal of Sociolinguistics* 11, 221-260.

[23] Drager, K. 2010. Sociophonetic variation in speech perception. *Lang. and Ling. Compass* 4, 473-480.

[24] Purnell, T., Idsardi, W., Baugh, J. 1999. Perceptual and phonetic experiments on American English dialect identification. *Journal of Language and Social Psychology* 18, 10-30.

[25] Hinton, L.N., Pollock, K.E. 2000. Regional variations in the phonological characteristics of African American Vernacular English. *World Englishes* 19, 59-71.

[26] Wolfram, W., Kohn, M.E. 2015. Regionality in the development of African American English, In: Lanehart, S., Editor. *The Oxford Handbook of African American Language*. Oxford: Oxford University Press, 140.

[27] Holt, Y.F. 2018. Mechanisms of vowel variation in African American English. *Journal of Speech, Language, and Hearing Research* 61, 197-209.

[28] Labov, W., Ash, S., Boberg, C. 2006. *Atlas of North American English*. New York: Mouton de Gruyter.

[29] Fridland, V., Bartlett, K. 2006. The social and linguistic conditioning of back vowel fronting across ethnic groups in Memphis, Tennessee. *English Language & Linguistics* 10, 1-22.

[30] Fridland, V. 2003. Network strength and the realization of the Southern Vowel Shift among African Americans in Memphis, Tennessee. *American Speech* 78, 3-30.

[31] Wolfram, W., Thomas, E. 2002. *The Development of African American English*. Oxford: John Wiley & Sons.

[32] Peirce, J.W. 2007. PsychoPy—psychophysics software in Python. *Journal of Neuroscience Methods* 162, 8-13.

[33] Studebaker, G.A. 1985. A rational arcsine transform. *Journal of Speech and Hearing Research* 28, 455-462.

[34] Perrachione, T.K., Chiao, J.Y., Wong, P.C.M. 2010. Asymmetric cultural effects on perceptual expertise underlie an own-race bias for voices. *Cognition* 114, 42-55.

[35] Foreman, C.G. 1999. Dialect identification from prosodic cues. *Proc. 14th ICPhS*. San Francisco, CA, 1237-1240.

[36] Hudson, A.I., Holbrook, A. 1981. A study of the frequency reading fundamental vocal of young black adults. *Journal of Speech, Language, and Hearing Research* 24, 197-201.