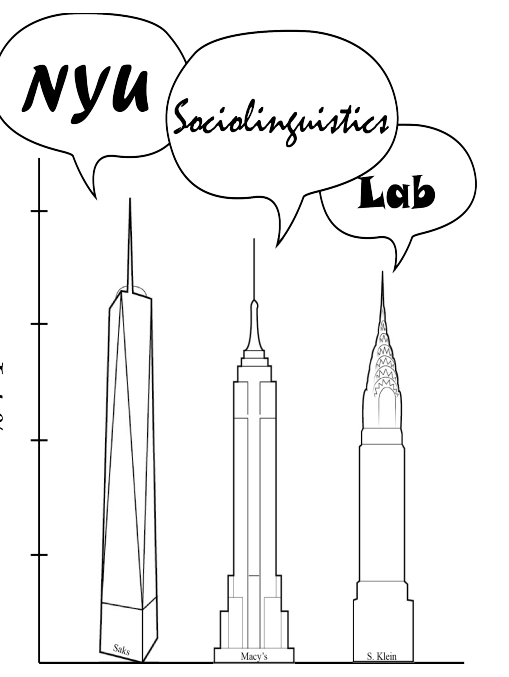




NYU

Revisiting linguistic profiling: Testing accuracy and the influence of prosodic variables on listener judgments of ethnicity

Nicole Holliday and Zachary Jagers, Department of Linguistics, New York University



1. Introduction

- How accurate are listeners at identifying the ethnicities of political figures? What linguistic cues provide the basis for these assessments?
- Purnell, Idsardi, and Baugh (1999) found listeners > 70% accurate at identifying dialects (Chicano English, African American English, Standard American English) based only on hearing a stimulus containing the token *hello*, as produced by Baugh in various guises, with implications for linguistic profiling.
- The current study examines whether a similar result is obtained with different tokens extracted from naturalistic speech of a variety of speakers.

Current Project

- Utterances of {city, many, people, women, issues, thank you} extracted from televised speeches by 7 male political figures from cities in NYC metro area of different stated ethnic identities: 3 black, 2 Latino, 2 white non-Hispanic.
- Tokens selected to avoid potential morphophonological variation and primarily capture how suprasegmental features contribute to listener identification of speaker ethnicity.
- Mechanical Turk experiment: The 42 tokens were compiled and presented in randomized order. Listeners (N=94) heard each token twice, then responded to forced multiple-choice question “What is the ethnicity of this speaker?”. Choices provided were “white”, “black”, and “Latino”. Listeners also responded to a short demographic questionnaire.



2. Analysis

- Previous research identifies suprasegmental features as significant phonetic correlates for ethnolect identification. Purnell et. al (1999) found Pitch Peak Ratio and Harmonics to Noise Ratio highest for black-identified voices.
- The following are variables identified for the current experiment. Statistical analyses were conducted on these features in Rbrul and R using LMER and Cluster and Regression Tree Analysis.

Selected Variables	Definition
HNR*	Ratio of periodic components to aperiodic components (Thomas 2011)
Intensity Ratio*	Change in intensity between two syllables (considering amplitude and frequency)
Intensity Avg*	Avg of amplitude as a function of frequency (in dB)
Shimmer Avg	Avg of local variation in amplitude between vocal fold vibrations (in dB) (Thomas 2011)
Jitter Avg*	Avg of local variation in F0 between vocal fold vibrations
Syllable Duration Ratio*	Measurement of length of 1 st syllable to length of 2 nd syllable
Pitch Peak Ratio (slope)*	Ratio of highest F0 to lowest F0 in sample (Hz)

3. Results

Listener Demographics

- One-way ANOVAs revealed no significant pattern with respect to accuracy or rating tokens as black by the following listener demographics:
 - Age (62% between 25-44)
 - Gender (49% female, 51% male)
 - Region
- We could not test for race because the sample was biased white (81%). This is comparable to general MTurk demographics (Berinsky et. al 2012).
- Individual listener also not significant, indicating no important idiosyncratic differences in accuracy or inclination to ID tokens as black.

Speaker Accuracy and Acoustics

- Across all speakers, ethnicity does not reliably predict accuracy.
- Cory Booker reliably and accurately identified as black, Bill de Blasio reliably and accurately identified as white, Andrew Cuomo misidentified as black with notable frequency.

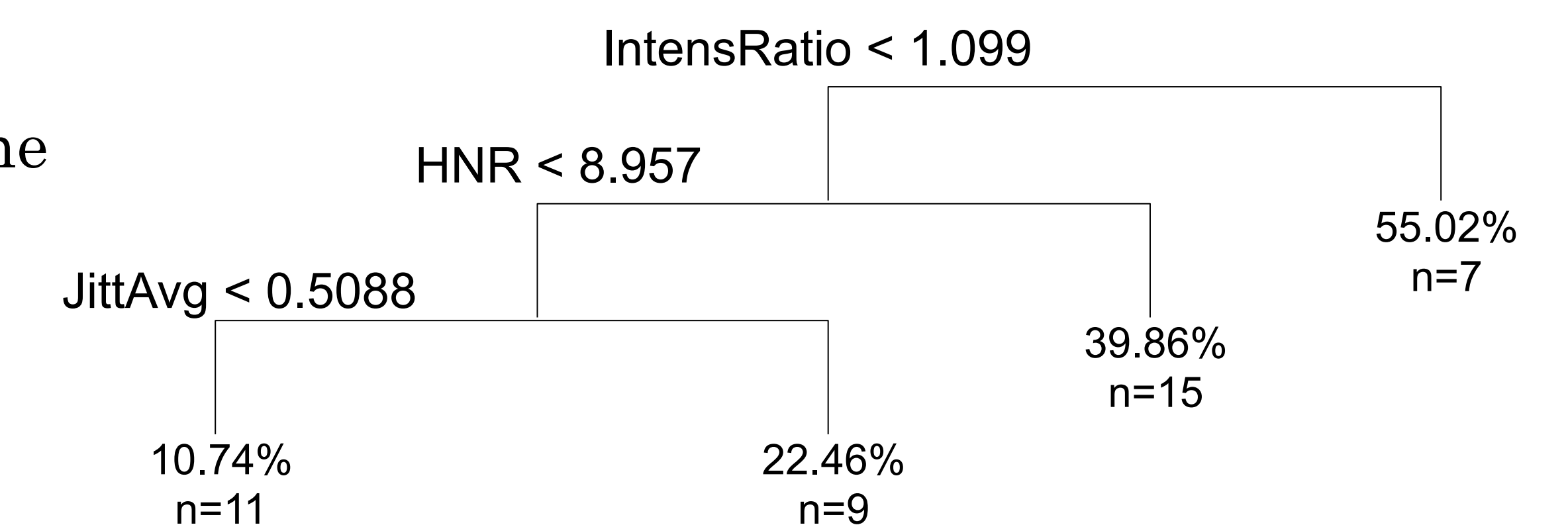
Speaker (Ethn)	Accuracy	Percent ID Black
Cory Booker (B)	71.63	71.63
Andrew Cuomo (W)	14.72	67.55
Bill de Blasio (W)	56.56	19.50

Rbrul Analysis

- Rbrul analysis tested ID black as a binary dependent variable with acoustic features and speaker as independent variables.
- A step-up/step-down analysis performed per response (N=3948) chose starred (*) variables in §2 table significant at p<.05.

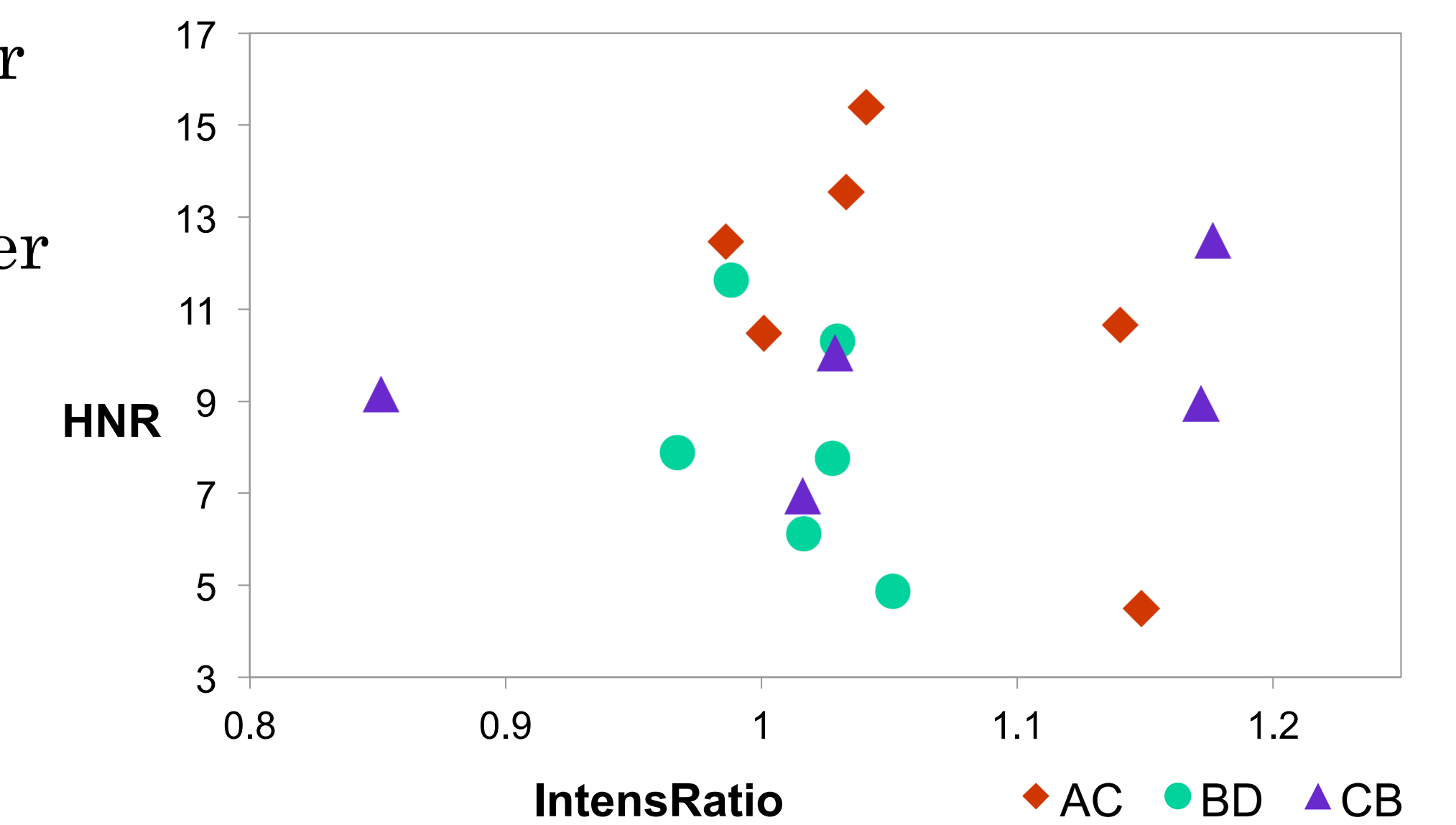
Cluster and Regression Tree Analysis

Cluster and Regression Tree (CART) Analysis revealed the following pattern (right) when tested against ID black percentage. This model works on group level, and is not predictable on token level.



CART Plot

- The plot (right) shows tokens for the 3 most reliably identified (or misidentified) speakers, along 2 strongest predictor variables.
- Cuomo patterns higher along HNR and Intensity Ratio than other white speakers, while de Blasio patterns lower on both features.
- Booker, the only reliably identified black speaker, varies more widely on these features but has tokens with the most co-occurrence of high HNR and high IntensRatio.



4. Discussion

- Results indicate that although listeners do not always make reliable judgments about speaker ethnicity based on one-word tokens, some suprasegmental features may correlate with a token or speaker being identified as black.
- In particular, Intensity Ratio, HNR, and Jitter Average seem most predictive of tokens' likelihood of being identified as black, even more reliably than individual or speaker ethnicity.
- Identifying the salient features that characterize ethnolects for different listeners may have special consequences for political figures, because they are often racialized explicitly and implicitly in public discourse with implications for their public personae.

Thank you to Gregory Guy, Sean Martin, Luiza Newlin-Lukowicz, and members of the NYU Sociolinguistics Lab for support and suggestions on this project. Also, thanks to all our politicians for their voices (especially Cory Booker and Andrew Cuomo).