You had me at "Hello": early racial dialect perception

William J. Idsardi[1]*, Mathias Scharinger[1], Philip J. Monahan[1,2], Thomas Purnell[3], John Baugh[4]

[1]Department of Linguistics, University of Maryland, College Park MD, 20742, USA.

[2]Basque Center on Cognition, Brain and Language, San Sebastián, 20009, Spain.

[3]Department of Linguistics, University of Wisconsin, Madison WI, 53706, USA.

[4]Department of Psychology and Program in African-American Studies, Washington University, St. Louis MO, 63130, USA.

*To whom correspondence should be addressed. Email: idsardi@umd.edu.

Racial discrimination, stereotyping and prejudice have been long-standing social problems within the United States – and are socio-political issues that continue to remain current even 45 years after the Civil Rights Act of 1964 *(1,2)*. Recent advances within the cognitive neurosciences have made it possible to begin to understand the brain bases of race perception, and promise to provide new insight into the biological and neurophysiological sources of racial discrimination. Current evidence from face perception indicates that early online brain responses are greater in magnitude to racial out-group members than to in-group members *(3,4)*. Given the demonstrable effects on early cortical processing, it has been claimed that the perception of race begins with the earliest perceptual stages of face recognition *(5)*. Visual cues, however, are not the only perceptual cues to race identification. Auditory information also allows listeners to

identify the race of a speaker (at least when dialect is strongly correlated with a particular racial group; e.g., African American English, AAE). Dialect identification and subsequent racial profiling and discrimination (as manifested in housing discrimination) has already been demonstrated with as little speech as *Hello* (*6*). We present here the first study demonstrating early, pre-attentive neural responses to racial dialect perception.

We recorded the neuromagnetic brain activity of 15 healthy right-handed native speakers of Standard American English (SAE), using a 157-channel whole-head magneto-encephalography (MEG) system (KIT, Japan, Figure 1A). We employed a passive oddball mismatch (MMN) design *(7)*, in two blocks counter-balanced across participants, Figure 1B. The MMN is an automatic pre-attentive electrophysiological response that is elicited when a change in a series of stimuli is detected. In each block, participants were habituated to ten randomly presented naturally-produced, acoustically distinct tokens of the word *Hello* produced by the same bi-dialectal speaker (from (*6*)). In one block SAE tokens served as the standards, in the other block AAE tokens were the standards. Crucially, standards occur more frequently (n = 700) than deviants (n = 100). Since all tokens (deviants and standards) were produced by the same bi-dialectal speaker, the tokens were matched for talker, voice and linguistic properties, and differed solely on dialect, therefore any elicited MMN would demonstrate that dialect information is extracted in pre-attentive speech processing. The root mean square (RMS) of the magnetic field amplitudes was calculated across 10 channels of interest over auditory cortex in each hemisphere. RMS amplitudes corresponding to standard and deviant tokens were averaged within participants and conditions in epochs of 800 ms, including a

pre-stimulus baseline of 200 ms. MMNs were defined on the basis of significant differences between RMS amplitudes of deviants and standards in the left hemisphere, Figure 1C. Deviants and standards differed in two prominent regions around 200 and 400 ms post stimulus onset, corresponding to the first and second vowels in *Hello*, yielding an early (180-230 ms) and late (360-410 ms) time window for statistical analysis.

For both dialects, mean deviant amplitudes were significantly larger than standard amplitudes in the early (SAE: $t = 9.94$, $p < 0.001$; AAE: $t = 6.14$, $p < 0.001$) and late (SAE: $t = 3.23$, $p < 0.01$; AAE: $t = 17.10$, $p < 0.001$) time windows. Linear mixed models (*8*) additionally show that SAE responses have larger mean amplitudes in the early time window ($F(1,42) = 4.21$, $p < 0.05$). Therefore, by 200ms subjects respond differentially to the two dialects.

Our results demonstrate that dialect categories are accessed early in automatic pre-attentive speech processing, perhaps as a consequence of the evolutionary demands for distinguishing in-group compatriots from others. The demonstration that dialect and racial identity modulate the automatic perceptual computations involved in speech processing suggests that racial identification cannot be "turned-off" in humans *(9)*, but that we must rely on later, higher-order, conscious cognitive faculties to eliminate any potential prejudice resulting from this initial racial identification.

References and Notes:

1. J. Baugh, *Am Speech* **75**, 362 (2000).

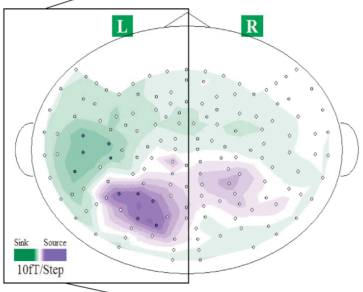
2. J. F. Dovidio, *Science* **326**, 1641 (2009).

3. P. Walker, L. Silvert, M. Hewsteon, A. C. Nobre, *Soc Cogn Affect Neurosci* **3**, 16 (2008).
4. T. A. Ito, G. Urland, *Cogn Affect Behav Neurosci* **5**, 21 (2005).
5. T. A. Ito, B. D. Bartholow, *Trends Cogn Sci* **13**, 524 (2009).
6. T. Purnell, W. Idsardi, J. Baugh, *J Lang Soc Psych* **18**, 10 (1999).
7. R. Näätänen, P. Paavilainen, T. Rinne, K. Alho, *Clin Neurophysiol* **118**, 2544 (2007).
8. J. C. Pinheiro, D. M. Bates, *Mixed Effects Models in S and S-Plus* (2002).
9. M. Weisbuch, K. Pauker, N. Ambady, *Science* **326**, 1711 (2009)
10. We thank David Poeppel for advice and support. This work was funded by the NIH, 7R01DC005660-07.

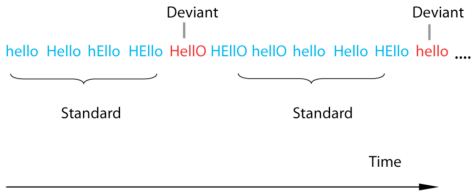
Figure 1 (A) Photograph of 157-channel whole-head MEG system. (B) Schematic of many-to-one oddball mismatch experimental design, involving several acoustically distinct tokens of each category (represented here with case differences). (C) Topography of MMN brain response, showing the magnetic field contour at 200ms, centered over left auditory cortex. Time-aligned speech waveform for a sample token of *Hello* and grand-averaged MEG responses by dialect and condition with early and late time windows highlighted with grey shading.

A. MEG whole-head system (157 channels)

MMN topography

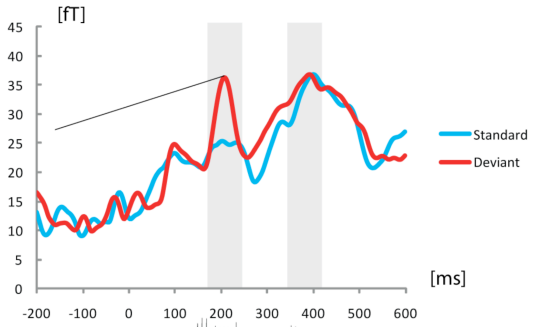


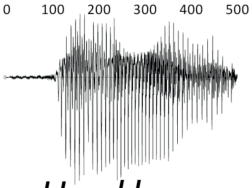
B. Passive odd-ball design



C. MEG responses in the left hemisphere

SAE





H ell o

AAE