Aerodynamic and Acoustic Characteristics of the Adult African American Voice

Christine M. Sapienza

Department of Communication Processes and Disorders, University of Florida

Summary: Laryngeal aerodynamic and acoustic characteristics of African American voice production were examined from vowel samples produced by ten adult female and ten adult male speakers. The data were compared with that for a control group consisting of ten adult female and ten adult male White speakers, matched for age, height, and weight. All measures were analyzed using Cspeech 4.0. Aerodynamic measurements, extracted from a glottal airflow waveform, included maximum flow declination rate, alternating glottal airflow, minimum glottal airflow, and airflow open quotient. Acoustic measures included fundamental frequency and sound pressure level. No significant mean differences between the African American and White speakers were found, except for maximum-flow declination rate. The White speakers produced significantly higher declination rates than the African American speakers. The factor of sex for the African American speakers was statistically significant for the measures of maximum-flow declination rate, alternating glottal airflow, open quotient, and fundamental frequency, consistent with the functioning of the White speakers. The results suggest that during vowel production, where the vocal tract is in a fairly static position, acoustic and aerodynamic characteristics for African American and White Speakers are comparable. Key Words: African American—Voice—Aerodynamic—Acoustic.

Normative databases are acquired from individuals judged to have normal speech, voice, and language abilities to establish standards of function. The majority of databases developed for standardizing the phonatory function of adult speakers have been acquired primarily from the voices of White speakers. Less empirical data are available detailing the phonatory function characteristics of other racial groups.

There are at least two implications for studying the vocal function of speakers from different racial groups. First, if functional distinctions exist between racial groups, current models that characterize normal voice production would have to consider the impact of racial variation. Second, population norms used for diagnosis and rehabilitation of voice disorders (1) would require revision so that the phonatory function characteristics of other racial groups could be incorporated into the existing database.

This study is a preliminary examination of laryngeal aerodynamic and acoustic characteristics of the African American voice, specifically describing aspects of glottal airflow, frequency, and sound pressure level during vowel production. Although many empirical studies have acquired acoustic data about the African American voice (2-7) most of these studies did not include a statistical control group for data comparison. Rather, the acquired data from the African American speaker was compared with previously reported data for White speakers. Although some of these studies have reported a trend for lower speaking fundamental frequency in young African American adult males when compared with that in young White males (2) others have found no significant differences between adult African American and White speakers for fundamental frequency sampled during vowel production (7)—the latter result recently replicated in a study by Mayo and Grant (6). Mayo and Grant (6) also

Accepted September 24, 1996.

Portions of this paper were presented at the 1995 American Speech-Language Hearing Annual Convention.

Address correspondence and reprint requests to Christine M. Sapienza, PhD, 63 Dauer Hall, Gainesville, FL 32611.
found no significant difference between the two groups in vocal perturbation, shimmer, or signal to noise ratio. The different findings of each study might be attributed to methodological factors, including the speaker's age, the influence of the speaker's dialect, and differences in recording procedures or in the equipment used to measure the fundamental frequency.

Despite the methodological differences between the mentioned studies, recent acoustic findings of Mayo and Manning (7) are interesting. Listeners can distinguish African American speakers from White speakers with high accuracy, while attending to short segments of vowel stimuli (8,9), yet fundamental frequency does not appear to be the determining factor for speaker identification. Mayo and Grant (6) hypothesized that speaker identification may be related to vowel-formant frequency differences among the two races. They compared formant frequencies from the voices of 20 African American males and 20 White American males and found the overall formant structure to be different between the two groups. Specifically, the frequency of the first (F1), second (F2), and third (F3) formants were significantly lower for the African American speakers than for the White speakers. Although the mean differences between the two groups were not great, the differences in formant frequencies could be related to variations in cephalofacial characteristics between the two groups—for example, the magnitude and anatomic pattern of the supralaryngeal structure (10,11). These results, however, were not verified by the findings of Walton and Orlikoff (12). Walton and Orlikoff found no significant differences in formant frequencies between 50 African American and 50 White speakers.

Empirical evidence documenting laryngeal anatomical or histological differences between African American and White adults is not extensive. Boshoff (13) examined the gross anatomy of 102 cadaveric larynges from Bantu South Africans and compared the data with 23 cadaveric larynges from White South Africans. Boshoff states that all of the "gross structures, the musculature, cartilages, membranes, ventricles, blood, and nerve supply of the South African Negro larynx show definite differences from those of the normal Caucasian" (p. 35). The principal anatomic findings of Boshoff's study were that the larynges of black South Africans demonstrated increased flexibility of the thyroid cartilage, longer cricoid height, larger vocal folds, the presence of a superior thyroarytenoid muscle, continuation of the oblique arytenoid muscle over the tips of the arytenoid, different type/shape of the laryngeal ventricle, and findings of "other aberrant muscles." Boshoff concluded from his dissections that "the South African Negro larynx was a more powerful organ than that of the Caucasian," as demonstrated by the laryngeal musculature in particular (p. 49).

If distinct differences in laryngeal anatomy exist between African Americans and Whites, one would presume differences in vocal fold function and alterations in the characteristics of the glottal airflow pulse. Currently, too little data are available to ascertain whether other African American speakers' vocal function is consistent with Boshoff's anatomical records from the larynges of black South Africans.

Glottal airflow measures have been used to estimate laryngeal physiological differences between such groups as children and adults (14) the young and the old (15), and the disordered and the normal (16); yet to date there is little use of these measures to examine empirically the variation between racial groups. Measures from a glottal airflow signal provide information that can be used to make hypotheses regarding vocal function and are collected in a relatively easy, noninvasive manner.

This investigation was designed to determine whether quantifiable differences in glottal airflow characteristics exist between adult African American speakers and White speakers. It is preliminary information that contributes to the existing normative database already obtained from White speakers. It provides further details of the African American voice by acquiring data from a glottal airflow waveform, a signal not examined previously for this group of speakers, and it may contribute information on the parameters used by listeners to identify African American speakers from White speakers. Examination of glottal airflow from sustained vowel production was chosen because it is a noninvasive procedure that allows phonatory function to be examined by yielding a signal without major influence of resonance characteristics of the vocal tract.

METHODS

Subjects
A total of 40 subjects, 20 African Americans and 20 Whites, were selected to participate in this study. Each group consisted of 10 females and 10 males. The age range of the African Americans was 18 to 28 years. Since such physical characteristics as age, height, and weight correlate to the size of the vocal folds and consequently effect vocal function (17,18) the White subjects were matched to the African American subjects (age, ±3 years, and height and weight ±5%). The mean age for the African American subjects was 21 ± 1.5 years and 21 ± 0.74 years for White subjects. Mean and standard deviations for height (cm) and weight (kg) were 168 ± 8.5 cm and 69 ± 7.35 kg, respectively, for the African American
subjects and 169 ± 6.4 cm and 68 ± 6.9 kg, respectively for the White subjects.

All subjects were required to meet the following screening criteria: (a) normal articulation, voice, resonance, and language, judged by a speech–language pathologist; (b) passing result from a hearing screening at 20 dB sound pressure level (SPL) at 0.5, 1, 2, and 4 kHz bilaterally; (c) freedom from symptoms of allergies or colds on the day of testing; (d) no history of voice problems, vocal pathology, respiratory, cardiovascular, or neurological disease; (e) nonsmoker; (f) no history of drug or alcohol abuse; (g) no history of professional singing and/or voice training. Additionally, all subjects reported good to excellent health on the day of testing.

**Subject personal data form**

The racial background of the subjects was determined, using a written questionnaire provided to each subject by the experimenter. The questionnaire was given to each potential subject at the beginning of the experiment and included questions relevant to family, health, and racial history. Each subject was asked to describe their racial/ethnic background to the best of their ability, to choose a category that best described the racial/ethnic background of their parents, and to describe the countries that their families claimed as places of origin. To be classified as African American or White, subjects were asked to acknowledge on the questionnaire that they considered themselves African American or White, and that both parents and grandparents were of the same race/ethnicity. Furthermore, only subjects who had been raised and had lived in the southeastern part of the United States for the first ten years of their lives were included in the study. This criterion was used to control major distinctions in regional dialectal variation. Finally, all subjects were students of the University of Florida, ensuring a comparable educational background, and all subjects spoke standard American English during the experiment.

**SPEECH TASK**

Oral airflow was obtained during three trials of sustained production of the vowel /a/. Subjects were instructed to sustain the vowel for approximately 5 seconds using a comfortable effort level. Subjects rested approximately 30 seconds between trials. Each subject’s productions were perceptually and instrumentally monitored on-line by the investigator to exclude any trial that was of extreme variation in SPL and/or frequency.

**Equipment**

A circumferentially vented, wire-screen pneumotachograph mask, coupled to a pressure transducer (Glottal Enterprises model MS 100 A-2), was used to sense a wideband oral airflow signal. The pressure transducer was calibrated with a known airflow from a rotameter (Glottal Enterprises MCU-2) prior to collecting data from each subject. Digital inverse filtering was completed using CSpeech 4.0 (19) to yield a glottal airflow waveform (20). To account for possible pressure-transducer drift, the pneumotachograph mask was removed from each subject’s face three times during the recording, at the beginning, middle, and end of the experiment. These “mask-off” conditions were used during data analysis to recalibrate the transducer’s electronic offset from baseline zero. Each of the subject’s productions were recorded to tape and monitored on-line using CSpeech. Following the subject’s vowel production, the digitized sample was inverse-filtered on-line. On-line observation and measurement of the inverse-filtered airflow waveform allowed the investigator to monitor the subject’s airflow values for each trial production and to note any extreme intrasubject variability in glottal airflow across the trials produced. Intrasubject variability was observed to be minimal across the three trials.

The airflow signal was sampled at a rate of 10 kHz, recorded and stored on VHS tape using a Vetter 3000A PCM digital data recorder. Measurement of the glottal airflow parameters was completed using a computer algorithm written within CSpeech, and a 486 workstation with a Data Translation 12-bit analog-to-digital signal processing board (Model DT2821-GSE).

Sound pressure level—obtained using the wideband pressure transducer inside the pneumotachograph mask calibrated with a tone produced by a signal generator (B & K Precision 3026) coupled to a speaker—was examined to observe the individual effort levels used to produce the vowel. A 500 Hz tone at 80 dB SPL, calibrated for a 15-cm mouth-to-microphone distance, was used (21).

**MEASUREMENTS**

A 200-ms window was extracted from the midpoint of each of the subject’s three vowel productions to obtain the measures. Data measurement from the midpoint of the vowel avoids any variability in fundamental frequency or sound pressure level that may be present during the onset or offset of sustained vowel prolongations. Figure 1 represents two cycles of vocal fold vibration from a glottal airflow waveform and delineates the points on the waveform used in the calculation of the glottal airflow measures.
Characteristics of the Adult African American Voice

Amplitude-based measures

Maximum-flow declination rate is obtained by differentiating the glottal airflow signal. It is the greatest negative peak of the differentiated glottal airflow waveform and corresponds to the fastest rate of change in airflow during adduction (Fig. 1). Maximum-flow declination rate provides information regarding the rate at which the glottal airflow shuts off during phonation (20).

Alternating glottal airflow is the amount of airflow from the maximum peak to the minimum valley (Fig. 1, X–Z) of the glottal airflow waveform and is the airflow that is modulated by the vocal folds during voicing (20).

Minimum glottal airflow is calculated from zero flow to the minimum amplitude of the glottal airflow waveform. It is the amount of airflow present during the closed portion of the glottal cycle (Fig. 1, point Z). Ideally, if the glottis is completely closed during phonation, minimum glottal airflow is zero. However, even in individuals with normal voices, minimum airflow is rarely zero (20). Theoretically, the higher the minimum airflow value the more incomplete the glottal closure.

Sound pressure level is obtained using the wideband pressure transducer calibrated for a 15-cm mouth to microphone distance. The wideband pressure transducer’s response extends from DC to 4,000 Hz (flat to 1,200 Hz); it therefore functions as a microphone for the measurement of sound pressure level (21).

Time-Based measures

Airflow open quotient is the time the glottis is open relative to the period of one vibratory cycle. Time points related to the instant of opening and closing are located on the glottal airflow waveform at 20% of the maximum amplitude (Fig. 1, Y–W/W–W') (20).

Fundamental frequency is computed from the derived glottal airflow waveform using a peak picking procedure programmed within CSpeech. Fundamental frequency is the physical correlate to the perception of pitch.

Statistical Analysis

Remeasurement of approximately 20% (n = 25) of the data for each dependent variable was completed. Pearson r correlations were performed to establish intrameasurer reliability. All measures were found to be reliable upon remeasurement at p ≤ 0.01. maximum-flow declination rate (r = 0.95); alternating glottal airflow (r = 0.98); minimum glottal airflow (r = 0.89); sound pressure level (r = 0.93); airflow open quotient (r = 0.94) and fundamental frequency (r = 0.96).

After the reliability testing, the values from the three trials of the sustained vowel production were used to compute the overall mean and standard deviation for the productions of each subject. The individual subject means were used to calculate group means and standard
deviations. A one-way analysis of variance (ANOVA) was completed for each dependent variable based on the group means. The probability level was set conservatively at \( p \leq 0.01 \).

**RESULTS**

Group mean and standard deviation data for each of the dependent measures are listed in Table 1.

**Statistical main effect—race**

No significant effects were found for race for any of the dependent variables, except for maximum-flow declination rate \( (F = 7.31, p = 0.012) \). African Americans \( (M = 213.01, SD = 81.97) \); Whites \( (M = 294.88, SD = 122.25) \).

**Statistical main effect—sex**

A statistically significant difference in maximum-flow declination rate was observed between sexes. Maximum-flow declination rate was significantly lower for the women than for the men \( (F = 9.61, p = 0.004) \): Women \( (M = 212.52, SD = 87.90) \); Men \( (M = 295.77, SD = 113.82) \).

Alternating glottal airflow was significantly lower for the women than for the men \( (F = 20.32, p = 0.004) \): Women \( (M = .146, SD = .105) \); Men \( (M = .285, SD = 0.088) \).

No statistically significant difference between sexes was observed for minimum glottal airflow \( (F = .180, p = .675) \) or for SPL \( (F = .356, p = .355) \).

Airflow open quotient was significantly higher for the women than for the men \( (F = 5.92, p = 0.010) \): Women \( (M = .608, SD = .055) \); Men \( (M = 0.548, SD = 0.095) \); as was fundamental frequency \( (F = 147.28, p = 0.004) \): Women \( (M = 218.99, SD = 26.96) \); Men \( (M = 124.24, SD = 22.04) \).

No statistically significant interactions between race and sex were found.

**DISCUSSION**

The results indicated few racial differences for any of the measures examined, except for maximum-flow declination rate. Significant differences were observed between the sexes for maximum-flow declination rate, alternating glottal airflow, open quotient, and fundamental frequency in both racial groups. No significant differences were found, either between sexes or between races, for minimum flow or sound pressure level. Finally, no interactions between race and sex were detected.

The measure of maximum-flow declination rate provides information regarding the rate at which glottal airflow shuts off during phonation. Sundberg and Gauffin (22) and Gauffin and Sundberg (23) found a strong correlation between the rate of airflow shut-off and the sound pressure level of the voice. It is widely accepted that the main excitation of the vocal tract is related to increased formant amplitude, which occurs at the time the differentiated airflow waveform reaches its maximum negative peak. Detailed models of vocal fold movement have been developed to describe the anatomic, physiologic and/or aerodynamic factors that are related to speed of glottal airflow shut-off (24,25). Scherer provides (26) an excellent overview of the relationship between the physiological signal and the mechanics of phonation and specifies a number of factors related to maximum-flow declination rate. Yet, all of this information is presumably based on White or Asian adult males and females.

Some of the factors related to an increase in maximum-flow declination rate include a larger ratio of the membranous to the cartilaginous portion of the vocal fold, increased thickness of the mucosa, increased vibra-

### Table 1. Group means and standard deviations for African American and white subjects for the aerodynamic and acoustic measures

<table>
<thead>
<tr>
<th></th>
<th>Maximum flow declination rate</th>
<th>Alternating glottal airflow</th>
<th>Minimum glottal airflow</th>
<th>Sound pressure level</th>
<th>Fundamental frequency</th>
<th>Airflow open quotient</th>
</tr>
</thead>
<tbody>
<tr>
<td>African American female</td>
<td>( M = -209.48 )</td>
<td>0.156</td>
<td>0.134</td>
<td>75.12</td>
<td>223.10</td>
<td>0.62</td>
</tr>
<tr>
<td></td>
<td>( SD = 78.40 )</td>
<td>0.114</td>
<td>0.062</td>
<td>2.25</td>
<td>29.41</td>
<td>0.04</td>
</tr>
<tr>
<td>White female</td>
<td>( M = -215.56 )</td>
<td>0.135</td>
<td>0.118</td>
<td>77.86</td>
<td>214.90</td>
<td>0.60</td>
</tr>
<tr>
<td></td>
<td>( SD = 97.39 )</td>
<td>0.095</td>
<td>0.084</td>
<td>2.54</td>
<td>24.51</td>
<td>0.07</td>
</tr>
<tr>
<td>African American male</td>
<td>( M = -217.33 )</td>
<td>0.304</td>
<td>0.108</td>
<td>76.88</td>
<td>123.95</td>
<td>0.52</td>
</tr>
<tr>
<td></td>
<td>( SD = 85.54 )</td>
<td>0.071</td>
<td>0.051</td>
<td>3.25</td>
<td>21.03</td>
<td>0.06</td>
</tr>
<tr>
<td>White male</td>
<td>( M = -374.20 )</td>
<td>0.365</td>
<td>0.083</td>
<td>77.56</td>
<td>124.53</td>
<td>0.58</td>
</tr>
<tr>
<td></td>
<td>( SD = 142.10 )</td>
<td>0.105</td>
<td>0.136</td>
<td>2.64</td>
<td>23.04</td>
<td>0.11</td>
</tr>
</tbody>
</table>

M, mean; SD, standard deviation.
tion of maximum-flow declination rate between the two racial groups (27).

Given that no other dependent variable extracted from the glottal airflow signal was statistically significant, the findings of a greater maximum-flow declination rate in the White subjects should be considered prudently until replication and further study of laryngeal anatomy and physiology of African Americans speakers is completed. How is it then that listeners are able to perceive differences between the voices of African American and White speakers during vowel production (7)? Although gross differences in the glottal airflow pulse are not apparent between the two racial groups, it may be that a combination of modest distinctions in such spectral characteristics as formant frequencies (7,12), formant amplitudes, and harmonic-to-formant ratios are the contributing factors responsible for speaker identification.

The statistical effects seen in data from the current study for the factor of sex for the African American speakers were not surprising. They correspond well with previous data from studies that have investigated sex differences relative to phonatory function in White subjects (20,28,29). The similar findings for the African American and White subjects suggest that the distinctions between the sexes in laryngeal anatomy and physiology are likely also similar. Certainly, smaller vocal folds are expected for women of any race, in that women are generally of smaller physical structure. Smaller vocal folds produce smaller vibrational amplitudes, which contribute to the findings of decreased maximum-flow declination rate, lower alternating glottal airflow, larger open quotients, and increased fundamental frequency during phonation for women when compared with the same measurements in men.

CONCLUSION

The speech task produced by the subjects was a sustained vowel production selected with the aim of limiting the vocal tract modifications that occur with coarticulation, and the linguistic and cultural factors that interact during more extemporaneous speech. The present findings suggest that for vowel production there are more similarities than differences in the vocal function between African American and White speakers. Although these data should be considered preliminary, future study should determine whether the current findings vary with different speaking tasks and explore the influence of speaker dialect, cultural, and regional variations. The current data suggest that the voice characteristics of African American speakers are likely related more to lin-
guistic and cultural influences than to laryngeal anatomical ones.

REFERENCES


