Effects of caffeine on vocal acoustic and aerodynamic measures of adult females

ABSTRACT

Purpose: To explore the effect of caffeine on voice, as evidenced by acoustic and aerodynamic measures, based on the fact that caffeine is considered to be a systemic dehydrating agent. Methods: The participants in this investigation were 58 females ranging in age from 18 to 35 years, who self-reported normal vocal production. The participants abstained from caffeine ingestion for 24 hours and from ingesting foods and liquids for 12 hours prior to the test. They were randomly assigned to one of two groups, control and experimental, and were individually tested. After the preliminary phase, participants in the experimental group ingested a 100 mg caffeine tablet; participants in the control group ingested placebo, and followed the same protocol as the experimental group. All participants were retested 30 minutes after ingesting caffeine or placebo. Results: Differences were not identified in any of the acoustic or aerodynamic measures across both groups. Conclusion: These results suggest that a conservative (100 mg) dosage of caffeine may not have an impact on vocal acoustics and aerodynamics.
INTRODUCTION

Prevention of voice disorders is an important aspect of vocology. In order to provide clinical advice and propose preventive programs for voice disorders, it is necessary to obtain accurate information and empirical data. Approaches used to avoid voice difficulties include knowledge about substances that may have a negative impact on voice production. One substance commonly associated with voice problems by causing systemic dehydration is caffeine. However, few investigations have examined the relationship between caffeine intake and voice performance. The objective of this exploratory study was to identify whether caffeine may be implicated in vocal fold function, using acoustic and aerodynamic voice parameters.

Voicing is a demanding achievement that requires a synchronized interaction of multiple physiological feats (3). There are at least three mutually dependent processes involved in voice production (i.e., respiration, phonation, and resonance). The key structure in voice production is the larynx, a critical instrument of voicing. In order to remain healthy and functional, the delicate components of the larynx require proper care.

Voice disorder symptoms are typically self-reported and accompanied by verifiable clinical signs (2). Nearly 9% of the U.S. population experience voice difficulties at some point during their lifetime (3); a larger incidence of voice disorders is typically associated with professional speakers (4,5). When the speaker relies on the voice for professional purposes, a voice dysfunction can become an occupational challenge, imposing vocational limitations that may lead to career jeopardy. Furthermore, costs related to treatment of voice disorders and work days lost due to voice problems are estimated in economic losses of billions of dollars in the U.S. annually (2). As such, the consequences of voice dysfunction may not be restricted to transitory disturbances (2), social position and professional limitations for speakers that result in a decrease in the quality of life, typically associated with reduction in physical and emotional well-being (5). In order to propose plans to optimize functional and work-related use of voice, and to reduce the economic impact of occupational voice disorders, vocologists need to identify factors that enhance voicing, as well as those that limit voice performance. This way, speakers, particularly those who habitually have a high vocal demand, can have access to scientific-based advisement in order to maintain healthy vocal folds, avoid vocal pathology, and possibly maximize vocal fold production.

Some voice problems are thought to be preventable by implementing healthy vocal behavior and avoidance of substances that could negatively impact the larynx (6). For example, systemic and surface body dehydration may be reflected in the vocal fold mucosa, increasing its viscosity and reducing mucosa mobility (7,8). Hence, increased systemic hydration reduces the adverse effect of excessive vocal effort on phonation threshold pressure (PTP), while superficial dehydration of the vocal folds increases vocal effort (9,10). Therefore, body dehydration may reduce the mobility of the vocal folds, affecting the laryngeal function by generating increased vocal effort, phonatory subglottic pressure, and voice perturbation, and, as a result, increasing the risk of phonotrauma (7,10). Excessively dry vocal folds are irritated by repeated impact more easily than vocal folds with normal lubrication (11). Thick and sticky secretions may increase the weight of a fold, impeding smooth vibratory patterns and predisposing the speaker to habituate coughing and throat cleaning, which may exacerbate irritation (6,7).

Despite limited empirical evidence, there are widespread statements recommending limitation or elimination of caffeine intake, some of which imbedded within professional literature, apparently because of the knowledge that caffeine is a diuretic and induces frequent urination, which ultimately can cause body dehydration (11,13). These recommendations tend to indicate that caffeine has a negative effect on human body; the reason for a negative impact on voice has generally been explained by how caffeine functions as a dehydrating agent, drawing fluid from interstitial cells (14). Caffeine is then considered to be a systemic dehydrating agent with consequent negative effects on the quality of voice.

The diuretic action of caffeine is rapid and of short duration (13). Moreover, dehydration does not have to be severe to have a negative effect on physical performance; even mild dehydration affects the body to some degree, leading to measurable limitations in body performance (12,15). Adverse influences of caffeine on voice have been found in an investigation about the effects of caffeine on the fundamental frequency (F0) of subjects during conversation, reading, and singing using an electrolaryngograph (10). The results revealed a substantial increase in irregularity of F0, indicating that caffeine may have a negative impact on voice quality (10).

Literature regarding the effects of caffeine in humans has measured its impact on systems other than the phonation. For instance, caffeine consumption has been associated with elevated blood pressure, a particular risk factor for coronary heart disease and stroke (12,16-24). Furthermore, when the use of caffeine becomes intense, a subsequent decrease in its ingestion may result in withdrawal indicators, leading to typical symptoms such as headache, fatigue, drowsiness, lethargy, irritability, increased muscle tension, nausea, and vomiting (25,26).

Conversely, possible caffeine effects on health reported in the literature also include positive outcomes (11). For example, caffeine ingestion has been associated with prevention of diabetes mellitus, Parkinson’s disease, colorectal cancer, and cirrhosis (11,12). Potential relevant beneficial impacts of caffeine on the body for this study include that caffeine intake is considered to temporarily relieve fatigue and to sharpen mental performance. Not surprisingly, caffeine is a common and appreciated substance, also by professional speakers.

A considerable portion of what is known about preventive measures for voice disorders is based on clinical judgments, not empirical research, although clinical advice should be based on scientific evidence (27). Empirical literature on caffeine and voice is very limited (10). As such, it is possible that vocologists who have interpolated from the general body systems literature, making the assumption that if caffeine has a negative impact on general body tissue, then it probably has a corresponding negative impact on the vocal folds and voicing. While the literature hints that caffeine may be deleterious to the functioning
of vocal folds, investigations are necessary to identify whether caffeine is directly implicated in its function, and, if so, the specific nature of the impact of caffeine on vocal fold function. Therefore, the aim of this study was to investigate the effects of a conservative dosage of caffeine (100 mg) on vocal acoustics and aerodynamics.

Physiological effects of caffeine include elevation of blood pressure and central nervous system stimulation. An excessive dose of caffeine (e.g., more than 500 mg per day for an adult individual or any regular amount of consumption by children) may present health risks\(^{19}\). Therefore, we chose to conduct an exploratory investigation about the effects of 100 mg of caffeine on vocal fold function by using two measures: acoustics and aerodynamics. This dosage of caffeine is equivalent to approximately 8 oz of brewed coffee\(^{11,12}\), although there are differences of caffeine level related to diverse brands and mode of preparation.

The independent variable in this study was caffeine, with two levels: Caffeine Condition and Non-Caffeine Condition. The dependent variables included the acoustic measurements of relative average perturbation (RAP), shimmer, sound pressure level (SPL), and the aerodynamic measure of airflow.

**METHODS**

Participants were 58 females ranging in age from 18 to 35 years, who self-reported normal vocal production. Following appropriate approval by the Institutional Review Board (protocol number 09433), the researcher contacted and explained this investigation to university instructors. After instructors’ endorsements, their students were invited to participate. The participants then received instructions about the study and were required to sign an informed consent form prior to research participation.

The participants in this investigation were females to avoid confound from variation in F0 across males and females, which could introduce unwanted variability in the data. Additional research control involved limiting age range of participation to 18 to 35 years of age, to avoid changes associated with the aging voice. Furthermore, individuals who self-reported normal voice were selected to avoid confounds associated with voice impairments. Participants included in the study also reported not ever having felt discomfort after ingesting caffeine, and rated their general health as good, with no history of high blood pressure or coronary disease.

Four days prior to the test day, each participant received a reminder e-mail with details about preparation for testing, as well as schedule observance. In an attempt to reduce variability in the degree of systemic hydration, participants were instructed to refrain from ingestion of food and liquids for 12 hours prior to testing. The testing was applied on one day, one participant per test day, always beginning at 8 a.m. Participants were randomly assigned to either a control or an experimental group.

Variability in F0 was controlled across speech sound tokens through an initial identification of a comfortable pitch by matching each participant’s emission of /a/ with a corresponding note using an electronic keyboard. Prior to the production of each new vocal token, the investigator again played the note on the keyboard, instructing the participant to produce the vowel at the same pitch as the note played on the keyboard.

The pretest and posttest procedures were identical, requiring three sustained productions of the vowel /a/ per participant, per test. Following the pretest, each participant of the Caffeine Condition group ingested a 100 mg caffeine tablet, while participants in the Non-Caffeine Condition group ingested a placebo tablet. Following completion of the pretest, all participants were required to remain in a quiet environment for 30 minutes while awaiting the posttest. A period of 30 minutes between pretest and posttest was used in this study because, while caffeine appears in tissue fluid approximately five minutes after ingestion, blood plasma levels of caffeine reach its peak approximately 30 minutes after ingestion\(^{24}\).

**Data collection**

Each participant’s target vowel production was captured three times using the Computerized Speech Lab model 4500 (CSL), and the Phonatory Aerodynamic System model 6600 (PAS), computer-based systems manufactured by KayPentax\(^{8}\). Relative average perturbation of RAP (jitter), shimmer, sound pressure level (SPL), and airflow were extracted from each token, forming the raw units for assessments.

Acoustic voice samples were collected using the Multi-Dimension Voice Program (MDVP) module of the CSL, which is equipped with analog input and output channels with a frequency response of 5 to 22 Hz. The MDVP is a multi-parameter acoustic analysis tool that acquires and analyzes acoustic voice parameters. A professional cardioid microphone (Shure dynamic model BG with frequency response of 85 Hz to 14 kHz) was placed on a stand for stability, with a 10 cm straw strategically positioned in order to maintain a standard mouth-to-microphone distance. Recommendations for optimal acquisition of vocal acoustic signals using off-axis positioning of 45° to 90° from the mouth axis were observed so as to reduce aerodynamic noise from the mouth during speech. Participants were instructed to maintain a 10 cm mouth-to-microphone distance. Speech sound tokens extending from second 1.5 to 2.5 from the onset of each target vowel sound produced by each participant were selected for analysis.

The PAS was used to collect aerodynamic data. The external PAS hardware module is an ergonomic device that has bilateral handles held by the participant during data acquisition, with integrated face mask, pneumotach, pressure transducer, and microphone. The microphone for pitch and amplitude calculations is mounted at a fixed distance and is pre-calibrated at the factory for each system to maximize accuracy.

**Variables**

The independent variable in this study was caffeine, with two levels: Caffeine Condition and Non-Caffeine Condition. Pharmacological properties of caffeine include a dehydration action that may affect laryngeal function. The dose of caffeine used in this investigation was 100 mg, an amount equivalent to
a typical 8 oz cup of coffee. The dependent variables included
the acoustic measures of RAP, shimmer, and SPL, and the
aerodynamic measure of airflow.

RAP and shimmer are acoustic measurements of insta-
bility or perturbation of vocal fold vibrations. RAP gives
an evaluation of the variability of the pitch period within
the analyzed voice sample at smoothing factor 3-period.
Shimmer percentage gives an evaluation of the variability
of the peak-to-peak amplitude within the voice sample.
Assessment of voice perturbation has been used extensively
in voice diagnostics, treatment documentation, and analysis
of voice quality and voice properties of a variety of groups.
These parameters can help in differentiating between healthy
and pathologic voices in some voice disorder types. Sound
pressure level (SPL) is as a measure of the voice intensity
in decibels.

Airflow measurements refer to the amount of air passing
through the region of the vocal folds. Increased effort leads to
a restriction of flow, resulting in greater pressure. Likewise,
with decreased effort, air flows easily through the system and
less pressure is needed to generate the flow.

Data analysis

Statistical comparisons of Non-Caffeine Condition and
Caffeine Condition values were calculated using SPSS software
version 21.0 in four univariate ANOVAs. The statistical profile
of relative average perturbation of RAP, shimmer, SPL, and
airflow was extracted for each token.

RESULTS

The main outcomes of this study are related to the three
acoustic measures and one aerodynamic measure described
in the research questions involving: (a) RAP, (b) shimmer,
(c) SPL, and (d) airflow. Because a change in the level of sig-
ificance will affect the risk for both Type I and Type II error,
the traditional alpha level of 0.05 was selected for all analyses
in this study.

Descriptive statistics for both Caffeine and Non-Caffeine
Condition groups, including means and standard deviations,
are described on Table 1.

Comparisons between groups revealed that results related to
the impact of the caffeine condition on voice acoustics as mea-
sured by RAP, shimmer, and SPL were not different, as follows:
RAP – F(1.57)=0.02, p>0.05; shimmer – F(1.57)=0.88, p>0.05;
SPL – F(1.57)=0.53, p>0.05 (Table 2). Similarly, comparisons
of results from the Caffeine Condition and Non-Caffeine
Condition did not lead to differences in voice aerodynamics as
measured by airflow – F(1.57)=0.65, p>0.42 (Table 3).

Further inspection of the descriptive values from each mea-
surement in this study demonstrated an increase in shimmer and
airflow in the Caffeine Condition. Conversely, RAP and SPL did
not reveal overall increase in the Caffeine Condition (Table 1).

DISCUSSION

Caffeine from dietary sources is pointed out as the most fre-
cquently and widely consumed stimulant in the world(28). Consumers
and professionals should be aware of the possible contribution of
dietary caffeine in general and specific health concerns.

In this study, ingestion of 100 mg of caffeine did not
adversely affect voice production. These results are compat-
ible with prior negative associations of caffeine and voice
performance with phonation threshold pressure and perceived
phonatory effort(29), as well as with indications of moderate
increase in fundamental frequency (F0) irregularity upon caf-
feine ingestion(10).

<table>
<thead>
<tr>
<th>Group</th>
<th>n</th>
<th>RAP Mean</th>
<th>RAP SD</th>
<th>Shimmer Mean</th>
<th>Shimmer SD</th>
<th>SPL Mean</th>
<th>SPL SD</th>
<th>Airflow Mean</th>
<th>Airflow SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>CC</td>
<td>29</td>
<td>1.08</td>
<td>0.64</td>
<td>5.46</td>
<td>3.02</td>
<td>57.34</td>
<td>5.32</td>
<td>0.15</td>
<td>0.14</td>
</tr>
<tr>
<td>NC</td>
<td>29</td>
<td>1.10</td>
<td>0.43</td>
<td>4.89</td>
<td>1.19</td>
<td>58.39</td>
<td>5.61</td>
<td>0.12</td>
<td>0.10</td>
</tr>
</tbody>
</table>

Table 1. Descriptive statistics

Legend: SD = standard deviation; CC = caffeine condition; NC = noncaffeine condition; RAP = relative average perturbation; SPL = sound pressure level

<table>
<thead>
<tr>
<th>Acoustic measures</th>
<th>Sum of squares</th>
<th>df</th>
<th>Mean square</th>
<th>F</th>
<th>p-value</th>
<th>Effect size</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jitter</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between groups</td>
<td>0.007</td>
<td>1</td>
<td>0.007</td>
<td>0.02</td>
<td>0.88</td>
<td>0.00</td>
<td>0.05</td>
</tr>
<tr>
<td>Within groups</td>
<td>17.001</td>
<td>56</td>
<td>17.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>83.104</td>
<td>57</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shimmer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between groups</td>
<td>4.70</td>
<td>1</td>
<td>4.70</td>
<td>0.88</td>
<td>0.35</td>
<td>0.02</td>
<td>0.15</td>
</tr>
<tr>
<td>Within groups</td>
<td>296.70</td>
<td>56</td>
<td>5.29</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>301.41</td>
<td>57</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between groups</td>
<td>16.11</td>
<td>1</td>
<td>16.11</td>
<td>0.53</td>
<td>0.46</td>
<td>0.01</td>
<td>0.11</td>
</tr>
<tr>
<td>Within groups</td>
<td>1667.21</td>
<td>56</td>
<td>29.95</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1693.32</td>
<td>57</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Caffeine impact on voice acoustics

ANOVA (α=0.05; p≤0.05)

Legend: df = degrees of freedom; F = distribution under the null hypothesis; SPL = sound pressure level

In this exploratory study about potential influences of a conservative dose of caffeine on voice performance, negative findings may be explained, if not expected, by the reduced level of caffeine involved in the treatment (i.e., 100 mg – Caffeine Condition). A conservative approach involved an initial minimal caffeine dosage level in order to avoid potential problems related to maximal dosage levels. Furthermore, it stands to reason to examine a reduced dose as a first step when investigating possible effects of caffeine on voice. Ultimately, if confirmed by supplementary studies, the information that conservative caffeine dosages do not affect voice negatively could be of interest to those who are attentive to voice health care practices and research. Additionally, clinical implications of this investigation are applicable to persons, such as professional speakers and singers, who are interested in maximizing vocal performance and avoiding vocal disorders while still partaking in caffeine. In summary, clinicians, researchers, and consumers may use the results of this essentially negative study as part of preliminary considerations suggesting that ingestion up to 100 mg of caffeine may not cause vocal degradation.

Although not different, some of the parameters investigated (i.e., shimmer and airflow) demonstrated changes in the direction anticipated (i.e., degraded) as compared to pretest values. These results may suggest that, while a conservative dose of caffeine (i.e., 100 mg) did not demonstrate a different impact on the selected voice acoustics and aerodynamics parameters, it is possible that higher caffeine consumption would produce a stronger effect on voice performance.

One hundred milligrams of caffeine did not generate a different degradation of vocal acoustics and aerodynamics. However, it is not uncommon for persons to ingest more than 100 mg of caffeine per day. Therefore, additional investigations are indicated to measure the impact of higher doses of caffeine ingestion on voice acoustics and aerodynamics. Furthermore, because caffeine is a drug that is metabolized in the body, the impact of caffeine ingestion must vary across time. It takes approximately five hours for the liver to metabolize half of a given dose of caffeine. Therefore, if eventually a caffeine-voice effect is found, then it would be prudent to explore the effect curve of caffeine on voice.

Some limitations of the methods used to control extraneous influences include the prohibition of generalization to males, to females outside the age range of this investigation’s participants, to persons with voice disorders, to those who have not abstained from food and drink for 12 hours, to amounts of caffeine ingested other than 100 mg, and to durations other than those used in this investigation. Additional methodological concerns include accuracy of participant self report for identification of prior voice disorders, and whether participants correctly followed the pretest protocol requiring (a) abstention from caffeine for 24 hours and (b) fasting for 12 hours. Further concerns are related to reliability of acoustic perturbation measurement and to reduced sensitiveness of acoustic and aerodynamic measures to identify differences between pre and posttest, which may explain in part why results were not significant.

Supplementary studies regarding the relationship between caffeine intake and voice performance plan to increase research control by observing menstrual cycle periods for possible fluctuation in hydration level of female participants, obtaining the weight of each participant, and analyzing urine and blood samples in order to validate body hydration levels.

It is also essential to investigate the effect of caffeine in individuals who use their voice daily for professional purposes. Moreover, inspecting the effects of caffeine on voice performance in users of different substances that contain caffeine constitutes one more branch of this research thread. Finally, extended research about the effect of the interaction of caffeine and hydration on the voice should be achieved.

### Table 3. Caffeine impact on voice aerodynamics

<table>
<thead>
<tr>
<th>Aerodynamic measure</th>
<th>Sum of squares</th>
<th>df</th>
<th>Mean square</th>
<th>F</th>
<th>p-value</th>
<th>Effect size</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>Airflow</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between groups</td>
<td>0.01</td>
<td>1</td>
<td>0.01</td>
<td>0.65</td>
<td>0.42</td>
<td>0.01</td>
<td>0.12</td>
</tr>
<tr>
<td>Within groups</td>
<td>0.90</td>
<td>57</td>
<td>0.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>0.91</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

ANOVA (α=0.05; p≤0.05)

Legend: df = degrees of freedom; F = distribution under the null hypothesis

CONCLUSION

Knowledge of substances that may affect voice performance is an essential component of voice health programs. Caffeine is commonly described as a dehydrating agent of the body system, and presumably of the voice as well. However, there is little evidence that caffeine intake is detrimental to voice production.

In this study, differences were not apparent in any of the acoustic or aerodynamic voice measures, suggesting that the ingestion of a 100 mg dosage of caffeine may not have a negative impact on vocal parameters when the protocol involves a period of vocal rest between pretest and posttest. Further research is warranted to examine if these results would be reproduced under high vocal demands.

These introductory findings may be helpful in efforts to elucidate the association of caffeine and voice. The application of a conservative dosage of caffeine to examine voice performance with objective acoustic and aerodynamic measures in a controlled environment was an attempt to have a preliminary look at the impact of caffeine intake in the voice, leading to the conclusion that, under standardized conditions, 100 mg of caffeine may not lead to deleterious voice effects.
Effects of caffeine on voice

*MCF carried out study design, data collection, data analysis and manuscript writing; KOS was responsible for study design and manuscript writing; AS carried out data tabulation and participants’ management.

REFERENCES