The Journal of Laryngology and Otology
April 1999, Vol. **113**, pp. 341–345

Effect of caffeine on the vocal folds: a pilot study

S. AKHTAR†, F.R.C.S., G. WOOD*, B.Sc, J. S. RUBIN†, F.R.C.S., F.A.C.S., P. E. O'FLYNN†, F.R.C.S., P. RATCLIFFE*, M.Sc.

Abstract
Caffeine is considered to be a dehydrating agent with detrimental effects on the quality of voice of persons ingesting it. This has led medical personnel dealing with voice disorders, especially in the case of professional voice users, to give advice against the use of caffeine. Yet this is an anecdotal truth as an extensive Medline literature search did not reveal any scientific evidence of caffeine being proven to have adverse effects on the vocal folds. We, therefore, initiated this pilot study to ascertain the connection between caffeine and voice quality on a laboratory basis. Two hundred and fifty mg of caffeine were provided to eight volunteers in tablet form, and blood levels along with laryngograph readings were recorded to document the changes produced. Analysing the irregularities of frequencies in a) free speech b) a reading passage and c) singing 'Happy Birthday', substantial changes were seen to authenticate the fact that caffeine does produce alterations in voice quality but these alterations have considerable intra-subject variability. A full study with wider parameters is to be performed on this subject as we consider it to be of importance in the management of voice disorders.

Key words: Caffeine; Dehydration; Vocal fold

Introduction

There are numerous pharmacological agents with effects on speech production and voice. Most produce transient and reversible changes. Caffeine is one such agent. A thorough Medline literature search has failed to reveal any scientific evidence to substantiate the generally accepted fact that caffeine is a dehydrating agent and therefore detrimental to laryngeal function. Yet adequate hydration of the vocal folds and avoidance of dehydrating agents and habits, are regular recommendations made by medical personnel and speech therapists to most patients with voice problems (Bradley, 1980; Lawrence, 1981; Verdolini, 1988.) Typical advice includes: ambient humidification, steam inhalation, increased water intake, along with refraining from drying conditions such as smoke, alcohol, diuretics, antihistamines, and caffeine (Bradley, 1980; Bastian and Lawrence, 1984; Sataloff, 1987; Verdolini, 1988).

Caffeine, theophylline and theobromine, are methylxanthines. They have been exploited for their stimulant action on the mood, elevation of fatigue, and increased work output. Of these caffeine is undoubtedly the most commonly ingested alkaloid. It is present in the seeds of the coffee plant, tea leaves, cocoa, and in soft drinks derived from the kola nut of the *Cola acuminata* tree. (Gillies *et al.*, 1986).

The pharmacological properties of caffeine include: relaxation of smooth muscle (notably bronchial muscle), producing bronchodilation; stimulation of the central nervous system, producing increased vigilance; stimulation of the cardiovascular system producing tachycardia and increased cardiac output; acting as a loop diuretic, producing mild diuresis in a manner similar to frusemide.

The sympathomimetic properties of caffeine added to its diuretic action lead to its being a dehydrating agent, a combination which we postulate can lead to changes in the voice.

To estimate the changes produced in voice quality an electrolaryngograph can be used. The electrolaryngograph displays conductance against time between two surface neck electrodes placed either side of the thyroid cartilage. When the vocal folds close, the conductance increases. The resulting periodic electrical signal has been called the Lx waveform. The phrases of vocal fold contact are reflected in the Lx waveform. The Lx signal, in addition to providing an estimation of the open and closed phase duration, also provides a convenient means of measuring fundamental frequency (Fx) parameters (Carlson, 1995).

Irregularity of the fundamental frequency is a measure based on taking a three bin width diagonal of the Cx plot to represent regular voice. All

From the Department of Speech and Language Therapy* and the Department of Otolaryngology/Head and Neck Surgery†, Royal National Throat, Nose and Ear Hospital, London, UK.
Accepted for publication: 21 December 1998.

fundamental frequency pairs outside this central band are calculated to present as a percentage of the total number of fundamental frequency pairs (PCLX Manual, 1992).

As the standard of care has improved, the need for objective assessment quantification of voice function has become more essential. This pilot study as a potential part of a larger one, examines irregularity as the sole parameter, using the laryngograph PCLX by Laryngograph Limited, to analyse voice and any irregularities thereof produced by a controlled and uniform intake of caffeine. In future studies the comparative effect of caffeine on professional and non-professional voice users will also be assessed.

Materials and methods

Eight volunteers, four males and four females, participated in this pilot study. Their ages ranged from 27 years to 55 years. Inclusion criteria were: 1) no evidence or history of voice disorders, 2) no evidence of current systemic illness, especially respiratory and cardiovascular ailments, 3) an informed consent. Subjects were requested to refrain from any form of caffeine intake for 24 hours prior to the day of the study itself. It was also ensured that on the morning of the study, the participants were involved in non-clinical work only. Ethical approval for the study was obtained from the Royal Free NHS Trust ethical committee.

On the morning of the study the daily average amount of caffeine consumed in all forms, calculated to be 250 mg, was administered in an over-the-counter preparation Proplus. Proplus contains 50 mg of pure caffeine in a lactulose base. Five tablets of it were taken in a single dose at the commencement of the clinical evaluation by each subject.

Blood concentrations were determined pre- and then one hour post- caffeine ingestion. To ensure ease and continuity of testing in the voice laboratory a staggered time-table of activities was arranged for the participants.

The pattern of vocal fold vibration was evaluated by the parameter of irregularity of fundamental frequency. Irregularities of frequency were measured using the laryngograph PCLX. They were recorded as pre-caffeine (base-line reading) and then as two consecutive hourly recordings post-caffeine ingestion. The irregularity of frequency of voice was scrutinized in three different voicing techniques: in free speech, in a reading passage and in the 'Happy Birthday' tune sung in a comfortable mid-register.

Recording was performed via placement of electrodes on to the thyroid cartilage with a microphone in close proximity to the mouth. No problems with obesity of the neck and consequent placement of the electrode were encountered. It was attempted to achieve the same electrode position every time a reading was taken. Irregularity changes were assessed based on the pre- caffeine recordings; that is the pre- caffeine irregularities were regarded as baseline for any particular individual. For statistical analysis a one-factor, within-subjects variability design was used to examine the effects of caffeine on the vocal folds. The independent variable caffeine was tested at three levels: i) pre-caffeine, ii) post-caffeine at one hour and iii) post-caffeine at two hours. The dependent variable was the irregularity in vocal fold performance as measured by reading, free speech and singing.

Results

The intake of caffeine itself was not associated with any major side-effects, except in one case where hyper-excitability was noted. In this same person post-ingestion levels were recorded to be highest of all the participants.

The caffeine concentrations in the subjects did show some variation as documented in Table I. In one case water had been consumed in between tests and, in another, lemon tea. However, on an average, the pre-ingestion levels were 1 mg/ml and the post-ingestion were 7.8 mg/ml. This demonstrates that adequate blood levels of caffeine had been achieved to render it a reliable study in all participants.

Irregularity values for the reading passage are shown in Table II, for free speech shown in Table III, and 'Happy Birthday' listed in Table IV. It is important to mention that 1) perceptual changes in quality of voice were not considered in our analysis and 2) as by definition irregularity has no value for 'normality', a voice with 15–20 per cent laryngographic variations is considered hoarse by some (Barry *et al.*, 1991). Clinical experience suggests that a normal male voice may display between five to 10 per cent irregularity (Carlson, 1995).

In the case of the reading passage the pre-caffeine levels of irregularities were less than 10 per cent 'normal' in three participants, slightly raised in two and grossly raised in three. In all three 'normals' it had returned to near normal by two hours post-caffeine ingestion. Overall the mean irregularities increased slightly on the first hour and further over the second hour. There was considerable variation seen however.

In the analysis of the singing of 'Happy Birthday' there were several case variabilities. All participants showed less than 10 per cent irregularity, pre-caffeine with a mean of 6.3. The mean irregularity was measured at one and then again at two hours. The mean was still under 10 per cent at two hours, although three of the eight had irregularity over this level.

TABLE I
CAFFEINE CONCENTRATIONS

Subject	Pre-caffeine	Post-caffeine
	Blood levels mg/l	
1.	2.3	8.2
2.	<1	7.3
3.	<1	5.0
4.	<1	6.0
5.	<1	8.5
6.	<1	6.3
7.	1.2	8.6
8.	<1	12.4

https://doi.org/10.1017/S0022215100143920 Published online by Cambridge University Press

TABLE II
IRREGULARITY FOR READING PASSAGE (IN %)

Subject	Pre-caffeine	1 Hr	2 Hr
1.	24.7	11.1	11.8
2.	12.3	19.3	33.9
3.	7.0	12.9	6.8
4.	7.2	15.2	8.7
5.	24.3	29.9	48.5
6.	13.4	16.5	23.9
7.	18.1	14.5	9.9
8.	9.4	14.9	11.6
mean	14.55	16.79	19.39
median	12.85	15.05	11.70
S.D.	7.12	5.82	14.93

TABLE IV
IRREGULARITY FOR SINGING OF "HAPPY BIRTHDAY" (IN %)

Subject	Pre-caffeine	1 Hr	2 Hr
1.	5.8	6.7	7.7
2.	2.7	5.8	13.8
3.	9.2	5.2	6.5
4.	1.6	6.9	3.1
5.	9.8	23.6	22.0
6.	10.0	4.8	12.6
7.	6.1	4.8	4.0
8.	5.2	2.6	5.7
mean	6.3	7.55	9.43
median	5.95	5.5	7.15
S.D.	3.18	6.62	6.35

Free speech proved to be most intriguing as the base line reading itself showed an increased level of irregularity in all participants, but one. In seven of eight subjects the irregularity increased in the first hour. There was a mean increase from 16 to 24 per cent in the first hour which overall stabilized in the second hour, although there was intra-subject variability.

Descriptive statistics of the results by means of ANOVA was performed. The standard deviation (SD) in Tables II, III and IV suggested that the homogeneity of variance assumption may have been violated. This indicated that the population variances were not homogenous and there is a need to increase the number of subjects in the main study. The large increase seen in the group SD in the free speech condition at one hour and the reading condition at two hours indicated that the residual effect of caffeine on the vocal folds was highly variable between individuals.

A one-way ANOVA (repeated measures) confirmed that the main effects of caffeine at all levels of the independent variable was statistically non-significant. However, highly significant main effects between subjects, as opposed to within subject effects, were seen in all three conditions (reading $F = 0.87$, $p<.001$, free speech; $F = 27.71$, $p<.001$, singing: $F = 22.19$ $p<.01$). This has suggested that the variation in irregularity were more due to individual differences than an effect of caffeine per se.

TABLE III
IRREGULARITY FOR FREE SPEECH (IN %)

Subject	Pre-caffeine	1 Hr	2 Hr
1.	7.8	13.7	18.8
2.	19.8	23.7	35.0
3.	14.0	17.4	10.2
4.	11.8	21.5	15.0
5.	25.2	61.8	58.3
6.	13.0	16.1	25.6
7.	24.2	21.8	19.4
8.	13.5	20.1	12.3
mean	16.16	24.51	24.33
median	13.75	20.80	19.10
S.D.	6.21	15.42	15.84

Discussion

Traditionally caffeine has been considered to be the most potent of the methylxanthines. Persons ingesting caffeine or caffeine-containing beverages usually experience less drowsiness, less fatigue, and a rapid, clearer flow of thought. As the dose of caffeine is increased, signs of progressive CNS stimulation are produced including nervousness or anxiety, restlessness and insomnia. Focal and generalized convulsions are produced at the dose of 10 grams, which is the toxic dose in man. Stimulation of the respiratory centre occurs with these higher doses (Hardman, *et al.*, 1995).

Ingestion of 80–250 mg of caffeine, the amount contained in one to three cups of coffee produces an increased capacity for sustained intellectual effort and decreases reaction time. The long-term ingestion of caffeine can produce tolerance, and this is an exceedingly common and benign dependence, seen in coffee drinkers. (Gillies *et al.*, 1986).

The methylxanthines are readily absorbed after oral, rectal or enteral administration. In the absence of food, uncoated tablets of caffeine produce maximum plasma concentrations within one hour. Methylxanthines are distributed into all body fluid compartments. They cross the placenta and are present in breast milk and are eliminated primarily by the liver. The half life is increased in pregnancy and by contraceptives and steroids (Hardman, *et al.*, 1995).

It is estimated that the average intake of caffeine in the USA is between 170 and 200 mg per day, similar figures apply for the UK (Hardman, *et al.*, 1995). In America 90 per cent of the intake is from coffee. Depending on the alkaloid content of the coffee bean and the method of brewing:

one cup of coffee contains 65–175 mg of caffeine;
one cup of tea contains 50 mg of caffeine;
one cup of cocoa contains 5 mg of caffeine;
12 oz/360 ml of a cola drink contains 40–50 mg of caffeine;
30 gm bar of chocolate contains 25 mg of caffeine.

Of note is the fact that there is, compared with 200 ml of a cup of tea, a sixth more caffeine in instant coffee, whereas filter coffee has two and a half times more caffeine.

https://doi.org/10.1017/S0022215100143920 Published online by Cambridge University Press

In caring for professional voice patients with voice problems, the clinician will frequently guide the patients to avoid various sources of caffeine, on the basis that caffeine will dehydrate the larynx and exacerbate the problem (Verdolini-Marston *et al.*, 1994a). The effects of dehydration on the larynx are said to include increased production and viscosity of mucus leading to inadequate vocal fold lubrication. Roping and stranding of mucus from one vocal fold to the other can also be seen and probably reflects the increased viscosity of mucus (Benninger *et al.*, 1993). Viscosity is measured in Poise (P). The viscosity of pure water is 0.01 P, whereas the viscosity of vocal fold tissue is ~1–10P. Mucus and sputum are ~0.1–1P (Titze and Talkin, 1979).

These signs are often associated with miscellaneous symptoms such as a sensation of dryness of the nose, mouth and throat, frequent throat clearing, sudden changes in voice quality, mouth breathing, and decreased durability and range of the voice with pitch breaks. Yet, despite these extensive claims there are no objective studies present to quantify these changes.

On the other hand, the beneficial effects of hydration have been described in patients with increased phonatory effort, with laryngeal nodules and polyps. (Verdolini-Marston *et al.*, 1994a). Specific claims have been made that hydration may contribute to the regression of oedema-based lesions and also to improvements in dysphonia (Verdolini-Marston *et al.*, 1994b). That phonatory effort may be improved with hydration treatment is also suggested by Lawrence (1981). The basis of hydration treatment is that it decreases the viscosity of the vocal folds which in turn decreases the phonatory threshold pressure (PTP) which is the minimal sub-glottic pressure required to initiate and sustain vocal fold oscillation. Verdolini (1988) also states that, despite the widespread claims of the benefits of hydration therapy, little experimental data is available about the actual impact of hydration on voice. Bough *et al.*, (1996) reported that, utilizing Jitter and Shimmer as standards, no definite correlation could be found between subjective feeling of the research subjects, temperature, sleep time, time of last meal and hydration levels.

In the assessment of vocal pathology an essential aim must be to seek parameters which directly relate to voice production. The laryngograph is ideal for this as it records the events in vocal fold vibrations, which are most relevant to quality of voice production. It measures the laryngograph output waveform (Lx). The changes in the quality of this waveform and the fundamental frequency (Fx) combine to produce the irregularities in frequencies. They provide information concerning the frequency of excitation of the vocal tract during speech, more accurately than any other signal. Laryngography is used to measure this parameter as it is a non-invasive method and unaffected by acoustic noise. It is also feasible for long-term voice analyses and management, dependant only on the availability of the Lx waveform.

This study attempted to evaluate the effect of a specified amount of ingested caffeine on voice production, looking specifically at one voicing parameter, that of irregularity of the fundamental frequency. This variability acoustically corresponds to irregularities in the fundamental pitch period of the vocal fold vibration and reflects the 'roughness' in voice quality, which in other words means that a rough or 'unstable' voice displays a high degree of irregularity (Carlson, 1995). It was thus felt to be a simple, reliable, repeatable and valuable parameter to investigate. Also as we were causing a moderate degree of mucosal dehydration, the vocal fold vibrations were expected to be distorted. These effects were considered to be best reflected by the irregularity of frequency produced. Not only was free speech investigated in this manner, the usual acoustic investigation, but connected speech and singing were also reviewed to allow for an evaluation closer to speech.

A number of interesting findings came out of this study. The mean percentage of irregularity increased over time in all three investigations, suggesting that there is indeed an effect on the vocal folds caused by caffeine ingestion; and that this effect is deleterious to the voice. However, there was considerable inter-subject variability in response to caffeine ingestion. The cause of this is unclear. It is particularly noticeable that these untrained subjects had far less irregularities in song than in either connected or free speech. This is most likely due to better breath support.

It must be recalled that this is a very small study with only eight subjects. It was not age or sex matched. The variation in irregularities also highlighted the often clinically encountered situation of the individual differences in response to caffeine. The results are of enough interest, however, to support a larger study perhaps comparing trained and untrained voice users and also to include age and sex matching.

Conclusion

Eight subjects were investigated for the effects of caffeine ingestion on the voice. A single voicing parameter, that of irregularity of the fundamental frequency was scrutinized. It was investigated in free speech, connected speech and song at one and two hours post-caffeine ingestion. Although considerable inter-subject variability was revealed, the mean percentage of irregularity increased at one hour then further at two hours post-ingestion especially for connected speech and for singing 'Happy Birthday'. In the case of free speech the irregularity reached a plateau at one hour to stay virtually unchanged thereafter.

References

Barry, W. J., Goldsmith, M., Fourcin, A. J., Fuller, H. (1991). Stability of voice frequency measures in speech. *Twelfth International Congress in Phonetic Science* **2:** 38–41.

Bastian, R. W., Lawrence, V. L. (1984) Hoarseness in singers. *National Association of Teachers of Singing Bulletin* **40(3):** 26–27.

Benninger, M. S., Jacobson, B. H., Johnson, A. F. (1993) In *Vocal Arts Medicine*. The care and prevention of professional voice disorders. 1st Edition, Thieme Medical Publishers, New York, pp 295–300.

Bough, D. Jr, Heur, R. J., Sataloff, R. T., Hills, J. R., Cater, J. R. (1996) Intra-subject variability of objective voice measures. *Journal of Voice* **10(2):** 166–174.

Bradley, M. (1980) Prevention and correction of vocal disorders in singers. *National Association of Teachers of Singing Bulletin* **36(5):** 38–41.

Carlson, E. (1995) Electrolaryngography in the assessment and treatment of incomplete mutation (puberphonia) in adults. *European Journal of Communication* **30:** 140–148.

Hardman, J. G., Limbird, L. E., Molinoff, P. B., Ruddon, R. W., Gilman, A. G. (eds.) (1995) *Goodman and Gilman's: The Pharmacological Basis of Therapeutics,* 9th Edition. McGraw Hill Co., New York, pp 670–680.

Gillies, H. C., Rogers, H. J. Spector, R. G., Trounce, J. R. (1986). *A Textbook of Clinical Pharmacology*. 2nd Edition. Hodder and Stoughton, London. pp 860–861.

Lawrence, V. L. (1981) Handy household hints: to sing or not to sing. *National Association of Teachers of Singing Bulletin* **37(3):** 23–25.

PCLX Laryngograph Manual (1992) page 2.39.

Sataloff, R. T. (1987) The professional voice: part I. Anatomy, function, and general health. *Journal of Voice* **1(1):** 92–104.

Titze, I. R. Talkin, D. T. (1979) A theoretical study of the effects of various laryngeal configurations on the acoustics of phonation. *Journal of the Acoustic Society of America* **66(1):** 60–74.

Verdolini, K. (1988) Practice good vocal health and prevent those voice disorders. *Choristers Guild Letters* **2:** 40–44.

Verdolini-Marston, K., Sandage, M., Titze, I. R. (1994a) Effect of hydration treatments on laryngeal nodules and polyps and related voice measures. *Journal of Voice* **8:** 30–47.

Verdolini-Marston, K., Titze, I. R., Fennell, A. (1994b) Dependence of Phonatory effort on hydration level. *Journal of Speech and Hearing Research* **37:** 1001–1007.

Address for correspondence:
Ms Shadaba Akhtar, F.R.C.S.,
E.N.T. SpR,
Royal National Throat Nose and Ear Hospital,
330–331 Gray's Inn Road,
London WC1X 8DA.