2014

Testing of the VocaLog Vocal Monitor

Jeff Searl
Kansas City, Kansas

Angela M. Dietsch
Kansas City, Kansas, angela.dietsch@unl.edu

Testing of the VocaLog Vocal Monitor

Jeff Searl and Angela Dietsch

Kansas City, Kansas

Summary:

Objective. To elucidate capabilities and limitations of the VocaLog, a device marketed to log-calibrated dB sound pressure level (SPL). Study Design. The study design varied depending on the experiment. All were prospective. Some were case series, and others were cohort studies without controls.

Method. Experiments were conducted to determine (1) whether the VocaLog logged phonatory activity and silence when it should, (2) if nonphonatory activities were detected, (3) correlation of VocaLog dB values to an external sound level meter (SLM), and (4) accuracy of phonation time (PT) and speaking time (ST) estimates from the VocaLog.

Results. Silence and phonatory activity were logged as such nearly 100% of the time. Nonphonatory activities were sometimes detected as dB values, including coughs, throat clear, belching, and swallows. The dB values from the VocaLog were strongly correlated with dB SPL from an external SLM. When on the neck, the device rarely picked up external sounds when the external noise was between 85 and 103 dB SPL. The VocaLog gave a reasonable estimate of ST but overestimated PT.

Conclusions. Overall, the VocaLog holds promise as means of indexing vocal loudness via calibrated dBSPL levels. However, some nonphonatory activity is also likely to be logged. The device provides a reasonable estimate of ST, but not PT.

Key Words: Voice monitor–Sound pressure level–Instrumentation.

Introduction

The ability to track aspects of the voice in natural communication exchanges and over extended periods outside the laboratory has been of interest to researchers and clinicians for several decades. There are a number of issues that complicate the matter including privacy of study participants and those with whom they might interact, ability to detect the subject's voice without contamination from other people or the environment, and a host of technical issues that impose limits on what aspects of the voice can be accurately tracked. Methods of detecting the voice have included use of noise exposure analyzers,[1] contact microphones and accelerometers[2–7] and electroglotography.[8] Some devices have been described that log data over time, others that provide real-time feedback to the user, and still others doing both.

The extent to which the development, capabilities, and limitations of the vocal accumulators and dosimeters has been described in the literature has varied from very little to quite comprehensive. Devices developed and used by Titze et al,[6,9,10] Hillman et al,[2,3] and Szabo et al[7] have been described in some detail allowing readers the opportunity to understand and assess the principles on which the instrumentation and software are based, the validity of the measures obtained, and potential limitations or concerns.

Recently, a vocal monitor called the VocaLog (Griffin Laboratories, Temecula, CA) became available on the market. Currently, there are no details in the research literature

about its development or use. According to the manufacturer's Web site, the device can be used to track "loudness," and via realtime cuing, it might be useful for encouraging louder voice for individuals with Parkinson disease, or for reducing "voice levels" for those who presumably are speaking too loudly. The device's ability to detect phonatory activity and accurately track some metric of loudness, however, is not known. The purpose of this set of studies was to evaluate the VocaLog monitor (VM) from a variety of perspectives including accuracy in detecting voicing, unintended detection of other sounds or movements, user feedback regarding wearing the device, and so forth. Because we were not part of the original development of the device and because we are not involved in the manufacture or sale of the devices, the study was approached from the end-users' perspective.

Overview of studies, instrumentation, and participants

Several experiments were carried out over a number of months to address various issues as detailed below. The type and number of experiments was not specified a priori but rather evolved once we had the monitors on site and had the opportunity to explore their use.

The intention was to examine a number of VMs, not just the performance of a single unit. Each experiment included a minimum of four units and most often eight were used. Each VM comprised the monitor itself, contact microphone, docking station, USB cable, and the VocaLog software (Version 1.2.4.2). Components other than the VM itself did not bear a manufacturer stamp and specifications did not accompany the devices, so it is not clear if the accessories are produced by Griffin Laboratories or if they are available from another vendor. The docking station is used to charge the VM and also to allow connection to a computer via USB cable. Connecting the vocal monitor to a computer is necessary for calibrating the device, downloading data from the monitor, and clearing the memory of the monitor for reuse. The calibration process is described in the Appendix.

Many of the experiments involved comparing the dB data logged by the VM to dB SPL values from an external sound level meter (SLM). In all cases, the external SLM was a Checkmate CM160 (Galaxy Audio, Wichita, KS), which was positioned 30 cm from the mouth and set to fast and A weighting. This SLM has output capabilities and dedicated software that allowed display of the meter's dB SPL in real-time. This feature was used in some experiments wherein it was necessary for participants to monitor their speech.

In addition to comparing dB values from the VM to dB SPL values from the SLM, we also focused on the potential use of the VM as a means of tabulating total seconds of phonation. Although the manufacturer does not advertise the device for this purpose, such a measure seemed possible given that the VM output can be exported as a string of 1-second dB averages. Tallying the number of 1-second averages that were greater than 0 dB might give a reasonable estimate of the total number of seconds of phonation. To begin understanding how accurately the VM might be used to track total seconds of phonation, acoustic recordings were made simultaneously with the VM data collection using a headset microphone (Shure SM150; Shure Incorporated, Niles, IL) positioned 15 cm from the corner of the mouth and routed to one channel of a DAT recorder (Tascam DP-A1; TEAC America, Montebello, CA). At a later time, the acoustic recording from the DAT tape was routed to a desktop computer through a digital input port of an MI/ODI/ O connector attached to a Wave-terminal 192L sound card (ESI Audiotechnik, Leonberg, Germany). The digital signal was opened in Multi-Speech as a time-by-amplitude waveform and spectrographic display. The speech signal of interest was manually bracketed within the software for the purpose of summing various durations. Duration values derived from counting the

number of 1-second VM intervals that were >0 dB could then be compared with the duration measures from the acoustic recording.

All experiments were completed on normal adult participants. Approval from the Human Subjects Committee at the authors' home institution was obtained before executing all experiments. These adults were 12 graduate students that ranged from age 22 to 34 years (11 females and one male). All had normal hearing by self-report and normal speech and voice (again by self-report but also confirmed by informal observation of the investigators). We did allow subjects to participate in more than one experiment, and most did.

Methods

Issue #1: Detection of voicing
Experiment 1: Absence of dB logging when not phonating. The first issue was to determine whether the device registered 0 dB when the device was being worn but the participant was not phonating. For this purpose, four adults sat quietly in the laboratory with a calibrated VM on the neck. Participants were instructed not to talk or make other laryngeal sounds such as sighing and throat clearing. We also asked subjects to refrain from swallowing saliva for as long as possible; the intent was to log several minutes of sustained silence and at this point, we had not determined whether the VM detected swallowing as a dB value. Most participants could go for 2–3 minutes without swallowing. If they swallowed or voiced before a minimum of 2 minutes had passed, that run was excluded from further analysis. Participants completed 3–5 runs (ie, run = 2 >minutes of silence and no swallowing) so that a total of 10 minutes of silence without swallowing were logged with the VM in place. Data were downloaded after each run was completed. For each subject, the initial 10 + minutes of nophonation data were collected on one device and then the entire process was repeated using a second VM (eight monitors total).

Experiment 2: Presence of dB logging during phonation. The same four participants and eight monitors were used for experiment 2. After completing the silent runs, participants performed five repetitions of sustained phonation on the vowel /i/, each lasting 10 seconds and separated by 30 seconds of silence. The silence breaks were intended to result in a string of consecutive zeroes in the VM data stream to demarcate the runs of non-zero dB values associated with vowel productions. Five vowels were produced with the instruction to "use your typical conversational loudness." The 10-second phonation duration was enforced by using a digital clock displayed for the subject that cued them when to start, showed a countdown, and cued them to stop. Each then completed five additional sustained vowels under instruction to "be half as loud as you are in conversation." This was a direct magnitude estimation task where the participant was told that their conversational loudness was "100," and for the "half loudness" trials, they should produce speech at a "50."

All productions were audio recorded using the DAT recorder and microphone setup described above. The lead author and a laboratory assistant reviewed all recordings in Multi-Speech (model 3700; KayPENTAX, Montvale NJ) and verified by consensus that each of the "typical" and "half" loud sustained vowels from all subjects was produced without phonation breaks. They then reviewed the VM output from each vowel trial to confirm whether or not it logged consecutive dB values greater than 0 dB during each vowel run. Outcomes were reported as the percentage of vowel trials wherein the VM output accurately reflected continuous detection of the voice (ie, continuous dB values >0 during the vowel).

Experiment 3: Logging of dB for noncommunication activities. The same four participants and eight monitors were used for this portion of the testing. With a VM calibrated and the contact microphone positioned on the neck, the following behaviors were elicited: swallowing (saliva, cup drink of water; typical and "effortful" swallows), throat clearing, and coughing. Each subject completed 25 repetitions of each activity and then repeated the protocol with a different monitor in place. Throat clearing was a brief (<1 second) soft vocalization; coughing was a single cough. Burping was also evaluated but only with one participant who could consistently burp on command. This participant was instructed to produce a single burp of 1–2 seconds in duration and 50 trials were completed. Acoustic recordings of throat clearing, coughing, and burping were obtained using the same DAT recorder and microphone.

We also were interested to see if the VM logged dB values when the contact microphone was (1) moved while in contact with the neck, (2) tapped by a finger, (3) grabbed on two sides by the fingers, and (4) bumped by clothing. Twenty-five repetitions of each of these four activities were assessed for each subject and monitor. When sliding the microphone off-midline, care was taken to only touch the curved metal neck bar and not the contact microphone head. For the bump with clothing, the wearer of the VM rustled his/her collar in a manner so that the clothing contacted the back of the microphone as it might during routine movements throughout the day.

Issue #2: VocaLog dB estimates relative to SLM dB SPL estimates
Experiment 4: VM versus SLM Estimates. It was of importance to know whether the dB values logged by the VM are accurate. However, there were challenges that complicated this assessment. One is that the VM uses a contact microphone detecting vibration through the neck that is referenced, via the calibration process, to a SLM built into the VM. This precludes simply playing an acoustic recording of a calibration tone that can be simultaneously detected by the VM and a trusted external SLM. A human who is producing voice must be involved in the VM calibration process because a contact microphone is used. A second challenge is that we are unaware of a means by which the user can output dB values from the VocaLog system in real-time for either the contact microphone or the SLM built into the VM. If that were possible, one could simultaneously record output from the VM and from a calibrated external SLM. If the signals were properly synchronized, the values could be compared to see if they at least increased and decreased consistently in a manner that could be quantified. A third issue was that details are not provided about the SLM that is built into the VM for the purpose of calibrating the contact microphone. Issues such as the response frequency and weighting, for example, are not reported in the product literature.

Our approach was to at least determine whether a VM device that is calibrated according to manufacturer directions responds in a consistent manner to variations in SPL as indexed by an external SLM that we knew to be calibrated and working properly. Because we do not know the specifics of the built-in SLM to which the VM system is calibrated, we did not expect that dB values being logged by the VM would be equal to the dB SPL values of our external SLM. However, the readings from both should at least vary consistently and in the same direction as the readings from the external SLM. Toward that end, we did the following:

1. A participant (n =4) was seated in a sound booth with a calibrated VM in place.
2. An external SLM was positioned 30 cm from the subject's mouth. This SLM was calibrated in the laboratory using the CM-C200 Calibrator (Galaxy Audio).
3. The subject produced a sustained vowel (/i/) for 5 seconds three times each at six

targeted dB SPL levels. The targeted
4. levels were 55, 60, 65, 70, 75, and 80 dB SPL as detected and displayed by the external SLM at 30 cm distance. For this task, participants were instructed to hold the voice as steady as possible at the targeted level, and real-time visual feedback was provided via a computer monitor, which displayed the output of the external SLM. The order of the dB SPL targets was constant across subjects starting at the lowest value (55 dB SPL) up to the highest (80 dB SPL). A 30-second pause between each vowel at a given target level was enforced; during this pause, the subject did not produce any vocalization. The intent was to have the contact microphone register a string of $30 zeroes in the data log so that it would be clear during later data extraction where the sustained vowel trials were. After doing three trials of /i/ at a specified dB SPL level with a 30-second pause between trials, a 60-second pause without voicing was enforced before proceeding to the next dB SPL target level. The four subjects, all practiced the task before data runs and three of the four had strong vocal performance backgrounds, which assisted in being able to consciously and fairly precisely manipulate the voice.
5. Subjects completed a similar set of dB SPL recordings with the VM and the external SLM while reading "The North Wind and the Sun" at the six target levels. This paragraph provided approximately 25–35 seconds of speech. The insertion of silent pauses between the three readings was maintained for the readings (30-second silence between trials at a given dB SPL target and 60second pause before shifting to the next dB SPL target). For the reading task, the dB SPL level displayed on the computer screen that served as feedback to the subject varied much more than during the sustained /i/ task given the nature of the speech sample. On the computer screen displaying the real time dB SPL of the external SLM, horizontal lines were placed 2 dB below and 2 dB above the target. The subject was instructed to keep the dB SPL of their voice at a level as close to the target and within the delineated range as possible.

Issue #3: Detection of environmental noise
Experiment 5: Contact microphone open to the air. If the VM detects environmental noise and logs it as a dB value, this could contaminate the voicing data. Furthermore, the VM does not have an "off" switch so it continues to log 1-second dB averages even when it is not being worn. To assess whether unintended loggings occurred when the VM was not being worn, eight VMs were conventionally calibrated (four subjects, two VMs trialed on each subject). Each calibrated contact microphone and VM was then removed from the neck and placed inside a sound booth with the contact microphone suspended in the air. All eight contact microphones were suspended in a cluster with the microphone sensing surfaces oriented in the same direction; a foam wedge with cutouts allowed us to hold all eight contact microphones firmly and in the same orientation without any contact microphone touching another. The Checkmate CM160 SLM was place in the booth with its microphone within the cluster of suspended contact microphones. The output from the CM160 SLM was routed to a laptop for real-time display of the dB SPL detected inside the booth in direct proximity to the cluster of contact microphones. Also inside the booth was an audio speaker (JBL 150; Harman International, Northridge, CA) positioned 30 cm in front of the cluster of microphones. An audio signal was played through this speaker from a laptop/ audio mixer arrangement under the investigator's control outside the booth. With this ar-

rangement, the researcher could play an acoustic sample from the laptop, monitor the dB SPL inside the booth via the Checkmate SLM, and adjust the volume to targeted dB SPL levels using the mixer.

A 200 and a 400 Hz sine waves, respectively, were played through the speaker for 2 minutes continuously at the following dB SPL levels as detected by the external SLM positioned inside the booth: 55, 70, 85, and 100 dB SPL. At this point in our experience with the devices, we had learned that a firm tap on the contact microphone was detected and logged by the VM 100% of the time as a 1-second dB value in the data output. As a means of parsing the VM output between each playing of the sine wave, which had the potential to be simply a long string of 0 values, we used a series of three taps spaced 3 seconds apart on each contact microphone. These taps were done simultaneously for all eight contact microphones using a long wooden rod that a research assistant in the booth positioned across all microphones. After the third tap was delivered, we waited 30 seconds before playing the next sine wave. The research assistant who was sitting in the sound booth throughout the experiment was instructed not to vocalize, remain stationary while audio signals were being played, and wear ear protection throughout.

In addition to the sine waves, a 5-minute sample from President Obama's speech at the 2011 White House Correspondents Dinner (downloaded from the internet) was played back at two loudness levels. The selection from this speech included crowd noise (laughter, clapping, and background chatter), connected speech, and orchestral music and singing. Through successive adjustments of the mixer output and analysis of the output from the external SLM positioned in the booth, a volume output of the JBL speaker was set so that at least 75% of the audio sample registered between 60 and 80 dB SPL (mean dB for the 5-minute sample was 71 dB SPL at this mixer setting). The second playback level was set so that at least 75% of the audio recording registered between 85 and 103 dB SPL (mean =96 dB SPL) by the external SLM. Again, three taps and a 30-second pause were used to separate the two playback levels.

Experiment 6: Contact microphone on the neck. To determine whether the VM registered environmental noises while being worn, a participant sat silently in the sound booth with a calibrated, actively logging VocaLog unit and contact microphone in place. The 200 and 400 Hz tone and the president's speech sample were played into the sound booth as described above. The VM and contact microphone were then removed and replaced with a different VM and microphone. Data collection proceeded in this way until a total of eight different and newly calibrated VMs were evaluated, all on the same neck. Ear protection was worn by the subject throughout this testing to protect them from the very loudest dB SPL levels.

Issue #4. Using the device to log seconds of phonatory activity
 Experiment 7: Logging Seconds of Phonatory Activity. At the time of this writing, the manufacturers were not marketing the device as a means of measuring phonation time (PT). However, with the ability to export a data string of 1-second average dB values, we were interested in determining how precisely a measure of PT (or perhaps speaking time [ST]) could be obtained from the VM output. From a data string of 1-second dB averages, one might simply be able to add up the total number of cells with a value >0 to get a rough estimate of the seconds of PT.

TABLE 1.
Seconds of Detection of Phonatory Activity by the VM
When Phonation Was Not Occurring

Subject	VM	Seconds of >0 dB During Silence Trials (of 600 s)
1	1	0
1	2	0
2	3	3 (0.5%)
2	4	0
3	5	1 (0.2%)
3	6	0
4	7	0
4	8	1 (0.2%)
All combined		5 of 4800 s (0.1%)

Abbreviation: VM, VocaLog monitor.

Eight subjects participated. A VM contact microphone and data logger were calibrated on each person's neck while she/ he sat in a sound booth. A headset microphone was placed for an acoustic recording with the signal routed to the Tascam DAT recorder. The speech protocol was as follows:

> a. Sustained /i/ (15 second) 33 at conversational loudness with a 30-second silent interval between vowel trials.
> b. 60-second silent interval before moving to the reading task.
> c. Reading of "The North Wind and the Sun" passage at conversational loudness.
> d. 120-second silent interval before moving to the "half loudness" task.
> e. Repeating steps a–d using "half loudness."

As stated previously, we are unaware of a way to output the dB values of the VM in real-time, precluding the opportunity to sync the data to another signal such as the acoustic recording from the headset microphone. Our best approach for comparing phonation duration measures was to export the 1-second dB averages from the VM and identify each stimulus production in this data string using the imposed silent intervals between trials and tasks. The duration of phonation activity for each sustained vowel and for each read paragraph was tabulated by summing up the number of cells with non-zero dB values that corresponded with each speech attempt. This sum (seconds) was compared with duration measures from the acoustic recording obtained with the headset microphone. After importing the DAT recording into Multi-Speech, cursors were manually placed at the start and stop of each sustained vowel and the duration was noted to the nearest 10th of a second. The waveform was inspected to ensure that phonation breaks did not occur during the sustained vowel. For the paragraph readings, two duration measures were obtained from the acoustic signal. The first was a measure of speech duration for the full paragraph with all pauses included (ST). The second was a measure of PT for which pauses and voiceless phonemes were excluded.

Issue #5: Impact of contact microphone shift on dB estimates
Experiment 8: Microphone Shifting Impact on dB Estimate. In our initial trials using the VM on lab personnel, it was noted that the microphone generally stayed in position on the neck

during routine movements. However, for individuals with smaller diameter necks, there was a greater tendency for the contact microphone to rotate off-midline. Upward or downward migration of the microphone position was not observed. Even for necks that were not thin, some minimal rotation (less than 1–2 cm) occurred occasionally. The flexible metal bar that wrapped around the neck to hold the microphone in contact with the skin could be bent to more firmly hold it in place, and this helped significantly in reducing movement. Nonetheless, we were interested in determining the impact of contact microphone position on dB values logged by the VM.

TABLE 2.
Percent of Sustained Vowel Trials Confirmed to be Free of Phonation Breaks That Were Detected as Continuous Phonation in the dB Output of the VM

Subject	VM	"Typical" Loudness (5 Trials per Subject per VM)	"Half" Loudness (5 Trials per Subject per VM)
1	1	100	100
1	2	100	100
2	3	80	100
2	4	100	100
3	5	100	100
3	6	100	100
4	7	100	100
4	8	100	100
All combined		97.5 (39/40)	100 (40/40)

Abbreviation: VM, VocaLog monitor.

We tested four VMs on four adults who produced sustained vowels (33) and read a passage at their typical and soft loudness with the contact microphone positioned in midline (calibration site) and at 1 cm increments off of midline (arbitrarily chosen to the right). A maximum displacement of 3 cm was tested. For these recordings, the individual was seated in the sound booth with the contact microphone on the neck and the VM calibrated. The SLM (Checkmate CM-160) was positioned 30 cm from the speaker's mouth. With the contact microphone in midline, the participant produced sustained vowels and read a paragraph in their conversational voice following steps a–e in experiment 7. This sequence was repeated with the contact microphone shifted 1, 2, and 3 cm from midline.

It was important for a participant to maintain a relatively constant dB SPL level throughout the conversational loudness productions. To help them in this task, the output of the Checkmate SLM was displayed on a computer screen in their view. For the conversational loudness vowels, they were instructed to say the vowel in their talking voice while watching the computer screen. They were instructed to keep the displayed value as constant as possible. For the paragraph reading, they were asked to keep the displayed value as close to the value that they had during the sustained vowel runs. They practiced this task and were informed that because of pauses and normal speech inflection that the dB SPL value would be changing more during the reading than the sustained vowels. In addition to displaying immediate feedback, the software for the Checkmate SLM was engaged to record the dB SPL throughout all speech trials.

The average dB value for each vowel and for each reading passage as logged by the VM was computed from the 1-second dB string exported from the VM. For comparison, output from the external SLM was displayed, the vowels and passages were bracketed manually by cursors and average dB SPL values for each production were derived using the Checkmate software.
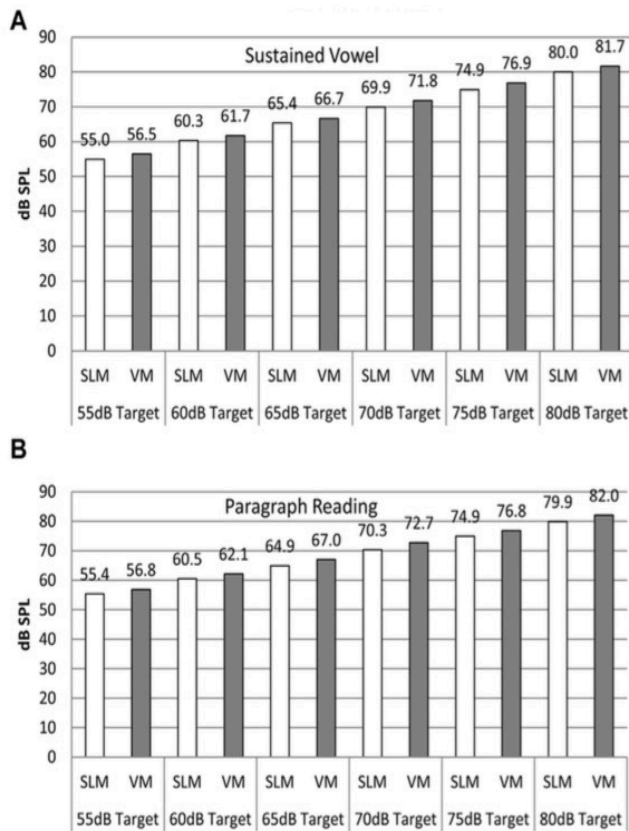
FIGURE 1. **A** and **B.** Group mean dB SPL values from a SLM and VM during sustained vowel production (**A**) and paragraph reading (**B**) at six target dB SPL levels. SLM, sound level meter; VM, VocaLog monitor.

Results

Experiment 1: Absence of dB logging during nonphonation

Five of eight VMs recorded 0 dB averages for all 600 seconds (10 minutes) when subjects were not phonating (Table 1). Although infrequent, the remaining three VM units recorded at least one 1-second dB average >0 dB suggesting phonation when none was occurring. A logging of a dB value >0 when no phonation was occurring could be considered a spurious or unexplained detection of phonatory activity. For this set of eight VMs, there were five 1-second averages of 4800 seconds (0.1%) that could be considered errors (false positives) in detection of voicing by the VM.

Experiment 2: Presence of dB logging during phonation

The purpose of this experiment was to answer the basic question: does the VM detect phonatory activity, as reflected in the devices dB log, when phonation is known to be present in a sustained vowel? Table 2 presents results for each subject and VM. Of the 40 sustained vowels at "typical" loudness level, 39 (97.5%) had output indicating continuous detection of phonation by the contact microphone. For the 40 vowels produced at "half" loudness, all had a continuous string of nonzero dB values in the VM output.

TABLE 4.
VM Detection of Environmental Noise With the Contact Microphone

	Microphone Not on Neck					Microphone on Neck				
		% Time With Logged dB Value		dB When Nonzero dB Was Logged			% Time With Logged dB Value		dB When Nonzero dB Was Logged	
	No. of Mics.					No. of Mics.				
Stimulus	Logging dB > 0	Mean	SD	Mean	SD	Logging dB > 0	Mean	SD	Mean	SD
Speech (dB SPL)										
60–80	0	0.0	0.0	—	—	0	0.0	0.0	—	—
85–105	8	14.3	5.1	55.4	1.4	2	0.2	0.3	58.0	4.2
400 Hz (dB SPL)										
55	0	0.0	0.0	—	—	0	0.0	0.0	—	—
70	0	0.0	0.0	—	—	0	0.0	0.0	—	—
85	0	0.0	0.0	—	—	0	0.0	0.0	—	—
100	0	0.0	0.0	—	—	0	0.0	0.0	—	—
200 Hz (dB SPL)										
55	0	0.0	0.0	—	—	0	0.0	0.0	—	—
70	0	0.0	0.0	—	—	0	0.0	0.0	—	—
85	5	2.6	3.6	56.3	6.1	0	0.0	0.0	—	—
100	2	0.8	1.8	60.7	10.4	0	0.0	0.0	—	—

Notes: Data are reported for trials with the contact microphone not on the neck and then on the neck.
Abbreviations: SD, standard deviation; VM, VocaLog monitor; Mics., Microphones.

Experiment 3: Logging of dB during noncommunication activities The VM logged 1-second dB averages greater than 0 for several non-communication activities (Table 3). Throat clearing and coughing registered as an event with a non-zero dB value 94% and 98% of the time. Burping registered as a dB event nearly 75% of the time.

The frequency of detection of swallowing as a dB event varied across tasks from as low as 9% (saliva swallow) to 66% (effortful water bolus swallow). The detection of swallowing appeared to be subject-dependent. Four of eight subjects registered 4% or less of each swallowing activity as a dB event. The remaining four subjects accounted for the large majority of swallow events logged as a non-zero dB value.

Intentional tapping on the microphone always registered as a non-zero dB value, whereas simply grabbing the side of the microphone or bumping it with clothing rarely did. Sliding of the microphone did result in logged events roughly one third of the time.

TABLE 5.
Sustained Vowel Duration as Estimated by the VM Output Compared With Measurement From the Acoustic Signal

Loudness Level	VM	Multi-Speech Phonation Time	t Value	P
Conversational	5.6 (1.9)	4.6 (1.7)	8.674	<0.000
Soft	5.5 (2.1)	4.7 (1.8)	5.900	<0.000

Abbreviation: VM, VocaLog monitor.

Experiment 4: VocaLog dB estimates relative to SLM dB SPL estimates
Figure 1A displays the dB SPL values from the external SLM and the VM during sustained vowel productions at 6 dB SPL target levels. The values from the VM and SLM differed by 1.3–1.9 dB across the 6 dB SPL target intervals (55–80 dB SPL). In all cases, the VM values were higher than the SLM (2.0–2.8% differences). A Pearson product moment correlation coefficient evaluating the strength of the relationship between values from the SLM and VM resulted in an r =0.999 indicating a strong, positive relationship between the two measures. A similar data display of dB values for the reading passage is shown in Figure 1B. The differences in dB SPL values between the SLM and VM were somewhat larger during the reading passage, ranging from 1.5 to 2.4 dB SPL (2.7–3.5%) with an r =0.999.

Experiment 5: Environmental noise detection— contact microphone open to the air
With calibrated VM microphones suspended in air (not on a neck), the 200 and 400 Hz wave and presidential speech recording did not register as dB values by the VMs when the stimuli were played at 80 dB SPL (Table 4). All eight VMs logged dB values during a portion of the speech sample when it was played back at the 85–103 dB SPL level. The mean duration for which dB values was logged was $14% of the speech (ie, 42 of 300 seconds) and the mean dB of those logged values was 55 dB. A few of the VM microphones detected the 200 Hz wave when played at the 85 and 100 dB SPL levels (<3% of the duration of the sine wave playback time).

Experiment 6: Environmental noise detection— contact microphone on the neck
With calibrated VM microphones on the neck, the only instance of dB detection of environmental noise by the VM occurred during playback of the speech sample at the 85–103 dB SPL level. Two of the eight VMs logged dB values during this stimulus production, representing 0.2% of the total seconds recorded across all VMs in this condition.

TABLE 6.
Paragraph Reading Duration as Estimated by the VM Output Compared With Measurement of PT and ST From the Acoustic Signal

Loudness Level	VM	PT	ST	F	P	Post Hoc Comparisons: VM Versus ...	P
Typical	34.8 (5.1)	19.5 (2.7)	36.0 (4.6)	46.31	<0.000	PT	<0.000
						ST	0.986
Soft	33.5 (1.5)	15.9 (2.5)	32.6 (2.7)	192.091	<0.000	PT	<0.000
						ST	0.974

Abbreviation: VM, VocaLog monitor.

Experiment 7: Using the device to log seconds of phonatory activity
Vowel durations as estimated from the VM output file differed from the vowel durations measured using Multi-Speech as evidenced by statistically significant t values for both conversational loudness (t =8.674, P < 0.000) and half-loud trials (t =5.900, P < 0.000). Based on group mean durations, the VM estimates of sustained vowel duration were approximately 20% longer than what was measured from the acoustic signal (Table 5).

A one-way analysis of variance (ANOVA) was statistically significant when the paragraph was read at a conversational loudness (F(3,28) =46.313, P < 0.000). The duration estimate from the VM was significantly longer than the PT measure from the acoustic signal (by $78%) but did not differ from the ST (Table 6). Results were similar for the paragraph reading done at half-loudness (F(3,28) =192.093, P < 0.000), wherein the VM duration was significantly longer than the PT measures (by $200%) but not the ST measure.

TABLE 7.
Post Hoc Testing Comparing the Percent Difference in dB SPL Measure From the VM Relative to the External SLM When the Contact Microphone is at Midline Versus 1, 2, and 3 cm Off-Midline

Midline Position Versus …	Vowel		Paragraph	
	Percent Difference	P	Percent Difference	P
1 cm off	2.25	0.999	3.53	0.406
2 cm off	6.70	**0.034**	6.20	**0.015**
3 cm off	13.55	**<0.000**	10.49	**<0.000**

Experiment 8: Impact of contact microphone rotation on dB estimates
This experiment involved subjects producing several trials of vowels and reading passages that could vary in dB SPL from one trial to the next despite instructions to use their conversational loudness. We first evaluated whether the subjects did, indeed, keep their intensity level constant across trials by looking at the dB SPL measure from the SLM positioned at 30 cm from the mouth for the vowel productions when the contact microphone was positioned midline and at 1, 2, and 3 cm offmidline. A repeated measures one-way ANOVA comparing the mean dB SPL from the SLM across four contact microphone positions resulted in a nonsignificant $F_{(3,29)}$ value of 0.317, $P =0.816$ indicating that subjects did keep dB SPL relatively constant across their vowel trials. Similarly, for the paragraph reading, the dB SPL from the SLM did not differ ($F_{(3,29)} =2.176$, $P =0.113$).

To determine whether VM estimations of dB differed as the contact microphone was positioned in increments off of midline, we calculated a percent difference in dB between the VM and SLM for each trial for each participant (percent difference =([VM dB _ SLM dB SPL] _ 1) 3 100). The percent difference was then used as the dependent measure in a repeated measures one-way ANOVA with contact microphone position as the independent variable. For the vowel productions, this resulted in an $F_{(3,28)} =14.218$ and $P < 0.000$. Post hoc testing that compared the percent difference at midline versus 1, 2, and 3 cm offmidline, respectively, is shown in Table 7. Figure 2A and B show the percent difference data at each contact microphone location for the vowels and the paragraph reading.

Discussion

The purpose of these experiments was to evaluate the performance of the VM in several ways to better establish its capabilities and limits. Results from the nine experiments can be summarized as follows:

1) The VM rarely logged a dB value during non-PT (experiment 1). Only one-tenth of 1% of the time tested was a spurious dB value logged during wearing of the device when a participant was intentionally being quiet.
2) The VM consistently logged a dB value during sustained phonation (experiment 2). In only one sustained vowel trial of 80, did the VM show a break in phonation detection (as reflected by a 0-dB interval within a sustained vowel trial confirmed to be free from phonation breaks).
3) Certain nonspeech activities showed up consistently as a dB value within the data log of the VM. Nearly all coughs and throat clears and most burps were logged as a dB values during a given period of data collection.

4) The dB values logged by the VM had a very strong, positive correlation with dB SPL values from an external SLM for both sustained vowels and paragraph readingThis suggests that the VM at least is tracking dB across a range of voice loudness   levels in a manner similar to a trusted device (ie, the calibrated external SLM).

5) When the VM's contact microphone was off the neck, it did occasionally log environmental noise (in this case, playback of a recorded speech) when the intensity of the environmental noise fell within the 85–103 dB SPL range. When the speech was played at this higher dB SPL level, approximately 14% of the time the VM registered a dB value. This is of some importance because it is possible that someone wearing the device will remove the contact microphone from their body during the day and at night. If loud noises, speech, or music are present, the VM log may show dB values even when the microphone is not being worn. Unless a means of tracking when the device is and is not on the neck is incorporated into its use, a practical solution might be to simply unplug the contact microphone from the data logger when the device is not being worn.

6) When the contact microphone was worn on a user's neck, the VM detected external sound or speech only rarely. For six of eight VMs, the external noise and speech was never logged as a dB value even when played at the 85–103 dB SPL range. Two of the VMs did register some external sounds as dB values although this occurred only 0.2% of the time during the speech played at the 85–103 dB SPL level. This provides a fair level of assurance to the clinician or researcher that when the contact microphone is known to be on the neck and non-zero dB values appear in the VM data log, there is only a small chance that vthe logged dB value is from another person or the environment.

7) The VM overestimated the seconds of PT when compared with PT measured from the acoustic signal. For sustained vowels, the overestimation was 20% and increased for the reading passage to nearly 80%. Interestingly, the estimation of seconds of PT from the VM did not differ significantly from a measure of ST from the acoustic signal. Recall that the ST measure included pauses and unvoiced speech time as well as voiced speech time. Presumably, the VM will calculate an average dB and report it as a 1-second dB average even if part of that 1-second window does not involve phonation. In addition, the current results suggest that the VM data string does closely estimate ST.

8) In trials assessing the effect of shifts in the VM microphone's position on the neck, the VM's dB estimates departed further and further from an SLM referent as the contact microphone shifted further away from midline. When the contact microphone was shifted 1 cm off the calibration point, the dB estimate from the VM did differ by about 2% from the SLM dB SPL estimate, but this was not a statistically significant difference. Presuming that the same result would occur if the microphone was shifted 1 cm in the opposite direction (recall that we only shift tested to the right), then it may be the case that there is about a 2 cm position on the neck that will give a reasonably close estimate to what would be measured if the microphone stayed at midline (ie, the calibration point).
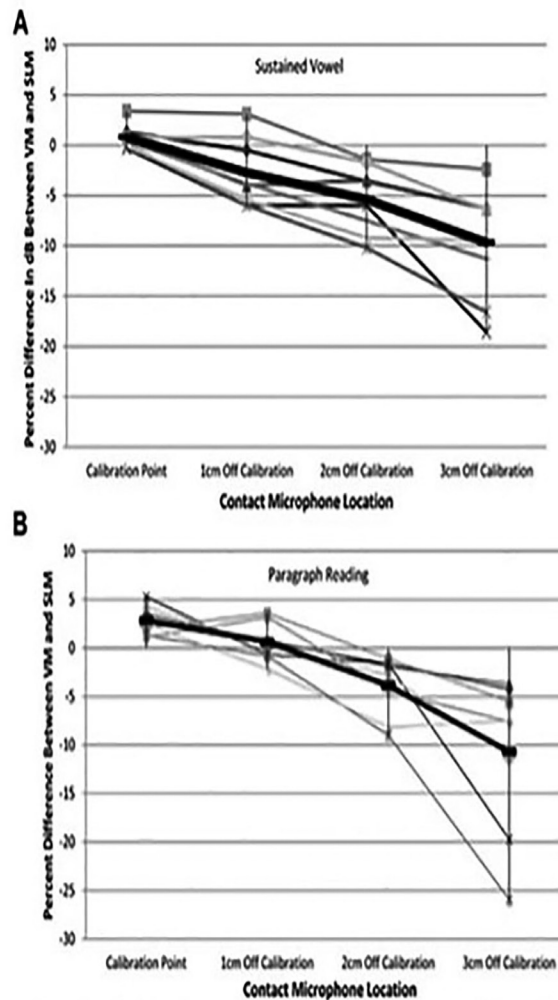
FIGURE 2. A and B. Percent difference in dB values for sustained vowels (A) and paragraph reading (B) as estimated by the VM and the external SLM at each contact microphone position. Data are plotted for the individual VMs, and the average percent difference is plotted as the thicker black line. SLM, sound level meter; VM, VocaLog monitor.

It was our intention to provide some guidance as to the capabilities of the VM for those who are considering its use in the absence of any other reports in the literature or from the manufacturer. The fact that we are not privy to the development, manufacture or sale of the device limits the extent to which we can understand how the device is performing. However, this set of studies provides at least some initial data about how the VM functions.

One other aspect of the VM was not incorporated into our testing. The device purportedly has the ability to use a dB threshold setting which, when crossed, triggers either an audible beep, a tactile vibration of the data logger, or both. The feedback settings and capabilities of the VM were not addressed at all in our testing but do deserve attention if clinicians or researchers are interested in using this function.

Acknowledgments

The authors extend our appreciation to Amy Seeland who helped with the project while a graduate student and to Drs Kelly Lyons and Rajesh Pahwa who helped recruit the subjects with Parkinson disease.

References

1. Airo E, Olkinurora P, Sala E. A method to measure speaking time and speech sound pressure level. Folia Phoniatr Logop. 2000;52:275–288.
2. Cheyne HA, Hanson HM, Genereux RP, Stevens KN, Hillman RE. Development and testing of a portable vocal accumulator. J Speech Lang Hear Res. 2003;46:1457–1467.
3. Hillman RE, Heaton JT, Masaki A, Zeitels SM, Cheyne HA. Ambulatory monitoring of disordered voices. Ann Otol Rhinol Laryngol. 2006;115: 795–801.
4. MasudaT,IkedaY,ManakoH,KomiyamaS.Analysisofvocalabuse:fluc tuations in phonation time and intensity in 4 groups of speakers. Acta Oto laryngol. 1993;113:547–552.
5. Ohlsson A, Brink O, Lofqvist A. A voice accumulator—validation and application. J Speech Hear Res. 1989;32:451–457.
6. Svec JG, Popolo PS, Titze IR. Measurement of vocal doses in speech: experimental procedure and signal processing. Logoped Phoniatr Vocol. 2003;28:181–192.
7. SzaboA,HammarbergB,KakanssonA,SoderstenM.Avoiceaccumulator device: evaluation based on studio and field recordings. Logoped Phoniatr Vocol. 2001;26:102–118.
8. Svec JG, Titze IR, Popolo PS. Estimation of sound pressure levels of voiced speech from skin vibration of the neck. J Acoust Soc Am. 2005;117:1386–1394.
9. Hunter EJ, Titze IR. Variations in intensity, fundamental frequency, and voicing for teachers in occupational versus nonoccupational settings. J Speech Lang Hear Res. 2010;53:862–875.
10. Titze IR, Svec JG, Popolo PS. Vocal dose measures: quantifying accumulated vibration exposure in vocal fold tissues. J Speech Lang Hear Res. 2003;46:919–932.

Appendix

VocaLog calibration procedures
1. VocaLog docking station is plugged into the wall outlet and connected (USB) to a laptop.
2. VocaLog monitor (VM) is placed in the docking station.
3. VocaLog software is opened and the ''Connect to VocaLog Unit'' option is chosen. A subchoice to ''Initialize'' the unit is selected to clear the memory and prepare it for use.
4. The ''Calibrate'' tab is selected and the procedure for ''soft-spoken'' is chosen.
5. Calibration requires the subject to produce the lowest audible sustained phonation and also a louder sustained phonation (both on /i/). The contact microphone is

detecting vibration on the neck, whereas the SLM built into the VM is detecting dB SPL. In the calibration window, a slider bar shows the dB SPL level as detected by the built-in SLM. Likewise, activity detected by the contact microphone on the subject's neck is also displayed as a slider bar. During a steady state portion of the sustained vowel, the researcher clicks the mouse to capture a pair of data: the dB SPL level from the built-in SLM and the activity from the contact microphone. Details are not provided but presumably a mathematical function is calculated that relates the dB SPL from the built-in SLM to the voltage activity from the contact microphone.

6. The calibration process is completed by finishing out with a few selections such as whether or not feedback is enabled (it was not for any portion of our testing), delay interval before providing feedback, and so forth. Once the selections are made, the VM is removed from the docking station and it begins logging 1-second average dB values referenced to the dB SPL of the built-in SLM.