

Research Note

Accuracy of Self-Reported Estimates of Daily Voice Use in Adults With Normal and Disordered Voices

Daryush D. Mehta,^{a,b,c} Harold A. Cheyne II,^d Asa Wehner,^a
James T. Heaton,^{a,b,c} and Robert E. Hillman^{a,b,c}

Purpose: Accurate estimation of daily patterns of vocal behavior is essential to understanding the role of voice use in voice disorders. Given that clinicians currently rely on patient self-report to assess daily vocal behaviors, this study sought to assess the accuracy with which adults with and without voice disorders can estimate their amount of daily voice use in terms of phonation time.

Method: Eighteen subjects (6 patients, 6 matched members of a control group without voice disorders, 6 low voice users) wore the accelerometer-based Ambulatory Phonation Monitor (APM; model 3200, KayPENTAX, Montvale, NJ) for at least 5 workdays. Subjects were instructed to provide hourly self-reports of time spent talking using a visual analog scale. Spearman correlation coefficients and errors between

self-reported and APM-based estimates of phonation time revealed subject- and group-specific characteristics.

Results: A majority of subjects exhibited a significant bias toward overestimating their phonation times, with an average absolute error of 113%. Correlation coefficients between self-reported and APM-based estimates of phonation time ranged from statistically nonsignificant to .91, reflecting large intersubject variability.

Conclusions: Subjects in all 3 groups were moderately accurate at estimating their hourly voice use, with a consistent bias toward overestimation. The results support the potential role that ambulatory monitoring could play in improving the clinical assessment of voice disorders.

Accurate estimation of voice use is important for understanding its role in the development of voice disorders, especially for professions requiring substantial voice use (Vilkman, 2004). Several voice-use-related measures have been developed to estimate ambulatory vocal function over the course of multiple days. Traditional measures of sound intensity, fundamental frequency, and phonation time have been proposed to yield daily phonation profiles, along with cumulative vocal-dose measures such as cycle dose and distance dose (Titze,

Švec, & Popolo, 2003). There have also been efforts to quantify how the durations of nonphonatory times distribute across entire days and longer, to begin attaining a better understanding of the potential role of “recovery time” in mitigating the occurrence of common voice disorders (Titze, Hunter, & Švec, 2007).

Clinicians can obtain information about the impact of voice disorders on the daily function of patients through the use of standardized self-report inventories such as the Voice-Related Quality of Life (Hogikyan & Sethuraman, 1999) and Voice Handicap Index (Jacobson et al., 1997) assessments. Actual estimates of typical voice use, however, are still usually obtained as part of the nonstandardized clinical interview process, and the validity and reliability of such highly subjective estimates is questionable. On the basis of our clinical experience, we also wondered whether some patients with voice disorders may be even less reliable than normal speakers in self-reporting their voice use because of reduced self-monitoring or awareness that could conceivably contribute to the development of a

^aCenter for Laryngeal Surgery and Voice Rehabilitation, Massachusetts General Hospital, Boston

^bHarvard Medical School, Boston, MA

^cInstitute of Health Professions, Massachusetts General Hospital, Boston

^dBioacoustic Research Program, Lab of Ornithology, Cornell University, Ithaca, NY

Correspondence to Daryush D. Mehta: daryush.mehta@alum.mit.edu

Editor: Krista Wilkinson

Associate Editor: Preeti Sivasankar

Received July 20, 2015

Revision received January 27, 2016

Accepted April 18, 2016

DOI: 10.1044/2016_AJSLP-15-0105

Disclosure: Robert E. Hillman has a financial interest in the Ambulatory Phonation Monitor (APM) based on a contractual agreement between Sensimetrics, Inc. (R & D for the initial version of the APM) and KayPENTAX, Inc. (manufacturer of the APM).

voice-use-related disorder. This is somewhat supported by anecdotal observations indicating that psychosocial factors may play a larger role in self-reports by people with voice disorders than those without (Watanabe, Shin, Oda, Fukaura, & Komiyama, 1987).

In a group of 12 teachers with normal voices, Rantala and Vilkmán (1999) found a moderate correlation coefficient of .66 between self-reported ratings of nonspecific “voice complaints” and an acoustic-based measure of cycle dose. Buekers, Bierens, Kingma, and Marres (1995) observed, however, that speakers with normal voices tended to overestimate acoustic-derived measures of vocal intensity, although no details regarding specific rating scales or statistical evidence were given. Of direct relevance to the current study, Ohlsson, Brink, and Löfqvist (1989) reported large discrepancies between self-reported speaking time (62%–71%) and phonation time derived from a contact-microphone-based ambulatory voice monitor (5%–7%); in that study, however, only speakers with normal voices were monitored, and a quantitative statistical analysis between self-reported and voice-monitor data was not carried out. In more recent work, Misono, Banks, Gaillard, Goding, and Yueh (2015) report that phonation times measured with a VocaLog (Griffin Laboratories, Temecula, CA) voice dosimeter correlated moderately ($r = .62$) with self-reported estimates of voice use by subjects with normal voices. However, subjects were monitored for only 2 hr, and there is some evidence that the VocaLog tends to overestimate phonation times (Van Stan, Gustafsson, Schalling, & Hillman, 2014).

As noted in several studies, caution is warranted when comparing results from different voice-use investigations to distinguish between speaking time and phonation time. Speaking time is typically estimated with ambulatory monitors that use microphones to capture the acoustic speech signal that includes voiced, unvoiced, and silence segments (Airo, Olkinuora, & Sala, 2000; Buekers et al., 1995; Sala et al., 2002; Södersten, Granqvist, Hammarberg, & Szabo, 2002). Phonation time has been studied in ambulatory settings using both neck-placed contact microphones (Ohlsson et al., 1989; Ryu, Komiyama, Kannae, & Watanabe, 1983; Szabo, Hammarberg, Håkansson, & Södersten, 2001; Watanabe et al., 1987; Watanabe, Komiyama, Ryu, & Kannae, 1984) and accelerometers (Cheyne, Hanson, Genreux, Stevens, & Hillman, 2003; Hillman, Heaton, Masaki, Zeitels, & Cheyne, 2006; Mehta, Zañartu, Feng, Cheyne, & Hillman, 2012; Popolo, Švec, & Titze, 2005; Titze et al., 2007) that respond primarily to neck skin vibration generated during phonation, with little activity during unvoiced speech sounds (Zañartu et al., 2009).

It should also be noted that, even though the technology for objective voice ambulatory monitoring has become commercially available in recent years (Van Stan et al., 2014), it has not been adopted for routine clinical use. This is at least partly due to a lack of compelling evidence that such monitoring can provide information that could improve the clinical voice-evaluation process. The need for such evidence was a major motivation for the current study.

The purpose of the current study was to assess the accuracy of self-reports of voice use by comparing individuals' estimates of their speaking times with measures of phonation time derived from an accelerometer-based ambulatory phonation monitor. We also sought to attain an initial sense of whether individuals with and without vocal-fold pathologies working in high-voice-use occupations differ from each other in the accuracy of their self-reported speaking time and whether individuals in low-voice-use occupations differ from those in high-voice-use occupations in the accuracy of their self-reported speaking time. In the current study, subjects were simply asked to estimate their speaking time, because voicing and phonation time were considered too conceptually abstract or challenging to directly estimate. Speaking-time estimates were then converted to estimates of phonation time to compare them with the accelerometer-based estimates of phonation time.

The results of this study should provide better insight into whether subjective self-reporting of voice use is associated with sufficient error to affect clinical decisions about the role of voice use in common voice disorders. The presence of such error would also lend support for the clinical use of ambulatory monitoring to improve the assessment of voice use.

Method

Subject Characteristics

Table 1 reports subject characteristics. The patient group comprised six individuals who had vocal-fold pathology associated with phonotrauma (vocal-fold nodules or polyps) before any treatment was administered. Each patient was in an occupation requiring substantial voice use and was asked to identify a colleague of the same sex in the same profession and as close to the same age as possible to serve as a matched control-group subject and yield a control group of six subjects. The patient and control data were gathered as part of a larger ongoing project addressing the role of voice use in phonotraumatic voice disorders, where this type of matching is essential—only this subset of that much larger subject group was asked to self-assess their amount of voice use. An additional six subjects were recruited from laboratory research staff to comprise a low-voice-use group. Because ambulatory voice monitoring was still relatively new, this group also served as an overall reality check for the study protocol and monitoring process (i.e., the lack of a significant difference in measured phonation times between this group and the groups of heavy voice users would signal a potential problem). All subjects underwent a laryngeal endoscopic examination by an experienced team (laryngologist and speech-language pathologist) to confirm the presence (in patients) or absence (in the control group and low voice users) of vocal-fold nodules prior to participation and group assignment. This study was preapproved by the Institutional Review Board of the Massachusetts Eye and Ear Infirmary.

Table 1. Subject identifier, age (in years), and occupation of subjects in the patient, control, and low-voice-user groups, with patient diagnoses.

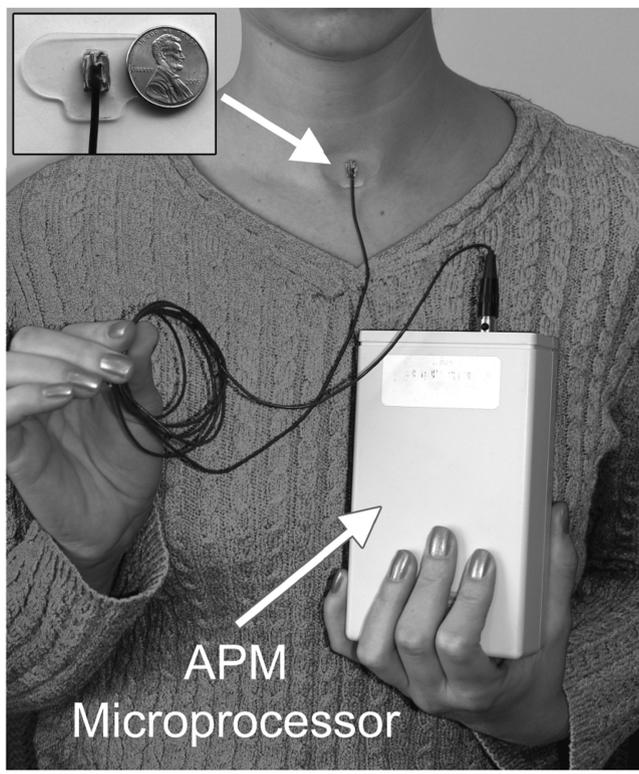
Sex	Patient	Control	Low Voice User
F	P1, 37, Nodules, Teacher	C1, 49, Teacher	
F	P2, 22, Nodules, Salesperson	C2, 23, Salesperson	
F	P3, 46, Nodules, Clinician	C3, 52, Clinician	LVU3, 52, Researcher (Histologist)
F	P4, 54, Nodules, Teacher	C4, 62, Teacher	LVU4, 59, Researcher (Histologist)
M	P5, 19, Polyp, Singer	C5, 20, Singer	LVU5, 29, Researcher (Engineer)
F	P6, 30, Nodules, Clinician	C6, 26, Clinician	LVU6, 29, Animal-care technician
M			LVU1, 31, Researcher (Physiologist)
M			LVU2, 48, Researcher (Engineer)

Note. Subjects appearing in the same row were matched by sex and approximate age.

Data Collection

Each subject wore an Ambulatory Phonation Monitor (APM; model 3200, KayPENTAX, Montvale, NJ) for at least five workdays. Figure 1 illustrates the position of a miniature accelerometer (model BU-27135, Knowles Electronics, Inc., Itasca, IL) attached to the front of the

Figure 1. Photograph showing a typical accelerometer attachment location on the neck surface above the sternal notch, with wire connection to the Ambulatory Phonation Monitor microprocessor box. The inset provides a magnified view of the accelerometer next to a U.S. penny for size comparison. Adapted from “Ambulatory Monitoring of Daily Voice Use,” by R. E. Hillman and D. D. Mehta, 2011, *SIG 3 Perspectives on Voice and Voice Disorders*, 21, p. 59. Copyright 2011 by the American Speech-Language-Hearing Association. Adapted with permission.



neck using a medical-grade adhesive (Skin-Bond, Smith & Nephew, London, United Kingdom; or model B-401, Factor II, Inc., Lakeside, AZ) just above the sternal notch to record neck-surface acceleration. Subjects were instructed not to alter their typical vocal behavior while being monitored. At the beginning of each day of data collection, the subject was fitted with the APM at the clinic or at the subject's place of work. An acoustic calibration routine was performed each morning to allow for the interpretation of acceleration amplitude in terms of sound pressure level (SPL; Cheyne et al., 2003).

After each hour of monitoring, subjects were instructed to fill out a voice-use diary to keep a written record of their activities throughout their workday. Subjects self-reported their “time spent talking” each hour by placing a mark on a visual analog scale ranging from *None* on the far left of the scale to *Continually* on the far right of the scale. In addition, the diary prompted subjects to specify the nature of the time they spent talking (percentage time devoted to face-to-face talking, talking to an audience, etc.). Subjects were, unfortunately, much less consistent in providing responses for these additional scales, so these data are not included in the current research note. Subjects were given the option of carrying a timer to provide an hourly reminder via an audible beep or tactile vibration. Visual analog scales were chosen on the basis of the design and rationale (e.g., ease of use) used in the Consensus Auditory-Perceptual Evaluation of Voice (Kempster, Gerratt, Verdolini Abbott, Barkmeier-Kraemer, & Hillman, 2009) and because we did not have any basis for determining how many subcategories or divisions to include in a numerical scale.

Data Analysis

The APM performed online processing of the accelerometer signal to yield estimates of SPL and fundamental frequency (f_0) every 125 ms (Cheyne et al., 2003). A voicing decision was made by first dividing each 125-ms frame into five 25-ms subintervals. If the root-mean-square values of two contiguous subintervals exceeded a preset threshold, the entire frame was considered voiced. For voiced frames, SPL was estimated by a linear function relating SPL to accelerometer root-mean-square level determined by the

acoustic calibration routine. An estimate of f_0 was computed from the largest peak in the autocorrelation function of the 125-ms frame. For unvoiced frames, SPL and f_0 values were set to 0.

Additional postprocessing constraints were placed prior to storing the parameters of a voiced frame: SPL had to be in the range of 50–130 dB SPL, and f_0 had to be in the range of 70–400 Hz. These limits were chosen on the basis of parameter ranges used in the commercial application of the APM and a widely used pitch-tracking software program (Computerized Speech Lab, KayPENTAX, Montvale, NJ). The limits sought to avoid frames that included spurious values that may have resulted from non-vocal activity (e.g., intense ambient noise and impulsive mechanical motion of the accelerometer) or vegetative sounds (e.g., laughing and coughing). The starting time index for each frame was saved in the APM data file to allow for synchronized comparisons between objective measures of vocal behavior and voice-use-diary information.

Only the hourly estimates of time spent talking were used in this study. Marks on the self-report scale for speaking time were converted to a percentage of the line length, with marks on the far left of the scale representing 0% and marks on the far right representing 100%. Some subjects chose to report their speaking time each hour as a percentage written above the visual analog scale instead of a mark on the scale; this practice was allowed. Because speaking time (termed *talking time* in the voice-use diary) consists of approximately 50% voiced speech and 50% unvoiced speech and silent periods (Fant, Kruckenberg, & Nord, 1991; Löfqvist & Mandersson, 1987), the self-reported percentages were divided by a factor of 2 to yield phonation-time estimates that could be compared with APM-derived phonation times.

Biases (average differences) between self-reported and APM-based phonation time were tested for statistical significance using paired t tests at a 95% significance level. Self-reported and APM-based phonation time per hour were converted to minutes per hour by multiplying percent phonation by 60 min. Covariance between self-reported and APM-based estimates of phonation time was examined using Spearman correlation coefficients for nonparametric data because of the non-Gaussian nature of the distributions of hourly phonation times. Tests for statistical pairwise differences were performed within each subject, between subject groups, and over all subject data pooled across all hourly data points.

Results

Total monitoring durations varied between 21 and 92 hr per subject. Compliance of subjects in completing their voice-use diary entries also varied. Some individuals wrote in detail about their activities during the day and faithfully reported an estimate of their talking time every hour. In contrast, other individuals were sparse in their descriptions and sometimes filled out their voice-use diary the following day. Subjects who wore the reminder timer reported that they tended not to use it after their first or second day

of recording, once they had become accustomed to making diary entries each hour. Due to compliance issues and intermittent technical difficulties, three subjects (P6, C4, LVU2) yielded only 21–23 hr of APM monitoring with corresponding self-report data. The remaining 15 subjects provided over 30 hr of data. Partial-hour data were not included.

Table 2 provides a comprehensive reporting of the ability of subjects to estimate their own voice use each hour. Correlation coefficients ranged from nonsignificant (P1, P2, C1, C4) to .91 (LVU5), reflecting large intersubject variability in the extent to which the measures covaried within subjects. The Bias column reflects any bias toward over- or underestimation of phonation time. Six subjects (P2, P3, C2, LVU1, LVU2, LVU5) did not exhibit a statistically significant self-report bias, indicating that they did not consistently under- or overestimate their voice use. The explained variance, however, was still typically less than 50% for these subjects. All individuals with a statistically significant bias (11 subjects) tended to overestimate their voice use (positive values of average bias), except for one (C1) who appeared to consistently underestimate her voice use. For these individuals, average absolute error ranges from 3.2 to 11.0 min/hr, with large variability indicated by standard deviations of the absolute error values. In terms of percentages, the absolute error ranges from 49.7% to 258.6%.

Group-based statistics in Table 2 show that patients and subjects in the control group were not significantly different from each other in terms of self-reported and APM-based phonation time. As expected, subjects in those two groups exhibited significantly higher voice use than subjects assigned to the low-voice-use group; the low voice users also reported significantly lower estimates of their phonation time. Pooling hourly data across all subjects, the overall absolute error is $113.3\% \pm 142.0$ percentage points (5.7 ± 5.0 min). These group-based results should be viewed as somewhat tentative given the small sample sizes.

Figure 2 displays the raw hourly data for the self-reported phonation times plotted against the corresponding APM-based phonation times for each of the three subject groups. Most data points lie above the unity-slope line, reflecting the fact that individuals in all three groups often tended to overestimate their phonation times. The consistent underestimation of phonation time by subject C1 is visible in Figure 2A.

Discussion

The goal of this study was to use improved ambulatory voice-monitoring technology to determine the accuracy with which three different groups of subjects could estimate their amount of hourly voice use (phonation time) over the course of a regular 5-day workweek. The three groups were patients with laryngeal pathologies in high-voice-use occupations; subjects with normal voices who were matched to the patients according to occupation, sex, and age; and

Table 2. Accuracy of voice-use estimation for each subject and subject group.

Subject ID	Duration (hr)	Self-report (min)	APM (min)	Correlation (ρ)	Bias (min)	Bias (%)	Absolute error (min)	Absolute error (%)
P1	37	21.9 (6.8)	16.5 (5.1)	n.s.	5.4 (7.2)	50.4 (94.9)	7.1 (5.5)	59.8 (89.0)
P2	34	9.9 (6.7)	7.6 (3.0)	n.s.	n.s.	n.s.	n.s.	n.s.
P3	40	10.5 (5.8)	9.9 (4.6)	.58	n.s.	n.s.	n.s.	n.s.
P4	40	18.2 (9.8)	11.6 (6.2)	.64	6.6 (7.6)	95.7 (126.3)	8.2 (5.8)	106.5 (117.1)
P5	40	4.9 (5.4)	2.9 (3.4)	.60	2.0 (4.7)	27.8 (177.0)	3.4 (3.7)	129.3 (122.2)
P6	23	17.8 (6.9)	10.5 (5.6)	.58	7.3 (5.6)	106.6 (111.3)	7.6 (5.1)	109.9 (108.0)
P group	214	13.5^b (9.2)	9.8^b (6.3)	.69	3.8^b (6.6)	52.9^{a,b} (120.2)	5.7^{a,b} (5.1)	85.9^{a,b} (99.3)
C1	37	11.3 (6.9)	17.0 (6.2)	n.s.	-5.7 (8.3)	-26.1 (48.4)	8.7 (4.9)	49.7 (22.5)
C2	40	9.8 (6.4)	9.0 (7.5)	.68	n.s.	n.s.	n.s.	n.s.
C3	40	18.5 (8.7)	8.4 (4.6)	.45	10.1 (7.7)	171.1 (191.9)	11.0 (6.4)	181.4 (182.0)
C4	21	23.9 (5.3)	16.1 (6.6)	n.s.	7.7 (7.4)	84.7 (121.0)	9.1 (5.6)	91.0 (116.1)
C5	40	13.5 (6.0)	4.5 (3.1)	.60	9.0 (5.0)	242.8 (212.4)	9.1 (4.9)	258.6 (192.3)
C6	32	9.3 (6.4)	6.7 (5.2)	.45	2.6 (6.3)	96.1 (148.4)	5.3 (4.2)	121.1 (128.1)
C group	210	13.8^c (8.2)	9.8^c (7.2)	.41	4.0^c (8.7)	105.6^a (183.5)	7.7^{a,c} (5.6)	133.4^a (164.3)
LVU1	30	5.5 (6.0)	5.9 (5.6)	.74	n.s.	n.s.	n.s.	n.s.
LVU2	21	6.9 (3.0)	7.4 (3.6)	.55	n.s.	n.s.	n.s.	n.s.
LVU3	40	7.5 (3.9)	3.8 (3.0)	.48	3.7 (2.6)	153.9 (178.4)	3.9 (2.3)	159.2 (173.6)
LVU4	40	6.9 (3.7)	5.0 (2.8)	.46	1.8 (3.7)	80.6 (175.0)	3.2 (2.6)	98.9 (165.1)
LVU5	40	6.4 (7.7)	5.9 (7.8)	.91	n.s.	n.s.	n.s.	n.s.
LVU6	36	13.5 (4.0)	5.6 (3.2)	.54	7.9 (3.9)	197.8 (143.9)	8.0 (3.6)	198.5 (142.9)
LVU group	207	7.8^{b,c} (5.7)	5.4^{b,c} (4.8)	.62	2.4^{b,c} (4.4)	97.0^b (167.6)	3.8^{b,c} (3.3)	121.2^b (150.9)
All subjects	631	11.8 (8.3)	8.4 (6.5)	.63	3.4 (6.8)	85.0 (160.6)	5.7 (5.0)	113.3 (142.0)

Note. Self-report and Ambulatory Phonation Monitor (APM) display mean (standard deviation [SD]) of self-reported (half of the judgment of talking time) and APM-based measures, respectively, of phonation time over the number of hours listed under Duration. Spearman rho correlation is computed for self-reported and APM-based phonation time (n.s. = not significant, $p \geq .05$). Bias quantifies mean (SD) of hourly estimation error by subtracting APM-based phonation time from half of the self-reported talking time, where the hourly error magnitude is used for computing the absolute error mean (SD). Boldface rows report mean (SD) after pooling all hourly data from patient (P), control (C), and low-voice-user (LVU) groups. Groups exhibiting statistically significant pairwise differences are marked by the same symbol: P-C (^a), P-LVU (^b), C-LVU (^c).

subjects with normal voices whose occupations required less voice use compared to the other groups.

There was a range of compliance associated with the task of self-reporting voice use that could clearly have affected the accuracy of these estimates. Although it might be possible to improve the accuracy of self-reported estimates by facilitating better task compliance (e.g., using additional incentives), it could also be argued that the observed variation mimics the type of variability in compliance that different patients typically display in following instructions related to their clinical care.

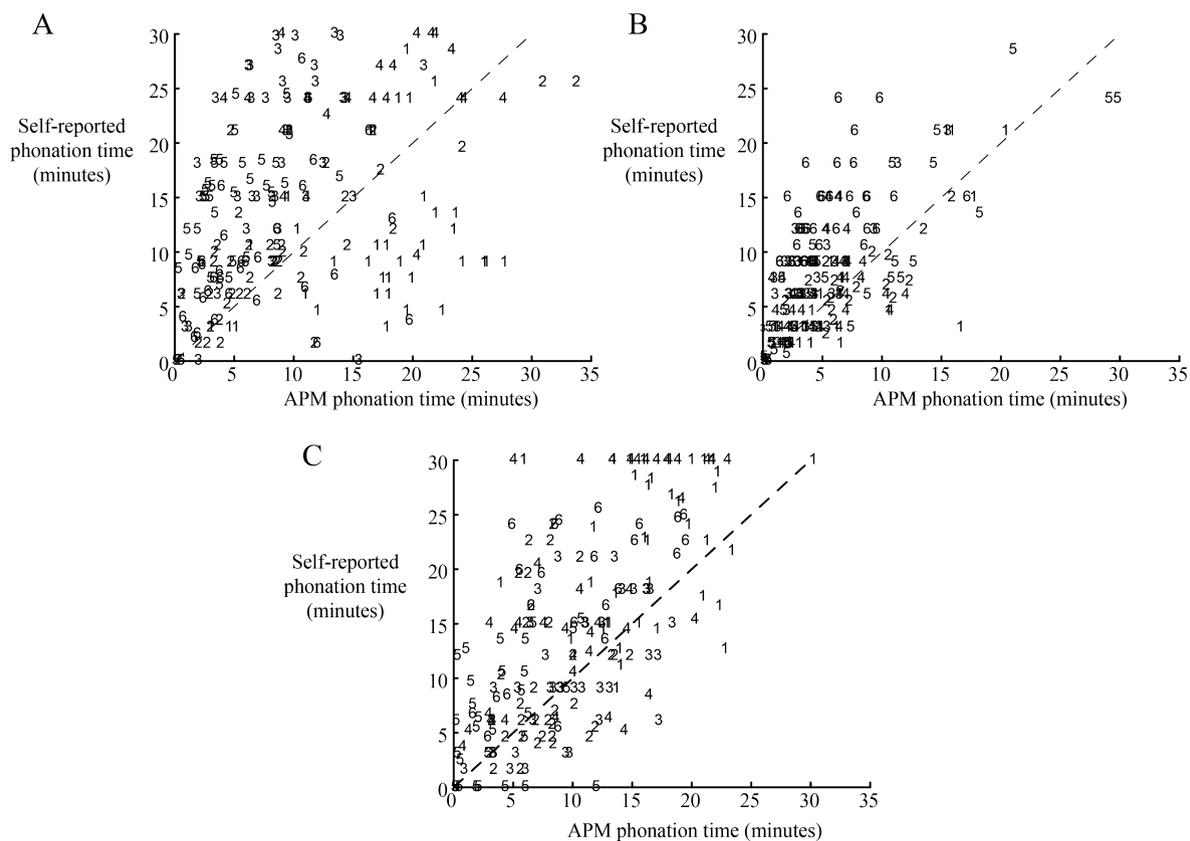
As expected, the low voice users spent less total time phonating per hour than did subjects in the patient or control groups (Table 2). It is unsurprising that the four pairs of teachers in the study (subject pairs C1-P1 and C4-P4) exhibited some of the highest average hourly phonation times. As a profession, teachers have been singled out as a group requiring extra care because of their higher risk for developing voice disorders (Roy et al., 2004; Vilkmann, 2004). Pooling data across all subjects (631 hr), the average APM-based phonation time was 8.4 ± 6.5 min/hr, corresponding to an average phonation time of 14.0%. This value is comparable to other APM-based values of phonation time for adults with normal voices reported by Watanabe et al. (1987; 11%) and Ohlsson et al. (1989; 8.6%) but less than the phonation times reported by Watanabe et al. (1984; 20%); Masuda, Ikeda, Manako, and Komiyama (1993; 21% for teachers, 22% for patients

with nodules); and Titze et al. (2007; 23% for teachers during the workday).

Valid assessment of whether individuals were accurate in their self-reported voice use could obviously be affected by how well phonation time was measured by the APM, particularly for disordered voices. This concern was somewhat mitigated by the results of our previous work comparing microphone- and accelerometer-based estimates of phonation time in normal and disordered voices (Hillman et al., 2006). That study provided evidence that the accelerometer signal was more robust than the microphone signal for analyzing severely disordered voices, which was attributed to the subglottal location of the accelerometer being largely immune to turbulent noise generation and vocal-tract resonances.

The average self-reported phonation time across all subjects was 11.8 ± 8.3 min/hr (Table 2), corresponding to an average hourly percent phonation time of 19.7%. Overall, individuals in the patient group tended to more accurately estimate their speaking times than did their control cohort, although the errors are large for all subjects, which casts doubt on the clinical validity of using patient self-reporting to assess voice use. Low voice users exhibited the lowest absolute error (3.8 ± 3.3) in terms of minutes per hour. Their estimation error, however, was significantly greater than that of patients when error was put in terms of percentages, probably owing to the inherently short phonation times that the low voice users exhibited. Moreover, although it is unknown to what degree the various groups

Figure 2. Scatter plots of self-reported phonation time versus Ambulatory Phonation Monitor (APM)–based phonation time, in minutes per 1-hr period. Dashed lines indicate perfect accuracy. Numbered points refer to subject identifiers in Table 1 for subjects in the (A) control group, (B) low-voice-user group, and (C) patient group.



followed instructions to complete the voice-use diary each hour of the day, the laboratory personnel comprising the low-voice-use group might have been more diligent in this task due to their scientific training. And last, as a group, the patients were no worse than the subjects in the control group in terms of the accuracy of their voice-use estimates, which is counter to the expectation that patients might perform more poorly at this task.

Because the method of reporting voice use is different in this study (voice use rated every hour) from what clinicians typically ask their patients (voice use rated over days), the result that individuals overestimated their hourly speaking time might not hold for determining their ability to estimate their speaking time throughout an entire day—although it is expected that estimates that are based on longer periods of time would be even less accurate. Even though study subjects were encouraged to be accurate in their voice-use-diary entries, there was still substantial variability within and among subjects regarding how diligently they completed their voice-use diaries. This calls into question the extent to which inaccuracies in self-reporting were due to reporting inconsistencies or to poor awareness

of their voice use. A distinct advantage of newer phonation monitors with a graphical user interface, such as the devices in the studies by Popolo et al. (2005) and Mehta et al. (2012), is the capability to query the user about their voice use or environment on a preset schedule rather than relying on subject compliance for periodic paper-and-pencil reporting.

Typical overestimation of voice use that was found in the current study has a potentially broader impact on clinical voice assessment. For example, adding the average absolute error per hour for estimating voice use (5.7 min) to the measured (APM monitoring) average phonation time per hour (8.4 min or 14% phonation time) yields a self-report of 14.1 min/hr, which represents 23.5% phonation time. In practical terms, converting these percent phonation values into talk time (multiplying by 2), individuals who are actually talking about a quarter of the time (28%) could report that they are talking closer to half of the time (47%). Such discrepancies could conceivably be of sufficient magnitude to influence clinical decisions about the role of voice use in the etiology and/or treatment of common behavior-related voice disorders. However, even though the overall results of this study indicate that most

individuals tend to significantly overestimate their voice use to varying degrees regardless of vocal status, this was not a universal finding, and the relatively small sample size (18 subjects) precludes recommending that clinicians simply assume that all patients are overestimating their voice use.

There are obviously other factors associated with voice use in addition to phonation duration that can potentially affect vocal health, including vocal intensity (loudness) and fundamental frequency (pitch). The results of the present study also cast doubt on the accuracy with which patients might be able to self-assess or self-monitor these parameters during activities of daily living, which is often required as part of the voice assessment and therapy process. Ambulatory voice monitors that have the capability to objectively measure multiple vocal parameters (e.g., phonation duration, intensity, and fundamental frequency) and provide real-time feedback have the potential to significantly improve the clinical management of behavior-based voice disorders (Llico et al., 2015; Van Stan, Mehta, & Hillman, 2015; Van Stan, Mehta, Zeitels, et al., 2015). This would include defining or establishing safe limits for healthy voice use that are based on the types of quantitative or cumulative measures of daily phonatory behavior that ambulatory monitoring can provide (Mehta et al., 2015; Titze & Hunter, 2015).

Conclusions

The results of this study confirmed that people are generally inaccurate at estimating their phonation (talk) time, exhibiting a consistent bias toward overestimation of hourly voice use in a manner that correlates to widely varying degrees with objective measures of phonation time. This supports the potential that ambulatory voice monitoring may have in improving clinical voice assessment.

References

- Airo, E., Olkinuora, P., & Sala, E. (2000). A method to measure speaking time and speech sound pressure level. *Folia Phoniatica et Logopaedica*, 52, 275–288.
- Buekers, R., Bierens, E., Kingma, H., & Marres, E. H. M. A. (1995). Vocal load as measured by the voice accumulator. *Folia Phoniatica et Logopaedica*, 47, 252–261.
- Cheyne, H. A., Hanson, H. M., Genreux, R. P., Stevens, K. N., & Hillman, R. E. (2003). Development and testing of a Portable Vocal Accumulator. *Journal of Speech, Language, and Hearing Research*, 46, 1457–1467.
- Fant, G., Kruckenberg, A., & Nord, L. (1991). Prosodic and segmental speaker variations. *Speech Communication*, 10, 521–531.
- Hillman, R. E., Heaton, J. T., Masaki, A., Zeitels, S. M., & Cheyne, H. A. (2006). Ambulatory monitoring of disordered voices. *Annals of Otology, Rhinology & Laryngology*, 115, 795–801.
- Hillman, R. E., & Mehta, D. D. (2011). Ambulatory monitoring of daily voice use. *Perspectives on Voice and Voice Disorders*, 21, 56–61.
- Hogikyan, N. D., & Sethuraman, G. (1999). Validation of an instrument to measure voice-related quality of life (V-RQOL). *Journal of Voice*, 13, 557–569.
- Jacobson, B. H., Johnson, A., Grywalski, C., Silbergleit, A., Jacobson, G., Benninger, M. S., & Newman, C. W. (1997). The Voice Handicap Index (VHI): Development and validation. *American Journal of Speech-Language Pathology*, 6(3), 66–70.
- Kempster, G. B., Gerratt, B. R., Verdolini Abbott, K., Barkmeier-Kraemer, J., & Hillman, R. E. (2009). Consensus Auditory-Perceptual Evaluation of Voice: Development of a standardized clinical protocol. *American Journal of Speech-Language Pathology*, 18, 124–132.
- Llico, A. F., Zañartu, M., González, A. J., Wodicka, G. R., Mehta, D. D., Van Stan, J. H., & Hillman, R. E. (2015). Real-time estimation of aerodynamic features for ambulatory voice biofeedback. *The Journal of the Acoustical Society of America*, 138, EL14–EL19.
- Löfqvist, A., & Mandersson, B. (1987). Long-time average spectrum of speech and voice analysis. *Folia Phoniatica*, 39, 221–229.
- Masuda, T., Ikeda, Y., Manako, H., & Komiyama, S. (1993). Analysis of vocal abuse: Fluctuations in phonation time and intensity in 4 groups of speakers. *Acta Oto-Laryngologica*, 113, 547–552.
- Mehta, D. D., Van Stan, J. H., Zañartu, M., Ghassemi, M., Gutttag, J. V., Espinoza, V. M., . . . Hillman, R. E. (2015). Using ambulatory voice monitoring to investigate common voice disorders: Research update. *Frontiers in Bioengineering and Biotechnology*, 3, 155.
- Mehta, D. D., Zañartu, M., Feng, S. W., Cheyne, H. A., II, & Hillman, R. E. (2012). Mobile voice health monitoring using a wearable accelerometer sensor and a smartphone platform. *IEEE Transactions on Biomedical Engineering*, 59, 3090–3096.
- Misono, S., Banks, K., Gaillard, P., Goding, G. S., Jr., & Yueh, B. (2015). The clinical utility of vocal dosimetry for assessing voice rest. *The Laryngoscope*, 125, 171–176.
- Ohlsson, A.-C., Brink, O., & Löfqvist, A. (1989). A voice accumulation—Validation and application. *Journal of Speech and Hearing Research*, 32, 451–457.
- Popolo, P. S., Švec, J. G., & Titze, I. R. (2005). Adaptation of a Pocket PC for use as a wearable voice dosimeter. *Journal of Speech, Language, and Hearing Research*, 48, 780–791.
- Rantala, L., & Vilkmann, E. (1999). Relationship between subjective voice complaints and acoustic parameters in female teachers' voices. *Journal of Voice*, 13, 484–495.
- Roy, N., Merrill, R. M., Thibeault, S., Parsa, R. A., Gray, S. D., & Smith, E. M. (2004). Prevalence of voice disorders in teachers and the general population. *Journal of Speech, Language, and Hearing Research*, 47, 281–293.
- Ryu, S., Komiyama, S., Kannae, S., & Watanabe, H. (1983). A newly devised speech accumulator. *ORL*, 45, 108–114.
- Sala, E., Airo, E., Olkinuora, P., Simberg, S., Ström, U., Laine, A., . . . Suonpää, J. (2002). Vocal loading among day care center teachers. *Logopedics Phoniatics Vocology*, 27, 21–28.
- Södersten, M., Granqvist, S., Hammarberg, B., & Szabo, A. (2002). Vocal behavior and vocal loading factors for preschool teachers at work studied with binaural DAT recordings. *Journal of Voice*, 16, 356–371.
- Szabo, A., Hammarberg, B., Håkansson, A., & Södersten, M. (2001). A voice accumulator device: Evaluation based on studio and field recordings. *Logopedics Phoniatics Vocology*, 26, 102–117.
- Titze, I. R., & Hunter, E. J. (2015). Comparison of vocal vibration-dose measures for potential-damage risk criteria. *Journal of Speech, Language, and Hearing Research*, 58, 1425–1439.

-
- Titze, I. R., Hunter, E. J., & Švec, J. G.** (2007). Voicing and silence periods in daily and weekly vocalizations of teachers. *The Journal of the Acoustical Society of America*, *121*, 469–478.
- Titze, I. R., Švec, J. G., & Popolo, P. S.** (2003). Vocal dose measures: Quantifying accumulated vibration exposure in vocal fold tissues. *Journal of Speech, Language, and Hearing Research*, *46*, 919–932.
- Van Stan, J. H., Gustafsson, J., Schalling, E., & Hillman, R. E.** (2014). Direct comparison of three commercially available devices for voice ambulatory monitoring and biofeedback. *SIG 3 Perspectives on Voice and Voice Disorders*, *24*, 80–86.
- Van Stan, J. H., Mehta, D. D., & Hillman, R. E.** (2015). The effect of voice ambulatory biofeedback on the daily performance and retention of a modified vocal motor behavior in participants with normal voices. *Journal of Speech, Language, and Hearing Research*, *58*, 713–721.
- Van Stan, J. H., Mehta, D. D., Zeitels, S. M., Burns, J. A., Barbu, A. M., & Hillman, R. E.** (2015). Average ambulatory measures of sound pressure level, fundamental frequency, and vocal dose do not differ between adult females with phonotraumatic lesions and matched control subjects. *Annals of Otolaryngology, Rhinology & Laryngology*, *124*, 864–874.
- Vilkman, E.** (2004). Occupational safety and health aspects of voice and speech professions. *Folia Phoniatrica et Logopaedica*, *56*, 220–253.
- Watanabe, H., Komiyama, S., Ryu, S., & Kannae, S.** (1984). A newly designed automatic phonometer. *Agressologie*, *25*, 1035–1037.
- Watanabe, H., Shin, T., Oda, M., Fukaura, J., & Komiyama, S.** (1987). Measurement of total actual speaking time in a patient with spastic dysphonia. *Folia Phoniatrica*, *39*, 65–70.
- Zañartu, M., Ho, J. C., Kraman, S. S., Pasterkamp, H., Huber, J. E., & Wodicka, G. R.** (2009). Air-borne and tissue-borne sensitivities of bioacoustic sensors used on the skin surface. *IEEE Transactions on Biomedical Engineering*, *56*, 443–451.