I International Conference on Occupational Voice
September 16, 17, 18 2016
Bogotá, Colombia

Vocal Dose Measures: General Rationale, Traditional Methods and Recent Advances

Eric J. Hunter
Communicative Sciences and Disorders
Michigan State University
1026 Red Cedar Road, East Lansing, Michigan 48824

Abstract
In an attempt to track the increased vocal health risks for occupational voice users, devices have been developed for ambulatory monitoring of voice use. These devices are often used either to correlate excessive voice use with voice problems or to identify voice overuse. The purpose of this presentation is to discuss benefits and pitfalls of vocal dose measurements in the context of the historical development of the methods. Discussed will be a short review of landmark papers, the effects of monitoring on teachers, and some difficulties working in schools. Currently, there are several commercial options available which can be used to address a range of research questions and clinical needs. Device uncertainty in recording will be discussed. Finally, while these devices are proven and are readily available for both clinical and research use, the cost of the technology can be a hindrance. Therefore, the presentation will end with a discussion of a research-grade voice monitoring system, made from off-the-shelf components from several commercial vendors for less than $500USD, which could be used to record several hours to several days of data.

1. Introduction

People working in vocally demanding professions voice problems more frequently than people in vocally non-demanding professions[1]. These occupational voice users are those who not only depend on a healthy, versatile voice as a tool for their profession but would be unable to perform their primary job responsibility in the event of the loss of vocal endurance and/or voice quality. These individuals include (but are not limited to) teachers, actors, singers, broadcasters, air traffic controllers, emergency dispatchers, and individuals involved in telephone customer service or marketing. In the United States (2000), these include approximately 37 million individuals, or nearly one quarter of the workforce. Individuals in such professions are affected by physiology[2], room acoustics[3], the extent of vocal loading[4], and recovery period[5]. Occupationally related laryngeal injuries can lead to missed work days or performances, lost revenue, significant rehabilitation periods, and the need to change of profession or retire early.

Because of a dependence on a healthy, versatile voice, an individuals in these professions is affected by the extent of vocal loading, a term used to quantify the demands placed on the voice mechanism by the way a voice is used and how much it is used[6]. Because it is valuable to quantify how much a voice is used, numerous full-day ambulatory voice monitors have been developed in the last few decades to facilitate multiple day measurement of voice. Additionally, recent scientific efforts to establish occupational safety criteria for vocalization has made such developments that much more important.

In this paper, a short review of several landmark papers in the development of ambulatory voice monitors, will be presented. Additionally, a brief review of some commercial options for monitor will be touched on, as well as the uncertainty and problems with the devices. Finally, new directions in analysis and costs will be discussed.

2. Historical Overview

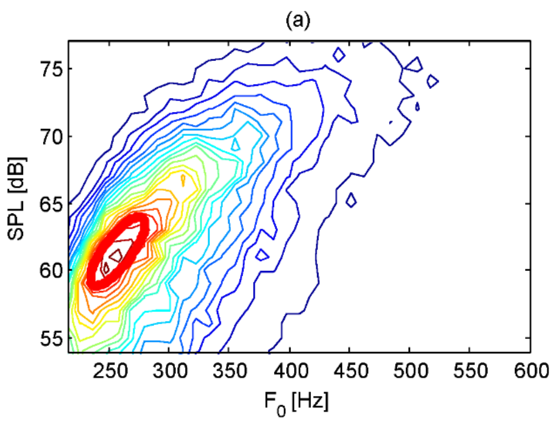
2.1. Brief History of Development

Voice accumulation and voice dosimetry devices have been developed for ambulatory monitoring of voice use, usually to either relate excessive voice use to voice problems or to identify voice overuse. One of the earliest devices (called a speaking timer) used a contact microphone which detected phonation time[7]. An additional early prototype, which logged phonation time and estimated vocal fundamental frequency, used a contact microphone[8] placed on the neck. Another early device also used a contact microphone but estimated phonation time and SPL[9]. Using a completely different approach to monitoring voice over the course of a day was a device that used a pair of noise exposure analyzers from which the phonation time was extracted[10]. By 2001, voice accumulators were collecting phonation time, estimated fundamental frequency, and approximating SPL[11]. From these studies, phonation time, long-term SPL, and F0 data on occupational voice users were added to the scientific knowledge base.

By the mid-2000, there were two distinct lines of work: one used digital recorders and multiple microphones[12,13], while the other was based on an accelerometer affixed to the neck[14,15]. These new devices (one being commercially available) allowed researchers to collect day-long and multiple-day phonation data.

2.2. Insights Gained

Such research allowed for new understanding of the effect of phonation on occupational voice users. For example, using hand vibration literature as a model, phonation exposure doses were suggested using empirical models and measured phonation time, phonation fundamental frequency and estimated SPL[16,17]. Other studies examined the measured effect of using voice amplification by school teachers[18] or tracked a performer's recovery after a performance[19]. More than just presenting phonation time, fundamental frequency and SPL, variations were used to look in more depth at how the voice was being used based on the three parameters. For example, these devices were used to show how a child would change his vocal style depending on the environment using Voice Use Profiles[20], all based on the three parameters (Fig. 1). In 2016, the accumulation durations of phonation voicing and silences were shown to be different when comparing teachers with and without voice disorders[21].



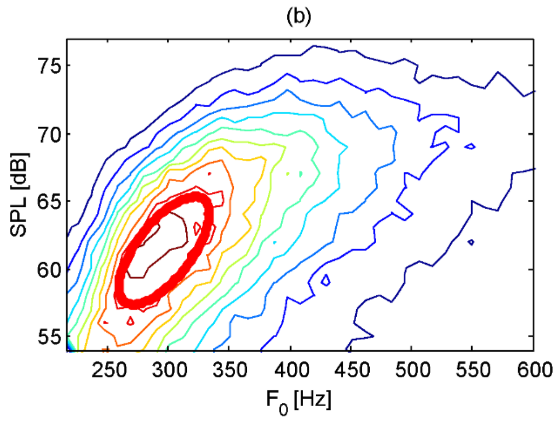


Fig 1. Two Voice Use Profile plots (contours) showing where the child spent time phonating (a) with adults, and (b) at preschool. The ellipse shows the most frequent 10% of voicing for the environment (after Hunter et al, 2012).

2.3. Best Practice & Revisions

As these devices became more widespread, studies began to examine the use of the devices itself to develop best practices (e.g., measurement and calibration[16,22], working with teachers[23], effect of using a monitor on the participating individual[24]). In a conference paper, it was presented that it would take about 12 hours of monitoring to reduce the percentage error on average F0 estimation to about 1%[25] (Fig. 2). Additionally, new analysis techniques have increased reliability and better quantify vocal vibration exposure. Lindstrom et al suggested a particular low pass filter which would remove many false positives of voicing[26]. Later, a revision of the empirical dose models was reported using neck accelerometer signals and endoscopic images[27], which resulted in estimated phonation damage risk criteria[28].

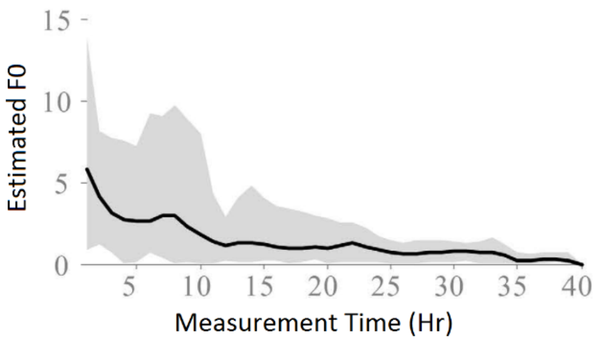


Fig. 2. Percentage error (multiple subjects) where dark line represents mean error and the gray represents range of error (min, max). Average F0 error decreased to about 1% after 12 hours. (After Mehta et al, 2012).

3. Current Uses and Advances

There are several options available currently, which can be used to address a range of research questions and clinical needs. These devices use either an accelerometer or contact microphone in contact with the neck, reducing interference of external noises. A portable electronic device processes the signal and stores skin vibration amplitude and fundamental frequency, which alleviates privacy concerns. Post-collection analyses are processed on a computer after capture is complete.

As of the time of this writing, three commercial products exist for purchase and/or lease: 1) VocaLog2 (Griffin Labs); 2) VoxLog (Sonvox AB); and 3) Voice-Care (PR.O.VOICE). A fourth previously available device, KayPENTAX's APM, has been discontinued.

Fig. 3. Four commercially produced devices.

Outside the commercial devices, a few have been created for research purposes. The National Center for Voice and Speech Dosimeter[15] has been described in dozens of studies but is not available for outside use. Another option currently being developed by Hillman et al uses a smartphone[29] as the processor with an accelerometer (Fig 4).

3.1. Current limitations

There are several limitations in the use of these devices. First, the cost of the traditional options can be prohibitive for small clinics, individual researchers, or studies in which multiple devices may be needed (with most devices costing several thousand USD, and the VocaLog2 costing the least at under a thousand USD). Second, the current commercial options only collecting phonation time, SPL, and F0. Finally, there are no standards for ambulatory voice monitoring (e.g. validity, accuracy, reliability,), and there has been little published on well some of these devices capture voice.

3.1.1. Prohibited Costs

Except for the VocaLog2, the cost of available devices prevents many research labs or clinics to have more than just one or two of these devices. Nevertheless, there are several possibilities on the horizon. First, the development of smartphone options like the one discussed above should be more cost-effective. For example, the makers of the Voice-Care recently announced a smartphone option (Fig. 4b), in which the user would only pay for the transducer and the phone application.

Fig. 4. Smartphone solutions. (a) Research device being developed by Hillman et al[29]. (b) The Voice-Care-App is currently being developed by the makers of the Voice-Care.

Another option that is currently available uses separate commercially available devices in tandem: the VoxLog's accelerometer/microphone collar with a digital recorder (e.g. TASCAM or Roland R-05 digital recorders, Fig 5). Because the accelerometer uses a standard 1/8-inch stereo jack, it can be used with many devices. However, this number is reduced by the accelerometer's need for a power supply. Nevertheless, this device has been successfully used in several studies[30] at a current cost of less than 500 USD.

Fig. 5 VoxLog collar and digital recorders. (left) TASCAM digital recorder ~100USD. (right) Roland R-05 digital recorder ~200USD.

The primary benefit of these types of systems is cost, both the initial investment and replacement if broken (about one tenth of the price of the commercial offerings). This cost savings also opens up the opportunity to run several devices at one time while having several backup components for a fraction of the initial investment.

As with any budget system, there are several limitations. First, there is no support for the device but the technical skills in research team and support obtained from the scientific community. Another limitation is potential privacy issues. The full vibration signal is recorded from the neck collar and is partially intelligible on the recording, adding complexity to recruiting and Institutional Review Board approval if used in a healthcare environment.

In addressing a clinical need or research question, if clinic/research personnel time is costly or if only one or two devices are needed, the commercial devices are adequate. If multiple devices are needed, cost is an issue, or if unique analysis is desired (with available technical expertise), a budget system similar to what is shown here may be more effective.

3.1.1. Limited Parameters

Traditionally ambulatory monitors of voice have logged phonation time, SPL and F0. While useful in recording how long, how loud and at what pitch someone is speaking, it does not provide information about vocal effort or any other spectral characteristic. There are some options. First, of the commercial devices, the VoxLog has an option to monitor the SPL of the environment. Second, the Voice-Care has an option to record more audio, so calculating more parameters (eg. LTAS) is possible. The NCVS Dosimeter[15] also logged spectral centroid (though never reported), while the Hillman et al's smartphone estimates the subglottal signal from the accelerometer to get an pseudo-inversed filtered subglottal impedence[31] (Fig 4).

The VoxLog collar used in tandem with a recording device provides the option of a full accelerometer signal for analysis, rather than the estimated F0 and level that is traditionally kept. Therefore, other analyses could be conducted (e.g., Cepstral Peak Prominence, Alpha Ratio, Long Term Average Spectrum) and scripts could be written into PRAAT to process dose.

3.1.1. Device Uncertainty

To begin to explore device uncertainty, a recent study of the four commercial devices and their uncertainty in estimating the SPL was conducted[32]. From the results, the device with the highest mean error was the APM, followed by the VoxLog, the VocaLog and the Voice-Care. For fundamental frequency specifically, the VoxLog had the highest mean error, followed by the Voice-Care, and the APM (the VocaLog doesn't monitor F0). While a first step, further study of the reliability and uncertainty is warranted.

4. Conclusions

People working in vocally demanding professions are at an elevated risk for voice issues. Ambulatory monitoring of voice use has provided tremendous insight underlying this risk. Currently, there are several devices a clinician or voice researcher can choose from, depending on use and costs. Within the next few years, the features promise to go up, with the cost maintained or decreasing. The options for researchers and clinicians to use these devices should increase rapidly.

Acknowledgements

Thank you to Laura M. Hunter for the editorial review of the document. This publication was prepared with support in part by the National Institute of Deafness and Other Communication Disorders of the National Institutes of Health under Award Number R01DC012315. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

5. Reference

1. Titze IR, Lemke J, Montequin D. Populations in the U.S. workforce who rely on voice as a primary tool of trade: a preliminary report. *J Voice*. 1997;11(3):254-259. doi:10.1016/S0892-1997(97)80002-1.
2. Hunter, Eric J. L Timothy W. Talking, teaching, and listening: Gender differences in teachers' responses to acoustic environments. In: *Science and Technology for a Quiet Europe*. ; 2016:721-726.
3. Bottalico P, Graetzer S, Hunter EJ. Effects of speech style, room acoustics, and vocal fatigue on vocal effort. *J Acoust Soc Am*. 2016;139(5):2870–2879.
4. 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(4):862–875.
5. Hunter EJ, Titze IR. Quantifying vocal fatigue recovery: Dynamic vocal recovery trajectories after a vocal loading exercise. *Ann Otol Rhinol Laryngol*. 2009;118(6):449-460.
6. Vilkman E. Occupational Safety and Health Aspects of Voice and Speech Professions. *Folia Phoniatr Logop*. 2004;56(4):220-253. doi:10.1159/000078344.

7. Watanabe H, Shin T, Oda M, Fukaura J, Komiyama S. Measurement of total actual speaking time in a patient with spastic dysphonia. *Folia Phoniatr Logop*. 1987;39(2):65–70.
8. Ohlsson A-C, Brink O, Lofqvist A. A voice accumulator—validation and application. *J Speech Lang Hear Res*. 1989;32(2):451–457.
9. Buekers R, Bierens E, Kingma H, Marres EHMA. Vocal Load as Measured by the Voice Accumulator. *Folia Phoniatr Logop*. 1995;47(5):252-261. doi:10.1159/000266359.
10. Airo E, Olkinuora P, Sala E. A Method to Measure Speaking Time and Speech Sound Pressure Level. *Folia Phoniatr Logop*. 2000;52(6):275-288. doi:10.1159/000021545.
11. Szabo A, Hammarberg B, Hakansson A, Sodersten M. A voice accumulator device: evaluation based on studio and field recordings. *Logoped Phoniatr Vocol*. 2001;26(3):102-117. doi:10.1080/14015430152728016.
12. Södersten M, Granqvist S, Hammarberg B, Szabo A. Vocal behavior and vocal loading factors for preschool teachers at work studied with binaural DAT recordings. *J Voice*. 2002;16(3):356–371.
13. Granqvist S. The self-to-other ratio applied as a phonation detector for voice accumulation. *Logoped Phoniatr Vocol*. 2003;28(2):71-80. doi:10.1080/14015430310011772.
14. Cheyne HA. Development and Testing of a Portable Vocal Accumulator. *J Speech Lang Hear Res*. 2003;46(6):1457-1467. doi:10.1044/1092-4388(2003/113).
15. Popolo PS, Svec JG, Titze IR. Adaptation of a Pocket PC for use as a wearable voice dosimeter. *J Speech Lang Hear Res JSLHR*. 2005;48(4):780-791. doi:10.1044/1092-4388(2005/054).
16. G. Švec J, S. Popolo P, R. Titze I. Measurement of vocal doses in speech: experimental procedure and signal processing. *Logoped Phoniatr Vocol*. 2003;28(4):181-192. doi:10.1080/14015430310018892.
17. Titze IR. Vocal Dose Measures: Quantifying Accumulated Vibration Exposure in Vocal Fold Tissues. *J Speech Lang Hear Res*. 2003;46(4):919-932. doi:10.1044/1092-4388(2003/072).
18. Gaskill CS, O'Brien SG, Tinter SR. The Effect of Voice Amplification on Occupational Vocal Dose in Elementary School Teachers. *J Voice*. 2012;26(5):667.e19-667.e27.
19. Carroll T, Nix J, Hunter E, Emerich K, Titze I, Abaza M. Objective measurement of vocal fatigue in classical singers: a vocal dosimetry pilot study. *Otolaryngol--Head Neck Surg*. 2006;135(4):595-602. doi:10.1016/j.otohns.2006.06.1268.
20. Hunter EJ, Halpern AE, Spielman JL. Impact of Four Nonclinical Speaking Environments on a Child's Fundamental Frequency and Voice Level: A Preliminary Case Study. *Lang Speech Hear Serv Sch*. 2012;43(3):253-263. doi:10.1044/0161-1461(2011/11-0002).
21. Bottalico P, Graetzer S, Astolfi A, Hunter EJ. Silence and Voicing Accumulations in Italian Primary School Teachers With and Without Voice Disorders. *J Voice*. In Press. doi:10.1016/j.jvoice.2016.05.009.
22. 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(3 Pt 1):1386-1394.
23. Nix J, Švec JG, Laukkanen A-M, Titze IR. Protocol Challenges for On-the-Job Voice Dosimetry of Teachers in the United States and Finland. *J Voice*. 2007;21(4):385-396. doi:10.1016/j.jvoice.2006.03.005.
24. Hunter EJ. Teacher response to ambulatory monitoring of voice. *Logoped Phoniatr Vocol*. 2012;37(3):133-135. doi:10.3109/14015439.2012.664657.
25. Mehta DD, Listfield RW, Cheyne H., et al. Duration of ambulatory monitoring needed to accurately estimate voice use. In: *INTERSPEECH*; 2012: 1335-1338. f.
26. Lindstrom F, Ren K, Li H, Waye KP. Comparison of two methods of voice activity detection in field studies. *J Speech Lang Hear Res*. 2009;52(6):1658+.
27. Popolo PS, Titze IR. Qualification of a quantitative laryngeal imaging system using videostroboscopy and videokymography. *Ann Otol Rhinol Laryngol*. 2008;117(6):404.
28. Titze IR, Hunter EJ. Comparison of Vocal Vibration Dose Measures for Potential Damage Risk Criteria. *J Speech Lang Hear Res*. July 2015. doi:10.1044/2015_JSLHR-S-13-0128.
29. Mehta DD, Zanartu M, Van Stan JH, Feng SW, Cheyne HA, Hillman RE. Smartphone-based detection of voice disorders by long-term monitoring of neck acceleration features. In: *2013 IEEE International Conference on Body Sensor Networks (BSN)*. ; 2013:1-6.
30. Schloneger MJ, Hunter EJ. Assessments of Voice Use and Voice Quality Among College/University Singing Students Ages 18–24 Through Ambulatory Monitoring With a Full Accelerometer Signal. *J Voice*. in press. doi:10.1016/j.jvoice.2015.12.018.
31. Zañartu M, Ho JC, Mehta DD, Hillman RE, Wodicka GR. Subglottal Impedance-Based Inverse Filtering of Voiced Sounds Using Neck Surface Acceleration. *IEEE Trans Audio Speech Lang Process*. 2013;21(9):1929-1939. doi:10.1109/TASL.2013.2263138.
32. Bottalico P, Passione II, Hunter EJ. Vocal dosimeter devices and their uncertainty. *J Acoust Soc Am*. 2016;139(4):2018-2018. doi:10.1121/1.4949949.