JSLHR

Research Note

Voice Relative Fundamental Frequency Via Neck-Skin Acceleration in Individuals With Voice Disorders

Yu-An S. Lien,[a] Carolyn R. Calabrese,[a] Carolyn M. Michener,[a] Elizabeth Heller Murray,[a] Jarrad H. Van Stan,[b,c] Daryush D. Mehta,[b,c,d] Robert E. Hillman,[b,c,d] J. Pieter Noordzij,[a,e] and Cara E. Stepp[a,e]

Purpose: This study investigated the use of neck-skin acceleration for relative fundamental frequency (RFF) analysis.
Method: Forty individuals with voice disorders associated with vocal hyperfunction and 20 age- and sex-matched control participants were recorded with a subglottal neck-surface accelerometer and a microphone while producing speech stimuli appropriate for RFF. Rater reliabilities, RFF means, and RFF standard deviations derived from the accelerometer were compared with those derived from the microphone.
Results: RFF estimated from the accelerometer had slightly higher intrarater reliability and identical interrater reliability compared with values estimated with the microphone. Although sensor type and the Vocal Cycle × Sensor and Vocal Cycle × Sensor × Group interactions showed significant effects on RFF means, the typical RFF pattern could be derived from either sensor. For both sensors, the RFF of individuals with vocal hyperfunction was lower than that of the controls. Sensor type and its interactions did not have significant effects on RFF standard deviations.
Conclusions: RFF can be reliably estimated using an accelerometer, but these values cannot be compared with those collected via microphone. Future studies are needed to determine the physiological basis of RFF and examine the effect of sensors on RFF in practical voice assessment and monitoring settings.

Vocal hyperfunction is often described as a chronic or recurring condition that likely results from exerting poorly regulated or excessive laryngeal tension during phonation (Hillman, Holmberg, Perkell, Walsh, & Vaughan, 1989). Current evaluation of vocal hyperfunction relies primarily on a clinician's interpretation on the basis of subjective measures of auditory perception, neck palpation, endoscopic imaging, and patient self-report of perceived fatigue or discomfort (Andrews, 1996; Behrman, 2005; Morrison, Nichol, & Rammage, 1986; Roy, Ford, & Bless, 1996). Initial studies support relative fundamental frequency (RFF) as a potential objective indicator of vocal hyperfunction and a promising acoustic correlate of vocal effort. RFF is defined as the 10 fundamental frequencies in voiced segments immediately prior to the onset and following the offset of unvoiced consonants, normalized to steady-state fundamental frequencies (see Figure 1, upper). In young speakers with healthy voices, the RFF in the offset vowel (offset RFF) remains unchanged or slightly decreases as it gets closer to the voiceless consonant; the RFF in the onset vowel (onset RFF) is highest closest to the voiceless consonant and decreases thereafter (Robb & Smith, 2002; Watson, 1998). However, for speakers with vocal hyperfunction, both offset and onset RFF tend to be lower compared with those of age-matched controls (Stepp, Hillman, & Heaton, 2010). Furthermore, RFF values significantly increase toward normative values following successful voice therapy (Stepp, Merchant, Heaton, & Hillman, 2011) and have also been shown to correlate with the perception of excessive vocal effort (Eadie & Stepp, 2013; Stepp, Sawin, & Eadie, 2012), a percept that is often used to describe hyperfunctional voices (Andrews, 1996; Hillman et al., 1989).

The majority of studies on RFF have estimated the measure from acoustic waveforms recorded using a microphone (Eadie & Stepp, 2013; Goberman & Blomgren, 2008;

[a]Boston University, Boston, MA
[b]MGH Institute of Health Professions, Boston, MA
[c]Center for Laryngeal Surgery & Voice Rehabilitation, Massachusetts General Hospital, Boston
[d]Harvard Medical School, Cambridge, MA
[e]Boston University School of Medicine, Boston, MA

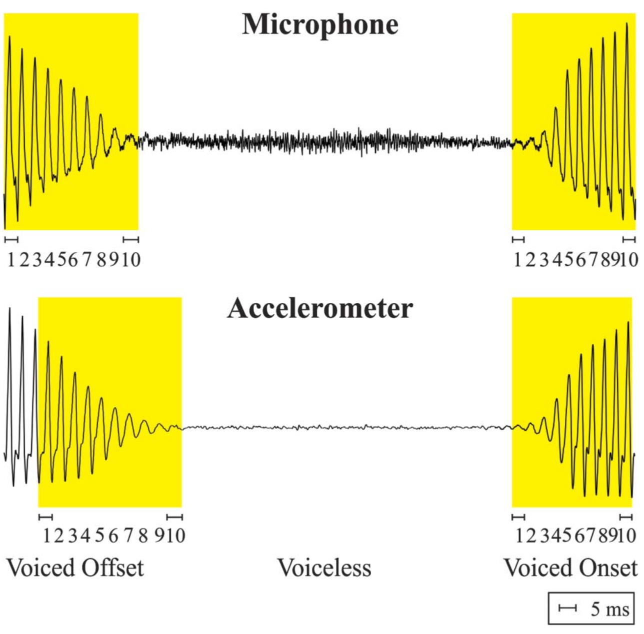
Correspondence to Cara E. Stepp: cstepp@bu.edu

Editor: Jody Kreiman
Associate Editor: Bruce Gerratt

Received April 4, 2015
Accepted June 25, 2015
DOI: 10.1044/2015_JSLHR-S-15-0126

Disclosure: Robert E. Hillman has a financial relationship with Pentax Medical with respect to the Ambulatory Phonation Monitor.

Figure 1. Upper panel: A waveform of the relative fundamental frequency (RFF) instance /ufu/ recorded using a microphone. Lower panel: A waveform of the same RFF instance /ufu/ recorded using a neck-placed accelerometer. The highlighted portion indicates the 10 offset and 10 onset vocal cycles used for RFF estimation in the recordings. The bar scales directly below the waveform denote the first and 10th vocal cycles for both offset and onset vowels. The time calibration bar below the lower waveform denotes the time scale of a 5-ms interval.



Robb & Smith, 2002; Stepp, 2013; Stepp et al., 2010, 2011, 2012; Watson, 1998). There is a long-standing interest in using neck-placed miniature accelerometers to assess vocal function, particularly for long-term ambulatory monitoring of voice (Cheyne, Hanson, Genereux, Stevens, & Hillman, 2003; Hillman, Heaton, Masaki, Zeitels, & Cheyne, 2006; Mehta, Zanartu, Feng, Cheyne, & Hillman, 2012; Ohlsson, Brink, & Lofqvist, 1989; Popolo, Scaronvec, & Titze, 2005; Ryu, Komiyama, Kannae, & Watanabe, 1983). For such applications, accelerometers offer advantages over microphones by being less sensitive to environmental noise and by preserving confidentiality because accelerometers (placed below the larynx on the neck) do not enable recording of the intelligible speech waveform (Cheyne et al., 2003; Ryu et al., 1983). Commercial devices that incorporate accelerometers for ambulatory voice monitoring, including the Ambulatory Phonation Monitor (Model 3200, KayPENTAX, Montvale, NJ) and VoxLog (Sonvox AB, Umeå, Sweden), can generally be used to estimate sound pressure level, fundamental frequency, and phonation time across extended periods of time as individuals engage in their activities of daily living (Van Stan, Gustafsson, Schalling, & Hillman, 2014). The clinical use of ambulatory monitoring may be greatly enhanced if RFF could also be extracted from the accelerometer signal, given its potential as a biomarker for vocal hyperfunction. We have recently shown that neck acceleration can provide reliable estimates of RFF in individuals with healthy voices (Lien & Stepp, 2014), but it is unclear whether this result extends to individuals with disordered voices related to vocal hyperfunction. Thus, the goal of this study was to compare the RFF estimates derived from a neck-placed accelerometer to those derived from a microphone by examining and comparing the inter- and intrarater reliabilities, mean RFF, and RFF standard deviations in individuals with and without voice disorders related to vocal hyperfunction.

Method

Participants

A control group comprised 20 adults (18 women, two men) aged 18–87 years (M = 33, SD = 17), all of whom reported no prior history of speech, language, or hearing disorders. A group with voice disorders consisted of 40 adults (35 women, five men) aged 18–75 years (M = 39, SD = 17). Individuals in the latter group were all diagnosed with vocal hyperfunction by a board-certified laryngologist. Of these individuals, 21 were diagnosed with vocal fold lesions (polyps, nodules, or cysts) and vocal hyperfunction, whereas the remaining were diagnosed with muscle tension dysphonia. The Consensus Auditory-Perceptual Evaluation of Voice ratings, performed by a licensed speech-language pathologist, indicated that this group consisted of individuals with a wide range of overall severity of dysphonia (M = 34, SD = 21, range = 1–86). All participants completed written consent in compliance with either the Boston University Institutional Review Board or the Massachusetts General Hospital Institutional Review Board.

Experimental Design

Two sets of equipment were used for the recordings. The first set comprised a Sennheiser PC131 headset microphone (Sennheiser, Wedemark, Germany) with a Knowles BU-21771 miniature accelerometer (Knowles Acoustics, Itasca, IL). The second set of equipment consisted of a Sennheiser MKE104 lavalier microphone (Sennheiser) with a Knowles BU-27135 miniature accelerometer (Knowles Acoustics). In both cases, the accelerometer was attached to the surface of the neck just above the sternal notch using 3M Model 2181 medical grade adhesive (3M, St. Paul, MN). In a quiet room, the microphone signals were recorded at a minimum of 20 kHz, the accelerometer signals were recorded at a minimum of 11025 Hz, and both signals used 16-bit resolution. The accelerometer signals were up-sampled to the sampling frequency of the microphone signal in postprocessing to facilitate signal alignment and comparison.

Each participant was instructed to read the same set of stimuli in his or her typical pitch and loudness. The stimuli consisted of four sentences and three uniform utterances specifically designed for RFF analysis. Each sentence was purposefully loaded with three voiced–voiceless–voiced instances in which the voiceless phoneme was either /f/ or /ʃ/ (e.g., "We feel you do fail in new fallen dew"). Each uniform utterance consisted of three repetitions of the same voiced–voiceless–voiced instance with the voiceless phoneme /f/ (e.g., "/ɑfɑ ɑfɑ ɑfɑ/"). These stimuli were selected because RFF estimated from instances containing /f/ and /ʃ/ have been shown to have lower intraspeaker variability than those estimated from other voiceless phonemes (Lien, Gattuccio, & Stepp, 2014). The experimenter modeled the uniform utterances for the participant prior to the recording. During the recording, if any stimulus was misarticulated or obviously glottalized, the experimenter instructed the participant to repeat it.

Estimation of RFF

RFF was estimated from two types of sensors—microphone and accelerometer—in a similar manner by three individuals trained in RFF analysis. The technique used is identical to the one described in Lien et al. (2014). First, each technician examined the voiced–voiceless–voiced instance in Praat (Version 5.3.04; Boersma & Weenink, 2012) to extract the 11 pulse timings immediately before and after the voiceless consonant. Samples were rejected if any phoneme was misarticulated, if the voiced section was glottalized, or if the voiced section did not contain at least 10 voicing cycles. Otherwise, the pulse timings were exported to Excel (Version 14) to calculate the instantaneous fundamental frequency (F0). The F0 were normalized to a reference fundamental frequency ($F0_{ref}$), transforming it into a semitone (ST) scale to calculate the RFF:

$$\text{RFF (with units in ST)} = 39.86 \times \log_{10}\left(F0 \,/\, F0_{ref}\right) \quad (1)$$

Similar to previous studies, the $F0_{ref}$ for offset cycles and onset cycles were the F0s for Offset Cycle 1 and Onset Cycle 10, respectively. The RFF values estimated from nominally three RFF instances in each utterance were used to calculate the utterance-level RFF means and standard deviations.

Reliability Procedures and Analysis

Interrater reliability of RFF estimates was analyzed using the intraclass correlation (Shrout & Fleiss, 1979), type (2,k). To determine the intrarater reliability, each technician re-estimated 15% of the total samples from each sensor (microphone and accelerometer) in a different sitting and Pearson product–moment correlation coefficients were calculated.

All other statistical analyses were performed using Minitab Statistical Software (Minitab Inc., State College, PA). A mixed-design analysis of variance was used to determine the effect of sensor (microphone or accelerometer) on both the utterance-level RFF means and standard deviations. The within-subject factors were sensor and vocal cycle (Offset 1–10 or Onset 1–10). The between-subjects factor was group (control or voice disorder). Main effects and all possible interactions were investigated. Effect sizes were quantified using the square partial curvilinear correlation (η_p^2), and values of .01, .09, and .25 were interpreted as small, medium, or large, respectively (Witte & Witte, 2010). A predetermined level of statistical significance ($p < .05$) was used for all analyses. All post hoc analyses were completed using Tukey's honestly significant difference tests.

Results

The intrarater Pearson product–moment correlation coefficients of the three technicians were .92, .95, and .97 for the accelerometer and slightly lower for the microphone, with values of .88, .91, and .95. The interrater reliabilities for the microphone and accelerometer were both .95.

A three-factor general linear model (see Table 1) indicated statistically significant main effects ($p < .05$) of vocal cycle (Offset 1–10 or Onset 1–10), sensor (microphone or accelerometer), group (control or voice disorder), and statistically significant Vocal Cycle × Sensor, Vocal Cycle × Group, and Vocal Cycle × Sensor × Group interactions on the utterance-level RFF means. There was no statistically significant Sensor × Group interaction on utterance-level RFF means and the effect sizes of sensor, and the Vocal Cycle × Sensor and Vocal Cycle × Sensor × Group interactions were all small ($\eta_p^2 < .01$). Similarly, the effect sizes of group and the Vocal Cycle × Group interaction were also small ($\eta_p^2 \leq .01$). Post hoc testing revealed that the utterance-level RFF means determined from the microphone were significantly ($p_{adj} < .05$) higher than those estimated using the accelerometer. To explore these differences in terms of the statistically significant Vocal Cycle × Sensor interaction, the utterance-level RFF means for each sensor were plotted as a function of vocal cycle (see Figure 2). Visual examination of the plot revealed that RFF means for the microphone were higher than those for the accelerometer for the cycles close to the voiceless consonant. The difference is most apparent in Offset Cycles 5–10 and Onset Cycle 1. To examine the differences in terms of the statistically significant Vocal Cycle × Sensor × Group interaction, the utterance-level RFF means for each group were plotted as a function of vocal cycle for each sensor (see Figure 3). Despite the differences as a function of sensor, in both sensors, utterance-level RFF means for Offset Cycles 7–10 and Onset

Table 1. Results of three-factor general linear model on utterance-level relative fundamental frequency means.

Effect	*df*	η_p^2	*F*	*p*
Vocal Cycle (Offset 1–10 or Onset 1–10)	19	.50	2242.3	<.001
Sensor (microphone or accelerometer)	1	<.01	177.9	<.001
Group (control or voice disorder)	1	<.01	4.2	.044
Vocal Cycle × Sensor	19	.01	25.4	<.001
Vocal Cycle × Group	19	.01	38.5	<.001
Sensor × Group	1	<.01	2.9	.089
Vocal Cycle × Sensor × Group	19	<.01	1.7	.026

Figure 2. Utterance-level relative fundamental frequency (RFF) means as a function of sensor (microphone or accelerometer) and vocal cycle (offset 1–10 or onset 1–10) in semitones (ST). Error bars indicate 95% confidence intervals for the means.

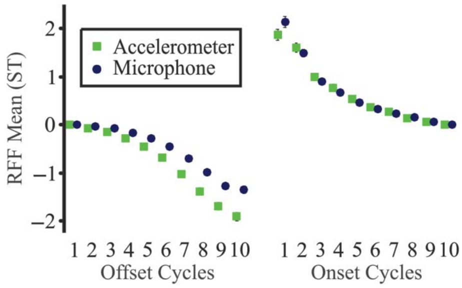


Table 2. Results of three-factor general linear model on utterance-level relative fundamental frequency standard deviations.

Effect	*df*	η_p^2	*F*	*p*
Vocal Cycle (Offset 1–10 or Onset 1–10)	19	.28	827.7	<.001
Sensor (microphone or accelerometer)	1	<.01	0.2	.655
Group (control or voice disorder)	1	.03	15.1	<.001
Vocal Cycle × Sensor	19	<.01	0.7	.799
Vocal Cycle × Group	19	<.01	20.1	<.001
Sensor × Group	1	<.01	2.8	.092
Vocal Cycle × Sensor × Group	19	<.01	1.0	.436

Cycles 1–2 were lower in the voice disorder group compared with the control group.

A three-factor general linear model (see Table 2) indicated statistically significant effects ($p < .001$) of vocal cycle, group, and the Vocal Cycle × Group interaction on utterance-level RFF standard deviations. The effect size of the Vocal Cycle × Group interaction was small ($\eta_p^2 < .01$), and the effect size of group was small to medium ($\eta_p^2 = .03$). There was no statistically significant effect of sensor or the Vocal Cycle × Sensor, Sensor × Group, and Vocal Cycle × Sensor × Group interactions.

Discussion

Reliability

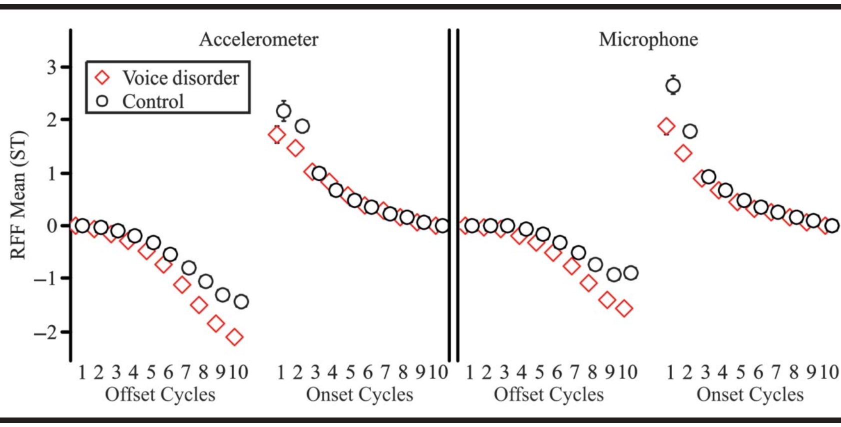
We found intrarater correlations of .88–.95 for the microphone and .92–.97 for the accelerometer. RFF estimated from an accelerometer is just as reliable as that derived from a microphone. These values were comparable to those found in the previous study of individuals with healthy voices by Lien and Stepp (2014) in which intrarater correlations were .90–.99 for the microphone and .91–.96 for the accelerometer. The interrater correlation coefficients found in this study were also quite similar to those found in the Lien and Stepp study.

Mean RFF

Statistically significant effects of sensor and the Vocal Cycle × Sensor, Vocal Cycle × Group, and Vocal Cycle × Sensor × Group interactions were found for utterance-level RFF means. The effect sizes of these factors were quite small ($\eta_p^2 \leq .01$), and a plot of the RFF means as a function of sensor and vocal cycle (see Figure 2) revealed that the typical RFF pattern (i.e., decreasing offset RFF that started at 0 ST and decreasing onset RFF that started at a positive value) was found for both the microphone and the accelerometer. However, the RFF means derived from the microphone were slightly higher compared with those derived from the accelerometer. The difference in RFF means was most notable in the offset vocal cycles, similar to the results of Lien and Stepp (2014). These results likely occur because the vocal cycles closest to the consonant that were concealed in the microphone due to coarticulation were captured in the accelerometer, leading to the observed lowered offset RFF.

Although the RFF pattern can be obtained using either a microphone or an accelerometer, the effect size of the Vocal Cycle × Sensor interaction was similar to the

Figure 3. Left: Utterance-level relative fundamental frequency (RFF) means as a function of group (control or voice disorder) and vocal cycle in semitones (ST) measured from microphone recordings. Right: Utterance-level RFF means as a function of group and vocal cycle in ST measured from accelerometer recordings. Error bars indicate 95% confidence intervals for the means.



effect size of the Vocal Cycle × Group interaction. This implies that the difference in RFF values estimated using different sensors can be as large as the difference in RFF values between individuals with and without voice disorders. The small effect size of the Vocal Cycle × Group interaction was anticipated because the statistical analysis was performed on all vocal cycles, including those that were close to the reference cycle (far from the voiceless consonant). The previous studies that also indicated significant effects of this interaction found that the largest difference in RFF means between the groups occurred in vocal cycles closest to the voiceless consonant (Stepp et al., 2010, 2011). Accordingly, one may choose to only examine the cycles closest to the voiceless consonant when using RFF to differentiate between groups (Eadie & Stepp, 2013). Given that the vocal cycles closest to the voiceless consonant are also the vocal cycles that differ the most between sensors, in order to provide valid information for voice assessments, the RFF measured using a microphone or an accelerometer must be compared with the standards established with the same sensor.

When the data collected from each sensor were independently examined by plotting the utterance-level RFF means as a function of group and vocal cycle (see Figure 3), the plot revealed that for both sensors, the RFF values in the voice disorder group were lower than those in the control group. The finding that individuals with voice disorders associated with vocal hyperfunction have lowered RFF values compared with those with healthy voices was consistent with the results from previous studies (Stepp et al., 2010, 2011). Our results support the validity of neck-skin acceleration for providing accurate RFF estimates in individuals with and without voice disorders.

RFF Standard Deviation

We found no statistically significant effects of sensor or the Vocal Cycle × Sensor, Sensor × Group, and Vocal Cycle × Sensor × Group interactions on utterance-level RFF standard deviations. This finding is likely due to the ability of the accelerometer to provide more robust estimates of fundamental frequency (lower RFF variability) and to reveal cycles closer to the voiceless consonant (higher RFF variability). Thus, its RFF variability is comparable to that derived from a microphone. In the study by Lien and Stepp (2014), significant effects of sensor and the Vocal Cycle × Sensor interaction were noted for RFF standard deviations, but the effect sizes of these factors were small ($\eta_p^2 \leq .01$). Overall, these outcomes imply that RFF can be both accurately and reliably estimated using an accelerometer.

Using Neck-Placed Accelerometry for RFF

At a glance, it may appear that there is no need for accelerometers in voice clinics because microphones are widely accessible in most clinics. Indeed, 75% of voice therapists have reported that they are likely to use acoustic measurements as a diagnostic tool (Behrman, 2005). Yet clinics can be rather noisy, and microphones tend to be more susceptible than accelerometers to environmental noise. The U.S. Environmental Protection Agency (1974) recommended that the indoor noise levels in hospitals should not exceed 45 dB SPL in the morning and 35 dB SPL at night. Yet studies have shown that sound pressure levels in hospitals frequently exceed the recommended levels and that the primary noise sources are inside the hospital (Aitken, 1982; Bayo, García, & García, 1995; Yassi, Gaborieau, Gillespie, & Elias, 1991). The mean sound pressure level in clinics has been found to be 59.4 dB SPL (Bayo et al., 1995). The signal-to-noise ratio of microphone recordings may be lower than the accelerometer recordings in these environments due to the high environmental noise levels.

Acoustic measures used in clinical assessments, such as sound pressure level and fundamental frequency, have also been adapted for ambulatory voice monitoring in devices such as the Ambulatory Phonation Monitor (KayPENTAX) and the VoxLog (Sonvox AB). Both of these devices contain miniature accelerometers that can be used for voice measurements. In addition to providing feedback on the basis of sound pressure level and fundamental frequency, our study indicates that the accelerometer in these devices can also be used to potentially provide feedback on the basis of RFF. The main obstacle to implementing RFF in voice monitoring is the lack of an algorithm for real-time estimation of RFF, so the data needs to be saved and transferred to a technician for analysis. Future work should explore automation of RFF estimation so that these ambulatory monitoring devices can be used to provide real-time RFF-based feedback to users. In conjunction, the physiological mechanisms behind RFF should also be examined to facilitate clinical interpretation.

Study Limitations

This study was conducted under quiet conditions, and the recording length was generally less than 5 min per participant. These conditions were advantageous for obtaining robust acoustic and acceleration signals for RFF estimation. In practice, the conditions may be different and the results of the effects of sensor type on RFF that we found in this study may not hold under practical conditions. As an illustration, the well-known Lombard effect, in which vocal intensity tends to rise with the increase in noise level, has been shown to not always hold for actual vocal behavior in the workplace (Lindstrom, Waye, Södersten, McAllister, & Ternström, 2011). Furthermore, fundamental frequency and sound pressure level have been shown to be higher when measured during vocal loading tests conducted in a clinical environment compared with measurements acquired in real teaching environments (Echternach, Nusseck, Dippold, Spahn, & Richter, 2014). Future studies are needed to determine whether our results are still valid under more practical clinical and ambulatory monitoring settings.

Acknowledgments

This work was supported by National Institute on Deafness and Other Communication Disorders Grants DC012651 (awarded to Cara E. Stepp) and DC011588 (awarded to Robert E. Hillman) and a New Century Scholars Research Grant from the American Speech-Language-Hearing Foundation (awarded to Cara E. Stepp). Thanks to Defne Abur, Christina Stevens, and Alexandra Martinson for assistance with data recording and to Melissa Cooke, Amanda Fryd, and Molly Bresnahan for help with recording segmentation. We disclose that author Robert E. Hillman has a financial relationship with Pentax Medical with respect to the Ambulatory Phonation Monitor.

References

Aitken, R. (1982). Quantitative noise analysis in a modern hospital. *Archives of Environmental Health: An International Journal, 37,* 361–364.

Andrews, M. L. (1996). Treatment of vocal hyperfunction in adolescents. *Language, Speech, and Hearing Services in Schools, 27,* 251–256.

Bayo, M. V., García, A. M., & García, A. (1995). Noise levels in an urban hospital and workers' subjective responses. *Archives of Environmental Health: An International Journal, 50,* 247–251.

Behrman, A. (2005). Common practices of voice therapists in the evaluation of patients. *Journal of Voice, 19,* 454–469.

Boersma, P., & Weenink, D. (2012). *Praat: Doing phonetics by computer: Version 5.3.04.* [Computer software]. Available from http://www.praat.org/

Cheyne, H. A., Hanson, H. M., Genereux, 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.

Eadie, T. L., & Stepp, C. E. (2013). Acoustic correlate of vocal effort in spasmodic dysphonia. *Annals of Otology, Rhinology & Laryngology, 122,* 169–176.

Echternach, M., Nusseck, M., Dippold, S., Spahn, C., & Richter, B. (2014). Fundamental frequency, sound pressure level and vocal dose of a vocal loading test in comparison to a real teaching situation. *European Archives of Oto-Rhino-Laryngology, 271,* 3263–3268.

Goberman, A. M., & Blomgren, M. (2008). Fundamental frequency change during offset and onset of voicing in individuals with Parkinson disease. *Journal of Voice, 22,* 178–191.

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., Holmberg, E. B., Perkell, J. S., Walsh, M., & Vaughan, C. (1989). Objective assessment of vocal hyperfunction: An experimental framework and initial results. *Journal of Speech and Hearing Research, 32,* 373–392.

Lien, Y.-A. S., Gattuccio, C. I., & Stepp, C. E. (2014). Effects of phonetic context on relative fundamental frequency. *Journal of Speech, Language, and Hearing Research, 57,* 1259–1267.

Lien, Y.-A. S., & Stepp, C. E. (2014). Comparison of voice relative fundamental frequency estimates derived from an accelerometer signal and low-pass filtered and unprocessed microphone signals. *The Journal of the Acoustical Society of America, 135,* 2977–2985.

Lindstrom, F., Waye, K. P., Södersten, M., McAllister, A., & Ternström, S. (2011). Observations of the relationship between noise exposure and preschool teacher voice usage in day-care center environments. *Journal of Voice, 25,* 166–172.

Mehta, D. D., Zanartu, M., Feng, S. W., Cheyne, H. A., & 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.

Morrison, M. D., Nichol, H., & Rammage, L. A. (1986). Diagnostic criteria in functional dysphonia. *Laryngoscope, 96,* 1–8.

Ohlsson, A.-C., Brink, O., & Lofqvist, A. (1989). A voice accumulator—Validation and application. *Journal of Speech, Language, and Hearing Research, 32,* 451–457.

Popolo, P. S., Scaronvec, 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.

Robb, M. P., & Smith, A. B. (2002). Fundamental frequency onset and offset behavior: A comparative study of children and adults. *Journal of Speech, Language, and Hearing Research, 45,* 446–456.

Roy, N., Ford, C. N., & Bless, D. M. (1996). Muscle tension dysphonia and spasmodic dysphonia: The role of manual laryngeal tension reduction in diagnosis and management. *Annals of Otology, Rhinology & Laryngology, 105,* 851–856.

Ryu, S., Komiyama, S., Kannae, S., & Watanabe, H. (1983). A newly devised speech accumulator. *Journal for Oto-Rhino-Laryngology, Head and Neck Surgery, 45,* 108–114.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86,* 420–428.

Stepp, C. E. (2013). Relative fundamental frequency during vocal onset and offset in older speakers with and without Parkinson's disease. *The Journal of the Acoustical Society of America, 133,* 1637–1643.

Stepp, C. E., Hillman, R. E., & Heaton, J. T. (2010). The impact of vocal hyperfunction on relative fundamental frequency during voicing offset and onset. *Journal of Speech, Language, and Hearing Research, 53,* 1220–1226.

Stepp, C. E., Merchant, G. R., Heaton, J. T., & Hillman, R. E. (2011). Effects of voice therapy on relative fundamental frequency during voicing offset and onset in patients with vocal hyperfunction. *Journal of Speech, Language, and Hearing Research, 54,* 1260–1266.

Stepp, C. E., Sawin, D. E., & Eadie, T. L. (2012). The relationship between perception of vocal effort and relative fundamental frequency during voicing offset and onset. *Journal of Speech, Language, and Hearing Research, 55,* 1887–1896.

U.S. Environmental Protection Agency. (1974). Information on levels of environmental noise requisite to protect public health and welfare with an adequate margin of safety (Report No. 550-9-74-0004). Washington, DC: Author.

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.

Watson, B. C. (1998). Fundamental frequency during phonetically governed devoicing in normal young and aged speakers. *The Journal of the Acoustical Society of America, 103,* 3642–3647.

Witte, R. S., & Witte, J. S. (2010). *Statistics* (9th ed.). Hoboken, NJ: Wiley.

Yassi, A., Gaborieau, D., Gillespie, I., & Elias, J. (1991). The noise hazard in a large health care facility. *Journal of Occupational and Environmental Medicine, 33,* 1067–1070.