Adaptation of a Pocket PC for Use as a Wearable Voice Dosimeter

THEORETICAL/REVIEW ARTICLE

Peter S. Popolo
National Center for Voice and Speech,
The Denver Center for the Performing Arts,
Denver, CO, and
University of Iowa, Iowa City

Jan G. Švec
National Center for Voice and Speech,
The Denver Center for the Performing Arts,
Denver, CO, and
University Medical Center Groningen,
Groningen, the Netherlands

Ingo R. Titze
University of Iowa, Iowa City, and
National Center for Voice and Speech,
The Denver Center for the Performing Arts,
Denver, CO

This article deals with the adaptation of a commercially available Pocket PC for use as a voice dosimeter, a wearable device that measures the vocal dose of teachers or other individuals on the job, at home, and elsewhere during the course of an entire day. An engineering approach for designing a voice dosimeter is described, and design data are presented. Technical issues include transducer selection, dynamic range, frequency response, memory requirements, power requirements, attachment, cables, connections, and data collection. Advantages and disadvantages of the design are discussed.

KEY WORDS: Pocket PC, voice dosimetry, voice accumulation

In voice research there is often a need for a simple, reliable, and inexpensive means of collecting voice data from research participants in the field, that is, in the environment where their voice activities normally occur. One example is in the vocally demanding profession of classroom teaching, in which there is a need to measure the amount of voice use among teachers and to monitor the occurrence of voice problems in order to establish occupational safety limits. Another example is in the area of voice therapy for individuals with Parkinson's disease, where there is a need to adapt clinician-directed treatment to home self-training while preserving feedback and clinician monitoring (Halpern, Matos, Ramig, Spielman, & Bennett, 2003). In these cases, the data collection instrument needs to be wearable (or at least portable), must be easily operated by someone with minimal technical aptitude, and in the case of teachers, must be able to be used during the course of normal workday or leisure time activities.

In both of these voice applications, commercially available Pocket PCs have been successfully adapted for the collection and real-time processing of voice data (Halpern et al., 2003; Popolo, Švec, Rogge-Miller, & Titze, 2002; Švec, Titze, & Popolo, 2003). This article describes the process by which a Pocket PC was adapted for use as a voice dosimeter in a study of voice use among teachers. This device could also be used to study voice in many different groups with similar requirements. Because there is no commercial counterpart to our customized device, sufficient technical detail is given here to help other researchers and clinicians who may be tempted to repeat the time-intensive process of developing such devices.

A voice dosimeter is a device that extracts and stores fundamental voice parameters from the wearer's voice signal for subsequent calculation of vocal dose. Vocal dose has been defined as vocal fold tissue exposure to vibration over time (Švec, Popolo, & Titze, 2003; Titze, Švec, & Popolo, 2003). The aim of voice dosimetry is to measure the intensity, frequency, and duration of a participant's vocal activity in terms of sound pressure level (SPL), fundamental frequency (F_0), and voicing time. Data collection is conducted over the entire day, starting as soon as practical after awakening and ending just before bedtime. The data can be used by speech scientists and clinicians to help determine the complex relationships among voice use, vocal fatigue, and recovery time.

The voice dosimeter can be considered a successor to the voice accumulator, prototypes of which were developed by different research groups during the last two decades (Buekers, Bierens, Kingma, & Marres, 1995; Ohlsson, Brink, & Löfqvist, 1989; Ryu, Komiyama, Kannae, & Watanabe, 1983; Szabo, Hammarberg, Håkansson, & Södersten, 2001). Because there was no commercially available version of the voice accumulator at the time this study was initiated, the decision was made to develop a custom device. The Compaq iPAQ 3765 Pocket PC was selected as the platform for our voice dosimeter from the available models on the basis of price and availability.

To determine the suitability of a Pocket PC for voice dosimetry, the following questions related to the characteristics of the Pocket PC hardware needed to be answered:

1. What is the proper transducer for recording the voice signal, and can such a transducer be integrated with the Pocket PC?
2. Is the dynamic range of the Pocket PC's sound capture hardware compatible with the dynamic range of voice?
3. What is the overall frequency response of the dosimeter, including the input transducer and the Pocket PC?

Because a typical day for the teachers in this study can easily exceed 12 hr, the following questions related to the long-term use of the device also needed to be addressed:

4. How much data storage is required, and does the Pocket PC provide enough storage capacity?
5. Is the device capable of operating for at least 12 hr on a single charge, or is an additional power supply needed? If so, what type should be used?

Finally, the following questions are related to the general design of the dosimeter:

6. What is the best means of attachment of the transducer to the participant?
7. What are the considerations in selecting the cable and connectors for the transducer?
8. How can the daily collection of a complete set of data be ensured?

An engineering approach to answering the above questions is described in the body of this article, and the technical data used to design the voice dosimeter are presented.

Throughout this article, the word *wearer* refers to the participant whose voice data are being collected by the dosimeter, and the word *end-user* refers to any of the researchers who analyze or use these data.

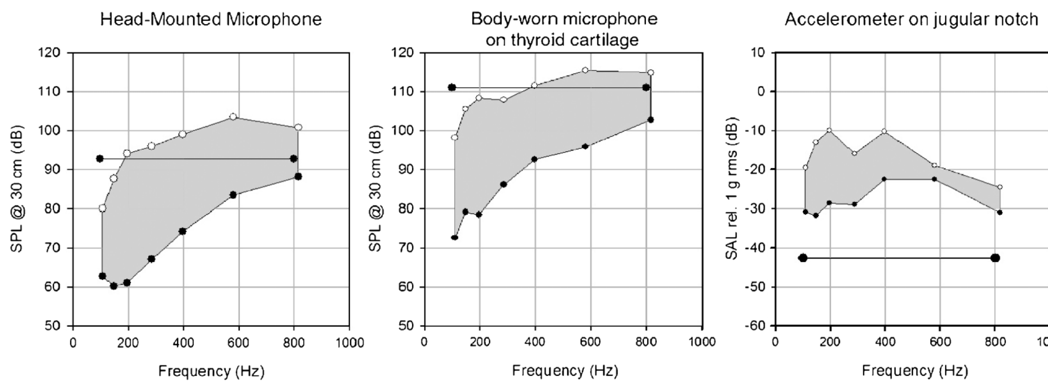
Method and Results

The format of this article departs from that of a standard research article in order to document the engineering design process used to develop the voice dosimeter. The Methods and Results sections are combined, and each of the above questions is treated separately as a stand-alone subsection. The experimental methods and equipment setup for measuring the device characteristics in Questions 1–3 are presented in the following sections, respectively, immediately followed by the result of each experiment: *Selection of a Transducer for Voice Dosimetry, Dynamic Range,* and *Frequency Response*. The *Memory Requirements* and *Power Requirements* sections address the main considerations involved in answering Questions 4 and 5, respectively. The problems associated with the attachment of the transducer to a participant (Question 6) and the selection of cables and connectors (Question 7) are addressed in the *Attachment* and *Cables and Connections* sections, respectively. Finally, some specific considerations in the collection of data in the field are described in the section entitled *Ensuring Completeness of Data* (Question 8).

Selection of a Transducer for the Voice Dosimeter

Various types of input transducers have appeared in the recent literature about voice accumulators: Buekers et al. (1995) used a head-mounted microphone (Sennheiser MKE 48 PU); Ryu et al. (1983) used a contact microphone with a gasket that, when taped to the anterior neck, maintained a sealed column of air between the skin and the microphone diaphragm; Ohlsson et al. (1989) used a piezo-ceramic contact microphone placed on the front of the participant's neck; Szabo et al. (2001) used a contact microphone attached to the front part of the neck to register vocal

Figure 1. Voice range profiles taken with three different transducer types and their relationship to the background noise (97 dB, six-talker babble), demonstrating the discrimination of each type of transducer. The horizontal line in each plot is the background noise level relative to the speaker's softest and loudest phonation levels. The body-worn microphone (b) appears to be even more sensitive to the background noise than the head-mounted microphone (a), which slightly suppresses it due to its placement closer to the mouth than the reference distance of 30 cm. Only the accelerometer (c) suppresses the background noise sufficiently so that its level is below the level of the softest phonations.



fold vibrations; and a miniature accelerometer (Knowles Electronics, Model BU7135) was used by Hillman and Cheyne (2003) and Cheyne, Hanson, Genereux, Stevens, and Hillman (2003).

An experiment was devised to test the usefulness of the signals registered by various transducers and to determine the amount of speaker discrimination versus background noise provided by each. We chose to use our own version of a head-mounted microphone, a body-worn microphone (instead of a contact microphone), and the Knowles Electronics BU-7134 accelerometer used by Hillman and Cheyne (2003) and Cheyne et al. (2003). The head-mounted microphone was made from an electret condenser microphone with an omnidirectional pick-up pattern (Vivanco EM-116), with a mouth-to-microphone distance of about 5 cm. This microphone was used instead of one with a directional pickup pattern to avoid the "proximity effect," which causes a boost of lower frequencies at close distance (Švec, Popolo, & Titze, 2003). The body-worn microphone was a homemade version of a "body mic" (such as those used by stage actors, typically taped to the actor's temple or clipped to an article of clothing), using an inexpensive omnidirectional condenser microphone element (Radio Shack, Part Number 270-084) taped to the skin over the thyroid cartilage with Johnson & Johnson First Aid All-Purpose Cloth Tape. The accelerometer was taped to the skin at the jugular notch (anterior part of the neck below the larynx, between the cricoid cartilage and the sternum), also with cloth tape.

A participant (the second author) was seated in an IAC Sound Booth, approximately 2.3 m^3, wearing the head-mounted microphone, the body-worn microphone, and the accelerometer. The cables from all three transducers were led out of the sound booth, and each transducer was connected to a separate channel of a Kay Elemetrics Computerized Speech Laboratory (CSL) 4400 system. All channels were sampled at 44100 Hz.

A Brüel & Kjær 2238 sound level meter, set to linear frequency weighting and fast frequency response, was positioned 30 cm in front of the participant inside the booth and connected to the fourth channel of the CSL system during the SPL calibration procedure. The sound level meter was calibrated to absolute SPL using a Brüel & Kjær 4231 calibrator, and the participant then produced a sustained /a/ recorded simultaneously in CSL with the head-mounted microphone and the sound level meter at 30 cm. In this way, the microphone signal could ultimately be related to SPL at 30 cm. A detailed description of the calibration procedure is given in a publication by Švec, Popolo, and Titze (2003). The body-worn microphone, not yet attached to the participant, was placed next to the sound level meter microphone and calibrated at the same time. The root-mean-squared (RMS) skin acceleration level (SAL) measured by the accelerometer was calibrated in deibels relative to 1 g (the acceleration due to gravity on the earth's surface), using the manufacturer's published sensitivity data of the accelerometer.

With all three transducers attached, a simplified voice-range profile (VRP) of the participant was measured, consisting of soft and loud phonations on the vowel /a/ for different pitches (about three pitches per octave). After the VRP was completed, six-talker babble (simulated background noise consisting of the speech of six simultaneous talkers) was played over a loudspeaker with an SPL of 97 dB (linear frequency weighting scale) to determine how the background noise is registered by the different sensors. The 97 dB SPL was adjusted before the experiment. It was measured with the sound level meter placed at the location of the participant's neck when seated in front of the loudspeaker and was kept constant during the experiment. The results are shown in Figure 1a–1c. The soft phonation data points (filled circles), in decibels relative to 30 cm, are at the bottom of each curve, and the loud data points (empty circles) are at the top. The background noise level picked up by the transducer in each case is shown as a horizontal line. Figure 1 shows that only the loudest phonations can be detected above the background noise for both the head-mounted microphone and the contact microphone. However, for the accelerometer, the entire VRP is 10–30 dB above the noise (picked up by the accelerometer as induced vibration in the participant's body). The accelerometer was thus selected as the input transducer of choice for the voice dosimeter. The ramifications of this choice are further discussed in the Discussion and Conclusions section below.

Dynamic Range

The dynamic range of the Pocket PC's sound capture hardware was measured by the direct input of a voltage signal from an HP8904A multifunction synthesizer to the microphone input of the Pocket PC via the modified input connector. In this way, the measurement was independent of the type of input transducer used. The input signal was a sinewave with a frequency of 500 Hz, and the amplitude was varied over a 60-dB range of 5 millivolts (mV) to 0.005 mV in 2-dB steps. The recorded WAV file was analyzed in MATLAB (Mathworks, Inc.).

Figure 2 shows the resultant curve of the dynamic range of the sound capture hardware at 500 Hz. The linear region of the curve corresponds to a total range of 36 dB (about 0.012 mV to about 0.8 mV), when the automatic gain control (AGC) feature of the device was disabled. It was necessary to turn off the AGC feature in order to measure the true amplitude variations in the voice signal; otherwise, the gain of the Pocket PC's sound capture hardware would vary inversely with the intensity of the input signal to keep the recording amplitude nearly constant. The maximum input amplitude that resulted in a linearly related output (corresponding to the origin in the graph of Figure 2) was found to be about 0.8 mV, and the minimum input amplitude that could be measured before it was lost in the device's noise floor was about 0.012 mV. This noise threshold was about 1.5% of maximal input level. A 10-dB attenuator was attached to the accelerometer in order to fit the dynamic range of the Pocket PC to the voice range picked up by the accelerometer (i.e., the

Figure 2. Dynamic range of the iPAQ 3765 Pocket PC sound capture hardware at 500 Hz.

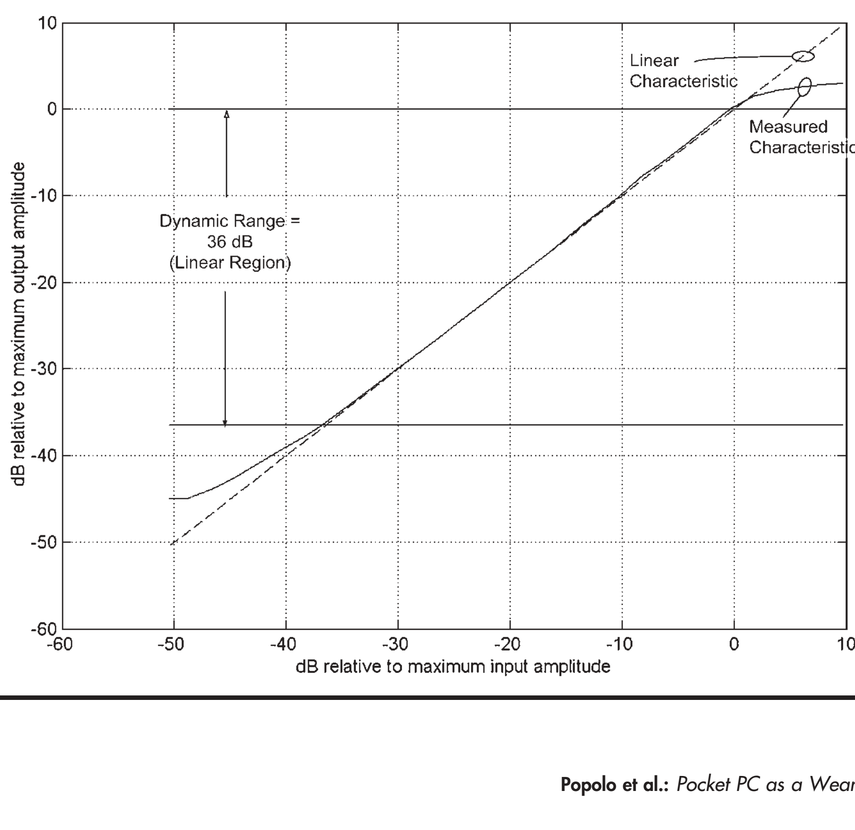
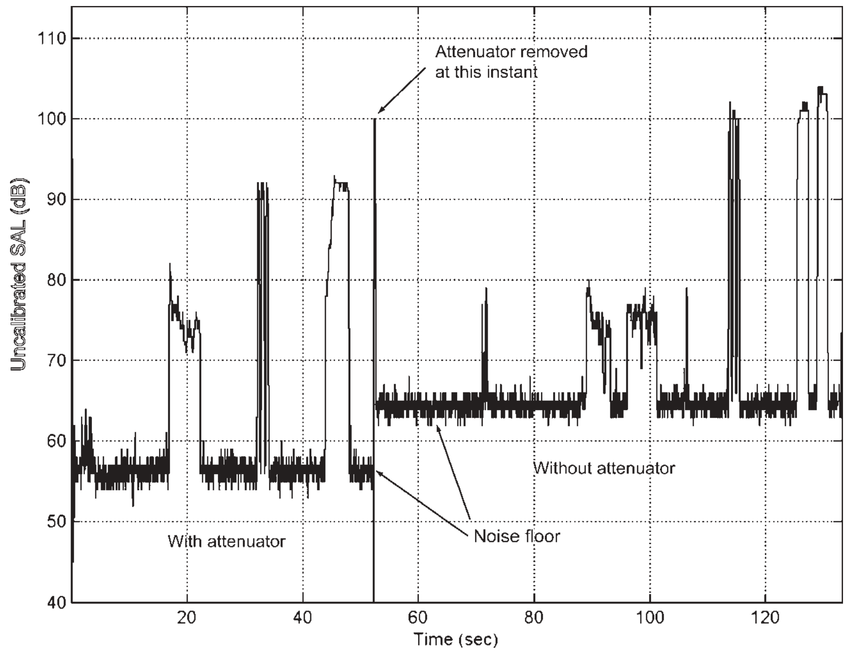


Figure 3. Noise floor of Pocket PC and accelerometer with and without attenuator.



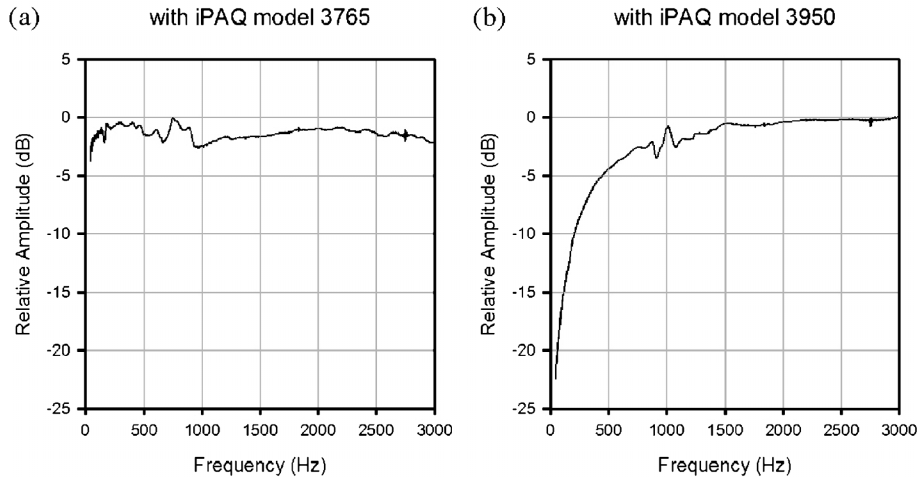
voice signal was attenuated so that loud speaking would not exceed the Pocket PC's 36-dB input range). The attenuator was also found to lower the Pocket PC's noise floor, which increased the useable dynamic range by a few decibels. Figure 3 shows about 2 min of a recording of normal and moderately loud phonations, both with and without the attenuator. It is seen that in the first 50 s, recorded with the attenuator in-line, the overall noise floor was lowered, as was the level of the recorded voice signal.

Frequency Response

The frequency response was measured for the combination of the Pocket PC, the attenuator, and the accelerometer. A calibrated vibration reference was required to provide a swept-frequency input to the accelerometer. This reference was constructed from a RadioShack 6.5 in. woofer (Catalog Number 40-1033) driven by a MATLAB-generated linearly swept signal, from 45 Hz to 3000 Hz, whose amplitude was proportional to the inverse of the speaker element's frequency response in order to present a constant-amplitude signal to the transducer over the entire frequency range (the frequency response of the speaker element was obtained experimentally by using a PCB Electronics Model 252C22 calibrated accelerometer, with a flat frequency response over the measurement range, attached directly to the center of the speaker cone). The measured accelerometer was firmly attached to the top surface of the calibrated accelerometer so that they both moved as one mass when the speaker was vibrating. A signal corresponding to the acceleration level versus frequency was recorded by the Pocket PC as a WAV file. The frequency response of the dosimeter recording system was obtained by comparing the recorded levels to those registered by the calibration accelerometer, recorded with the CSL system.

Figure 4a shows the combined frequency response of the dosimeter, consisting of the Compaq iPAQ Model 3765 Pocket PC, the attenuator, and the accelerometer. The amplitude variation is within ±2 dB (4 dB total) over the frequency range 45 Hz to 3000 Hz. Since the skin is known to act as a low-pass filter (Hess, 1983), and the acceleration signal transduced though the skin has little energy above 1000 Hz (Cheyne et al., 2003), the upper limit of 3000 Hz chosen for the frequency response measurement includes all the frequencies of interest for the voice dosimetry application. In Figure 4b, the frequency response of a dosimeter built using a newer model Pocket PC shows a very sharp roll-off for frequencies below about 500 Hz, which would have an adverse effect on the measurement of the SPL of voice signals. This difference in the frequency response of two different models of Pocket PCs from the same manufacturer demonstrates the importance of measuring the characteristics of a commercial device when adapting it to the purpose of scientific data collection in voice and speech.

Figure 4. Frequency response of dosimeter (Pocket PC + attenuator + accelerometer) with two different Compaq iPAQ models. In (a) Model 3765 was used, which has a relatively flat response of less than ±2 dB variation over 45 Hz to 3000 Hz, and in (b) Model 3950 was used, which has a very sharp roll-off at frequencies below about 500 Hz.



Memory Requirements

The goal of the dosimeter is to store the SAL, F_0, and a third parameter that provides information about the accelerometer spectrum (spectral centroid, or the so-called "frequency energy center" [FEC] described by Švec, Titze, & Popolo, 2003). These parameters are measured in real time and stored every 30 ms in a binary file for postprocessing. Each parameter is stored as an unsigned 8-bit integer (1 byte). Supplementary data, such as session start time, stop time, number of frames processed, self-rating test answers, power status, and error messages, are stored in separate binary and text data files and are used for the subsequent calculation of dose as well as to determine overall data integrity. These supplementary data files typically require less than 10 KB of storage per day. All file names consist of the participant's code name, the start date, and time in HH:MM:SS format (e.g., M001_01-05-2004_073519_SPL-F0.bin).

The amount of storage required is thus 3 bytes per 30 ms frame, or 100 bytes/second. In 1 hr, this is equal to 352 KB (1 KB = 1024 bytes), which in a 14-hr day totals about 5 MB. With 32 MB of RAM, the Pocket PC can store about 5 days of data before the memory exceeds 80% capacity; in practice, the data will be retrieved from the Pocket PC every 1 or 2 days to prevent a catastrophic loss of data. The data files are transferred to a PC or laptop using Microsoft ActiveSync, the built in file-transfer protocol used by the Pocket PC.

Power Requirements

To address the problem of extended operation of the device between chargings, two different external power supplies were investigated. The first one was a custom-made battery module that takes four disposable AA batteries. It was found that the voltage applied to the Pocket PC needed to be limited with a 5.1-volt zener diode, as shown in Figure 5; otherwise, the voltage from a fresh set of four batteries (exceeding 6 volts) would cause an increase in the electronic noise floor of the device, as shown in Figure 6. This noise would contaminate the audio input signal, affecting the SPL and F_0 measurements. Limiting the voltage of the fresh batteries to V_{zener} eliminated the noise

Figure 5. Electric diagram of the custom-made battery module for the dosimeter, showing a zener diode shunted across the series connection of four disposable AA batteries, limiting the voltage applied to the Pocket PC to 5.1 volts.

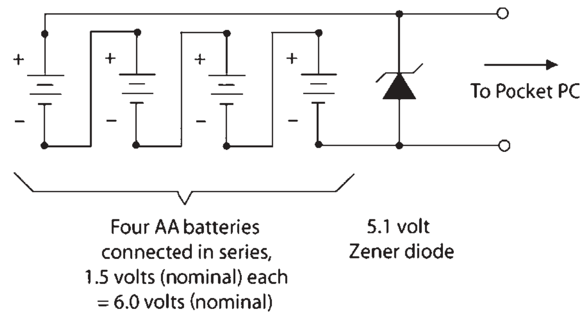
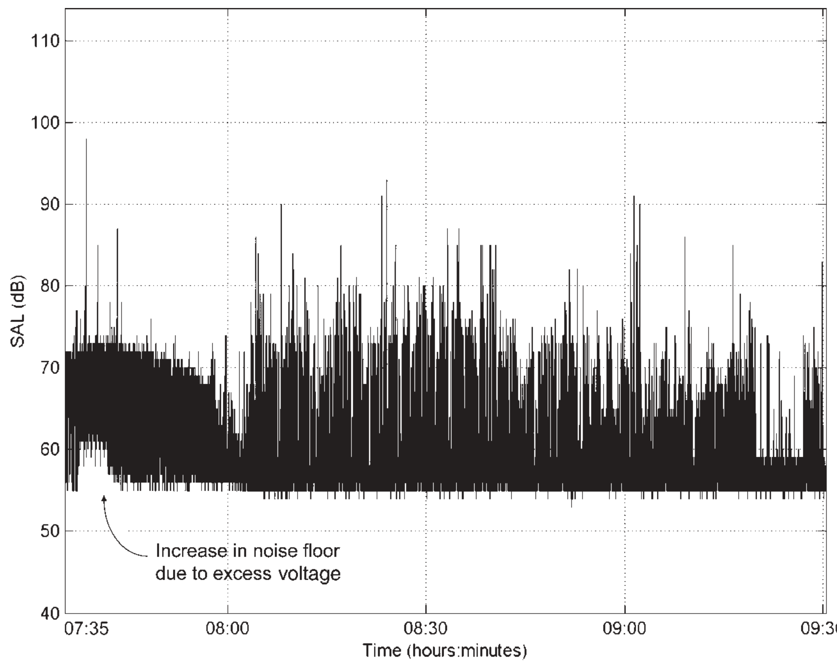


Figure 6. Increase in noise floor due to the excessive voltage from the power supply with four fresh AA batteries without the zener diode, seen in the first 30 min of the recording. As the batteries become depleted, the proper voltage is achieved and the noise floor drops, no longer interfering with the skin acceleration level (SAL) measurement. The dark downward-curving band up to Time 08:00 is noise with a few signal spikes. From Time 08:00 to 09:30, the noise band is combined with the signal, which is seen as many vertical spikes.



problem but reduced the operation time to about 8 hr. Still, this was a great improvement over the time obtainable on internal battery alone (2–3 hr). This passive approach to limiting the voltage was chosen for its simplicity. It should be noted, however, that it is not as efficient with respect to battery life as an active circuit such as a voltage regulator. For a full day of data collection, the batteries need to be changed by the wearer at about midday.

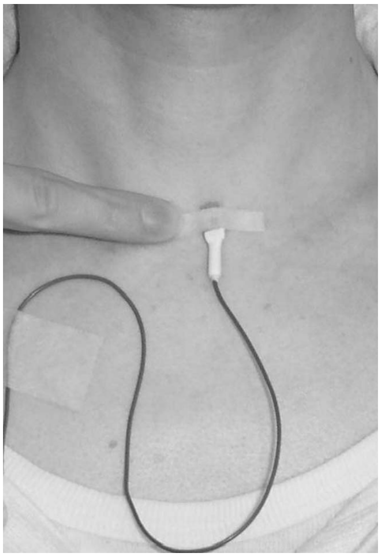
Another solution that worked well was to use a commercially available rechargeable module, manufactured by Unity Digital for use with digital cameras. The output of this module is much closer to the required voltage and current for proper operation of the Pocket PC, but noise is sometimes present in the first few minutes of connecting a fully charged module to the voice dosimeter. This device has been used for as long as 17 hr.

Attachment

When using the accelerometer, a good contact with the skin is needed to ensure a reliable signal. Two types of adhesive were evaluated: Skin-Bond® (manufactured by Smith+Nephew), the use of which was reported by Cheyne et al. (2003), and Mastisol™ (manufactured by Ferndale Laboratories), a surgical adhesive used for wound closure.

With the Skin-Bond, the consistency when first applied is thin liquid, but it eventually dries to a rubber-cement-like consistency. It was found that proper application required waiting about 15 min for the glue on both the skin and the accelerometer to become tacky before the two could be adhered to each other; then it was necessary to apply pressure for about 5 min to ensure proper adhesion. It was found during field-testing that the attachment would sometimes weaken during the course of the day, creating a gap between the accelerometer and the skin, which degraded the recorded signal. The Mastisol was found to provide a more secure connection of the accelerometer to the skin, and was thus preferred by the authors of this article. Similar to spirit gum in consistency and strength of adhesion, there was almost no waiting time for it to become tacky, and one needed to apply pressure for only about a minute. When worn by two of the authors during field testing, the bond between the accelerometer and the skin lasted the entire day. A suture strip adhesive band was used in conjunction with the Mastisol to prevent inadvertent pulling from the cable strain or clothing. Figure 7 shows a photo of the accelerometer attachment to the wearer.

Figure 7. Photograph showing the attachment of the accelerometer to the wearer, using Mastisol™ surgical adhesive and a suture strip to secure the accelerometer to the skin of the anterior neck. A small loop in the cable is secured with hypoallergenic adhesive tape to prevent cable strain.



Cables and Connections

The cable connecting the accelerometer to the Pocket PC runs beneath the clothing from the jugular notch to a small pack worn around the waist; thus, the cable is required to be flexible and rugged. Because of its length (5 ft.), the cable also needs to be shielded for immunity to electronic noise while producing minimal strain on the accelerometer attachment. The cable chosen for the voice dosimeter is a single-conductor shielded cable with an outer diameter of about 1 mm, manufactured by Cooner Wire (Part No. NMUF 1/30-4046SJ).

A connector between the accelerometer and the Pocket PC allows the transducer to be replaced if necessary. The main requirements for the connector are ruggedness, security (i.e., a secure connection that would not become unintentionally disconnected), and reliability (i.e., consistently makes good contact, and would not wear out). A Mini-XLR connector, such as those used to connect a small body mic worn by an actor on stage to a wireless transmitter, was chosen as the connector for the voice dosimeter. The connector is about 4 cm long $\times$ 1 cm in diameter, weighs less than 10 g, and features a locking mechanism with excellent alignment of contacts. The 3-pin connector style, commonly used for a balanced microphone connection, was chosen and the extra pin was used as a redundant ground for the unbalanced connection to the accelerometer. This approach provided good protection against intermittent contact or failure of a solder joint within the connector.

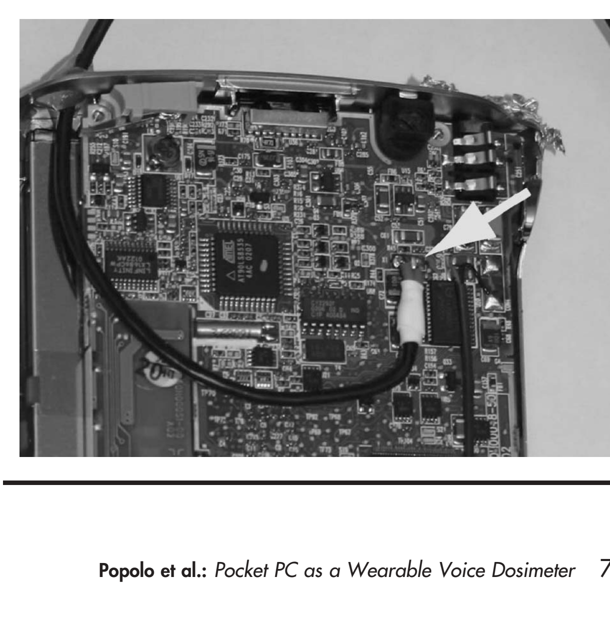
Figure 8 shows a photo of the internal modification of the Pocket PC to include the external cable and connector.

Ensuring Completeness of Data

Several features were included in the dosimeter software to aid the wearer in the operation of the device, and to ensure that the data were being collected properly (or at least alert the wearer or the end-user of a possible problem). One such feature was the disablement of the hardware buttons on the front and side of the Pocket PC that provide access to some of the native applications, to prevent these from interfering with the running of the dosimeter application. Another was the use of dialogue boxes or windows that appeared at 2-hr intervals throughout the data collection session, prompting the wearer to perform various self-rating tasks or "housekeeping" activities. The wearer was forced to select a response to each task before a window closed and the next one appeared, and the window remained visible until the wearer responded (but SPL and F_0 data collection continued during this time). The self-rating tasks required the wearer to test and rate his or her soft-voice quality (Bastian, Keidar, & Verdolini-Marston, 1990), effort to speak loudly, and level of laryngeal discomfort. When the tasks were completed, the time-stamped self-rating data were written to a log file for that session.

The housekeeping tasks were designed to provide the end-users with some information about the integrity of the data being collected throughout the day. The wearer was prompted to perform these at the

Figure 8. Internal modification of the Pocket PC, showing the attachment (at the arrow) of an external cable for the accelerometer.



same time as the self-rating tasks. These interactive tasks included a noise test and a counting test. The time-stamped results of these tests were also stored in the session log file. The noise test prompted the wearer to maintain 3 s of silence while the device's intrinsic noise floor level was measured and stored. The counting test asked the wearer to count loudly while a test was run to ensure that the accelerometer signal exceeded the noise floor by a predetermined amount. If the signal failed to exceed this amount, the wearer was prompted to check that the accelerometer was connected and properly attached. In this case, the stored results would provide the end-user with an indication (after the fact) that the accelerometer may have become unattached or been disconnected some time within the 2-hr period between the tests.

A safeguard was implemented in the dosimeter software to prevent the wearer from unintentionally stopping the dosimeter program by, for example, an inadvertent button press. This safeguard required the wearer to go through a three-step process to exit the application. First the wearer had to select a checkbox on the touch-screen to display a "Stop Recording" soft-button, after which an alert window appeared, making sure the command to exit was intentional. The wearer had to respond "Ok" to this window before the program would shut down; otherwise, the application continued to run uninterrupted.

A simplified flowchart of the dosimeter software, showing the sequence of wearer interactions with the device throughout the data collection session, is shown in Figure 9.

Figure 10 shows a sample of the processed data obtained from a full day of measurement with the Pocket PC voice dosimeter (the participant was a 38-year-old vocally healthy male member of the National Center for Voice and Speech staff in Denver, CO, measured on a moderate-to-high voice-use day). In addition to the measured parameters of SAL and F_0 (see Figure 10a and 10b), four different representations of dose are plotted versus time. Figure 10c shows the time dose per minute (D_t/min), that is, the amount of voicing time (in seconds) for each minute of the day. Voicing occurred for 5,082 s out of the 36,943 s measured

Figure 9. A simplified flowchart of the dosimeter software, showing the sequence of wearer interactions with the device throughout the data collection session.

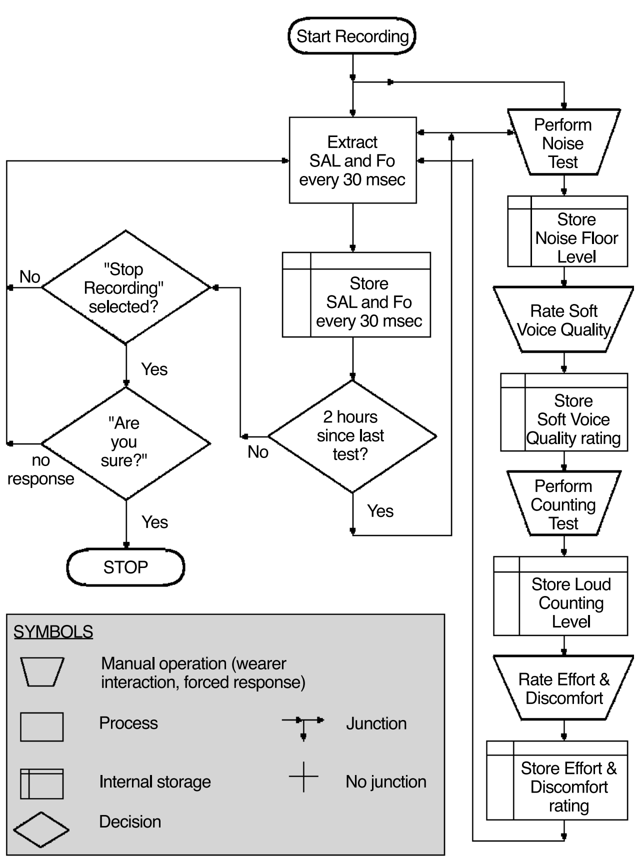
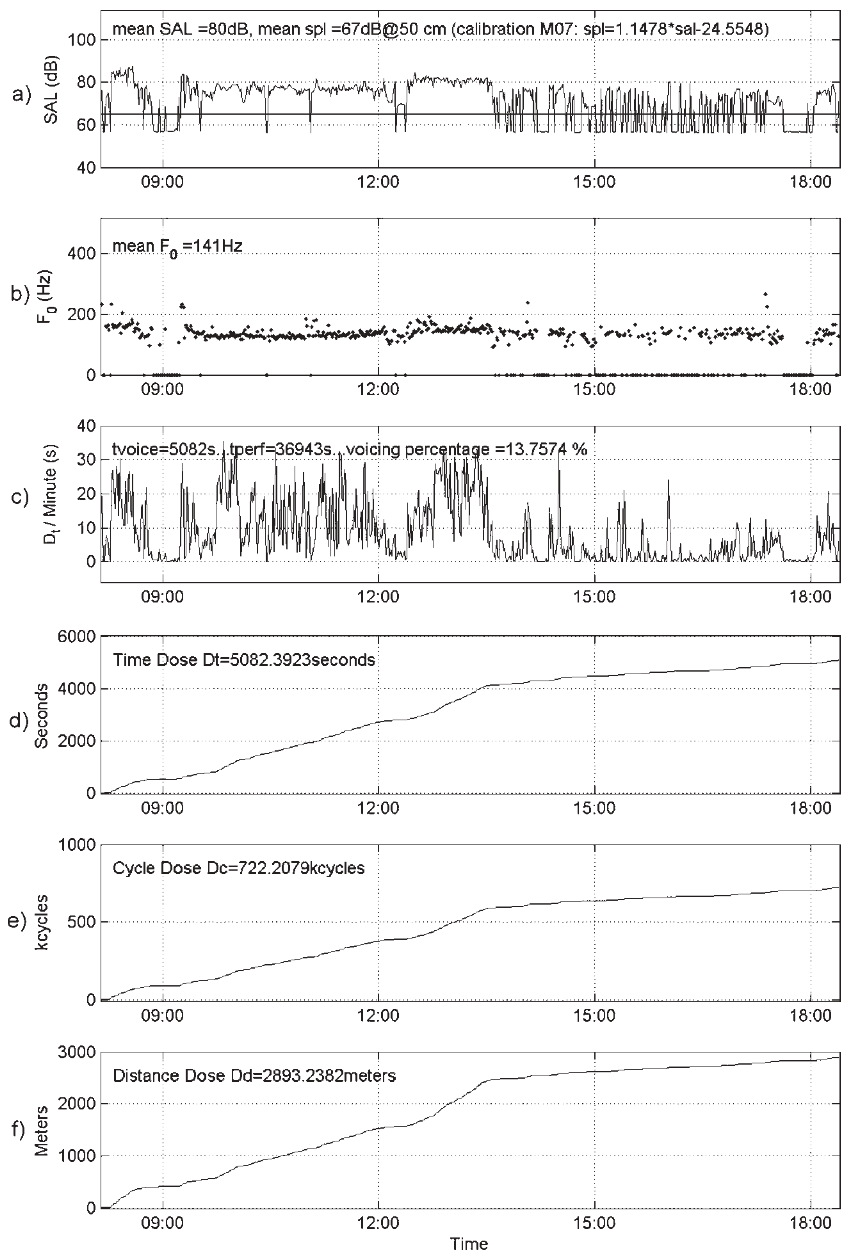


Figure 10. Processed data from a full day of measurement with the Pocket PC voice dosimeter. Each data point in plot (a) is the average SAL in decibels per minute (the dosimeter measures and stores SAL and F_0 data at a rate of once every 30 ms). Thus, the mean SAL of 80 dB also takes into account the actual number of seconds of phonation per minute, shown as D_t/minute in plot (c). Mean SPL (in decibels) is derived from the mean SAL according to the SAL-to-SPL calibration curve obtained for each participant. Plot (b) shows the F_0 contour, where each point is the average F_0 per minute. Plots (d)–(f) are the cumulative doses computed in terms of voicing time (in seconds), number of cycles of vocal fold oscillation (in kcycles), and distance traveled by the vocal folds in a vibratory trajectory (in meters).



throughout the day (i.e., roughly 1.5 hr out of 10 hr), corresponding to 13.75% voicing. Figure 10d shows the accumulation of voicing time during the entire day (time dose), which totals to 5,082 s. The corresponding cumulative cycle dose (number of cycles of vocal fold oscillation) is 722 kcycles (see Figure 10e), and the cumulative distance dose (distance traveled by the vocal folds in a vibratory trajectory) is 2,893 m (see Figure 10f). For a more complete description of the doses, see Titze et al. (2003) and Švec, Popolo, and Titze (2003).

During field tests of the dosimeter, it was found that inadvertent depression of the Pocket PC's external speaker, which also functions as a mouse button, produces the same effect as the soft "Ok" inside an application (i.e., closes the application). A plastic cover provided by the Pocket PC manufacturer as an accessory (see Figure 11) was found to be useful for protecting the dosimeter from this type of inadvertent shutdown.

Discussion and Conclusions

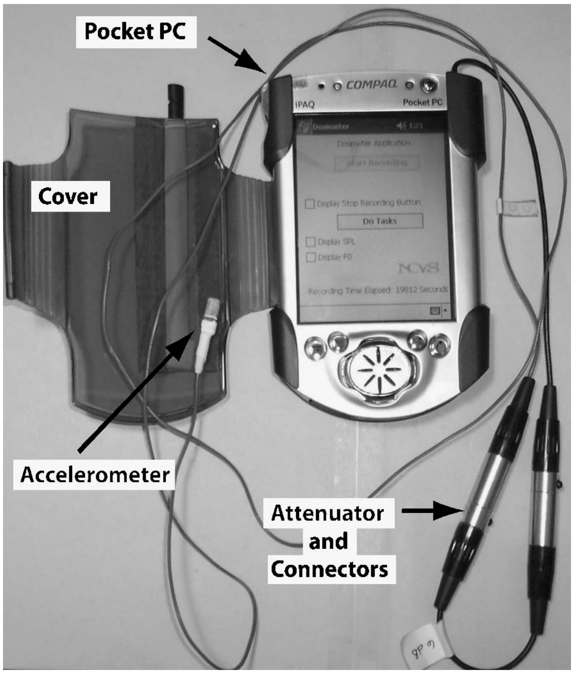
The final version of the Pocket PC voice dosimeter is shown in Figure 11. By using the features of a Pocket PC, the design task was greatly simplified with respect to building a custom device. These features include a powerful processor, a large amount of storage capacity, an operating system based on the familiar Microsoft Windows, an existing file-transfer protocol for moving data to a laptop or desktop computer, a touch-screen interface, built-in digital sound capture hardware, and a Microsoft-supported software developer's toolkit for writing application programs. The latter feature allows an inexpensive and rapid means of altering the dosimeter software as needed during the development and test phases.

A disadvantage associated with adapting a commercially available Pocket PC is the rapid "turnover" time by which consumer devices are updated. There is a risk that an important feature or characteristic of the device may change or be deleted when a newer model is introduced, impacting its use for the intended application. This was illustrated by the change in the frequency response of a newer model Pocket PC relative to the first model used for our application. Thus, it was important to completely characterize the device for the initial adaptation, as well each new model. This disadvantage was not a major concern for the present study, because the data collection phase was known to be of a finite duration and the number of participants was more or less fixed prior to the development of the device. Researchers facing a long-term data collection period, or those wishing to adapt a device to more general purposes, may need to consider a custom design, over which they would have more control.

The choice of an accelerometer as the input transducer for the voice dosimeter has both advantages and disadvantages. Unlike a microphone, an accelerometer picks up a vibration signal from a surface with which it is in direct contact. Thus, one advantage of an accelerometer is that errors due to background noise and speech are effectively eliminated. Also, the reduced harmonic content of an accelerometer signal makes frequency extraction easier than from a microphone signal. Yet another advantage is that the accelerometer signal can be used as a reliable means of voicing detection (since unvoiced signals are transmitted through the skin with significantly less energy than voiced signals). A disadvantage is that SAL must be converted to equivalent SPL for our dose calculations. This can be done by obtaining an SAL-to-SPL calibration curve for each participant for postprocessing of the data (Švec et al., 2005). An unexpected problem that was encountered with the accelerometer was that it picked up unwanted movement artifacts, including walking, climbing stairs, and riding in a motor vehicle, in addition to the desired voice signal, contaminating the dosimetry data. The solution to this problem will be described in a future paper.

In this article we have shown how a Pocket PC was successfully adapted for a specific application, namely, the collection and real-time processing of voice data for voice dosimetry of teachers. The individual components involved in sound capture have been characterized, and the requirements for use by a specific group of participants have been met. We expect that the Pocket PC can be adapted to other applications that go beyond voice accumulation. In particular, the device is sensitive to body noises of various types, including neck movement, heartbeat, and whole body vibration in activities such as walking and running. Furthermore, with the use of physiologic sensors (electrostatic, electromagnetic, etc.), many body functions could be monitored within and outside the context of voice and speech.

Figure 11. The Pocket PC voice dosimeter and its individual components.



Acknowledgments

This work was supported by National Institutes of Health Grant 5R01DC04224. Special appreciation is expressed to Sten Ternström for the original suggestion to use a Pocket PC.

References

Bastian, R. W., Keidar, A., & Verdolini-Marston, K. (1990). Simple vocal tasks for detecting vocal fold swelling. *Journal of Voice, 4,* 172–183.

Buekers, R., Bierens, E., Kingma, H., & Marres, E. H. M. A. (1995). Vocal load as measured by the voice accumulator. *Folia Phoniatrica, 47,* 252–261.

Cheyne, H. A., Hanson, H. M., Genereux, R. P., Stevens, K. N., & Hillman, R. E. (2003). Development and testing of a portable voice accumulator. *Journal of Speech, Language, and Hearing Research, 46,* 1457–1467.

Halpern, A. E., Matos, C., Ramig, L., Spielman, J., & Bennett, J. (2003, March 13–15). *PDA-enhanced speech treatment for Parkinson's disease*. Poster presented at the Fifth Annual Meeting of the American Society for Experimental NeuroTherapeutics, Washington, DC.

Hess, W. (1983). *Pitch determination of speech signals: Algorithms and devices*. Berlin: Springer-Verlag.

Hillman, R., & Cheyne, H. (2003). A portable vocal accumulator with biofeedback capability. In G. Schade, F. Müller, T. Wittenberg, & M. Hess (Eds.), *AQL Hamburg: Proceeding papers for the conference Advances in Quantitative Laryngology, Voice and Speech Research* (CD-ROM; p. 4p). Stuttgart, Germany: IRB Verlag.

Mathworks, Inc. (2002). Matlab (Version 6.5, Release 13) [Computer software]. Natick, MA: Author.

Ohlsson, A-C., Brink, O., & Löfqvist, A. (1989). A voice accumulator—Validation and application. *Journal of Speech and Hearing Research, 32,* 451–454.

Popolo, P. S., Švec, J. G., Rogge-Miller, K., & Titze, I. R. (2002, December 2–6). *Technical considerations in the design of a wearable voice dosimeter*. Poster presented at the First Pan-American/Iberian Meeting on Acoustics/144th Meeting of the Acoustical Society of America, Cancun, Mexico. Available from http://www.ncvs.org/ncvs/about/people/popolo/research.html

Ryu, S., Komiyama, S., Kannae, S., & Watanabe, H. (1983). A newly devised speech accumulator. *Journal of Otorhinolaryngology and Related Specialties, 45,* 108–114.

Švec, J. G., Popolo, P. S., & Titze, I. R. (2003). Measurement of vocal doses in speech: Experimental procedure and signal processing. *Logopedics, Phoniatrics and Vocology, 28,* 181–192.

Švec, J. G., Titze, I. R., & Popolo, P. S. (2003). Vocal dosimetry: Theoretical and practical issues. In G. Schade, F. Müller, T. Wittenberg, & M. Hess (Eds.), *AQL Hamburg: Proceeding papers for the conference Advances in Quantitative Laryngology, Voice and Speech Research* (CD-ROM; p. 8p). Stuttgart, Germany: IRB Verlag.

Švec, J. G., Titze, I. R., & Popolo, P. S. (2005). Estimation of sound pressure levels of voiced speech from skin vibration of the neck. *Journal of the Acoustical Society of America, 117,* 1386–1394.

Szabo, A., Hammarberg, B., Håkansson, A., & Södersten, M. (2001). A voice accumulator device: Evaluation based on studio and field recordings. *Logopedics, Phoniatrics and Vocology, 26,* 102–117.

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,* 922–935.

Received March 12, 2004

Accepted January 14, 2005

DOI: 10.1044/1092-4388(2005/054)

Contact author: Ingo R. Titze, National Center for Voice and Speech, The Denver Center for the Performing Arts, 1101 13th Street, 4th Floor, Denver, CO 80204. E-mail: ititze@dcpa.org