

## **General Statistics of the NCVS Self-Administered Vocal Rating (SAVRa)**

Eric J. Hunter

National Center for Voice and Speech,  
The Denver Center for the Performing Arts, 1101 13<sup>th</sup> Street, Denver, CO 80204, USA,

*The ability to measure and track vocal fatigue has been difficult. Subjective ratings and described symptoms have been the main indicators of vocal fatigue. One metric that has been used in several studies at the NCVS has been the SAVRa; a combination of three self administered subjective vocal ratings; level of laryngeal discomfort, vocal effort level, and the inability to perform a specific high pitch yet soft vocal task. This memo describes the current instructions and usage of the SAVRa, as well as general statistical results. Updates to this memo can be downloaded at <http://www.ncvs.org/ncvs/library/tech> .*

**Keywords:** vocal fatigue, subjective ratings, vocal fold swelling, soft voice, voice break, vocal effort, vocal loading.

### **1. Prelude: Vocal Fatigue?**

A common complaint of occupational voice users, and other heavy voice users, is a tired voice, or *vocal fatigue*. Nevertheless, vocal fatigue is neither an easily measurable symptom of a disease nor a phenotype of a genetic disposition. In their review of vocal fatigue, Welham and Maclagan (2003) state that “a link between vocal fatigue and other laryngeal pathologies is plausible, [but] it is unclear whether vocal fatigue primarily contributes to, results from, or exists independently of other voice conditions” (emphasis added). Adding to the confusion, the classification of vocal fatigue is used by patients as a lay description as well as by practitioners as a clinical diagnosis.

In this document, we assume that vocal fatigue is the result of an overused or under-recovered voice caused by excessive tissue vibration or repeated vocal fold posturing. Titze (1999) describes at least two potential aspects of this type of vocal fatigue: [1] laryngeal muscle fatigue, or reduction in short-term strength and speed of contraction as the muscle chemistry is reset and byproducts of muscle contractions are removed; and [2] laryngeal tissue fatigue, or damage to the lamina propria from excessive phonation (i.e., vocal fold vibration), sometimes called phonotrauma.

Laryngeal muscle fatigue is likely linked, at least in part, to the fact that phonation during everyday occupational voice use causes these muscles to be engaged in adduction and abduction more than 1800 times an hour (Titze et al., 2007). This type of muscle engagement might be

compared to repetitive submaximal isotonic contractions in larger muscles, in which 90% muscle recovery after fatigue from repeated contractions is complete in about 20 minutes (e.g., Yates et al., 1987). Nevertheless, while laryngeal muscles should also have measurable fatigue and recovery, it is not clear that they can be directly compared to larger skeletal muscles. For example, laryngeal muscles are fast twitch and slow fatiguing given their necessity for airway protection (e.g., Hunter and Titze, 2007).

The second component of vocal fatigue, or laryngeal tissue fatigue, likely stems from damage caused by vocal fold lamina propria vibration. High biomechanical stress during excessive phonation causes tissue damage as epithelial cells die and are shed, or as collagen and elastin fibers separate from the structural matrix at the sub-microscopic level (Gray and Titze, 1988). Molecular disruption, and at times biomechanical changes to tissue vibratory characteristics, could also be caused by fluid redistribution within tissue during phonation (Zhang et al., 2008) or by inflammation and fluid leakage from the vascular system due to vibration-induced high intravascular pressure (Czerwonka et al., 2008).

If vocal fatigue has both a muscular and a tissue (non-muscular) fatigue component, there should be some aspect of the voice production mechanism that would be compromised differentially by each of them. However, noninvasive vocal fatigue measures, such as acoustic metrics of the voice, have not been as fruitful. For example, in a study to evaluate whether a vocal endurance test could be used to effectively evaluate vocal fatigue, Buekers (1998) found that fatigue could not be conclusively identified using self-ratings (i.e., pain and fatigue), electroglottography (EGG), standard acoustic metrics (i.e., the Multi-Dimensional Voice Program), and pitch/loudness (monitored throughout the day on a subset of subjects). In another set of studies, Laukkanen et al. (2004, 2006) measured some changes in acoustic measures (e.g., jitter, mean F0, and alpha ratio) during a vocal loading task and in teachers before and after school. However, Laukkanen et al. (2008) later studied the relation between reported vocal fatigue symptoms and acoustic variables in 79 female primary school teachers and found that neither acoustic parameters nor voice production type had real relevance to reported vocal fatigue; in addition, they observed that the acoustic parameters only seemed to reflect an increased muscle activity in response to extended vocal use.

In contrast, certain perceptual ratings appear to be more able to capture the dynamic nuances of some aspects of vocal fatigue. McCabe and Titze (2002) postulated that subjective ratings of vocal effort and vocal quality could be used to dynamically track tissue fatigue and recovery as a continuum, rather than a dichotomy of either vocal function or failure; in their pilot study of four teachers, the subjective ratings of vocal function mostly recovered in 2 hours after a vocal loading task, with residual recovery taking several days. Further, Chang and Karnell (2004) also used dynamic tracking; in their study, they focused on tracing phonation threshold pressure and speaking effort level ratings in 10 subjects. Phonation threshold pressure and speaking effort level ratings were found to increase significantly during a vocal loading task. Interestingly, phonation threshold pressure recovered almost immediately after the vocal loading activity, while speaking effort level ratings stayed statistically elevated afterwards, recovering to statistically similar baseline levels within 2 hours. Another perceptual method which may prove to be a valuable self-evaluation task is the inability to produce soft voice in which subjects attempt to produce a

specific soft, high-pitched vocal task and rate their ability to produce it, specifically designed to correlate with a symptom of laryngeal tissue fatigue, vocal fold swelling (Bastian et al., 1990). Nevertheless, this rating has not been rigorously examined.

## 2. The SAVRa Parts

The purpose of this technical memo is to describe use of the self administered subjective vocal ratings (Self-Administered Vocal Ratings *ver. a* or SAVRa) at the NCVS as of May 2008. These tests have been used by individuals wearing the NCVS Dosimeter as well as in non-dosimeter studies where the ratings were just noted on log sheets. The dosimeter prompted the wearer to conduct these rating tests every two hours.

Three perceptual ratings make up the SAVRa which include three types of rating types (i.e., emotional, functional, and physical) found in the widely used Vocal Handicap Index (Jacobson et al., 1997). The first, a rating of the current speaking effort level (EFFT, 1-10 scale; 1 for no effort, 10 for an extreme effort to speak), is based on McCabe and Titze (2002) and Chang and Karnell (2004). The second is a rating of laryngeal discomfort (DISC, 1-10 scale; 1 for no discomfort, 10 for extreme discomfort) and place of the discomfort (1-4 scale; 1-outside the larynx, 2 inside the larynx, 3 – both inside and outside, 4 - neither). The third rating is a modified soft voice functional task and ability rating (IPSV evaluated on a 1-10 scale, 1 used for an unproblematic soft voice, 10 for extreme problems with producing the soft voice).

For evaluating EFFT or the *speaking effort levels*, the subjects were asked to rate how much effort he/she currently needed to speak loudly on the scale from 1 (least effort) to 10 (maximal effort, corresponding to the thought, “I’m so tired it’s taking all my effort to speak”). Counting “one-two-three” was recommended as the testing utterance.

The *laryngeal discomfort levels* or DISC was evaluated by sensation of any feelings of achy muscles, soreness or pain inside or around the larynx, or inability to speak because of discomfort. To evaluate the place or location of laryngeal discomfort the subject was advised to say “one-two-three” and swallow in order to decide whether the discomfort was outside the larynx (usually related to achy external laryngeal muscles), inside the larynx (presumably related to fatigued internal laryngeal muscles and irritated tissues of the vocal folds and the larynx), both outside and inside, or neither inside nor outside (when no discomfort is perceived).

The task associated with the IPSV (inability to produce soft voice) consisted of four phonation tasks: 1) sustained /i/ as softly as possible on a comfortable pitch, 2) glide from low to high pitch on /i/ vowel as softly as possible, 3) soft, high-pitched repetitions of the syllables /hee-hee-hee-hee-hee/ and 4) soft and high-pitched singing of the first bars of “Happy Birthday”. The subject evaluated the ease of the soft voice production in these utterances and identified signs of degraded voice production such as phonation predictability, presence of aphonic and hoarse segments, voice breaks, unevenness of the repeated phonations, delayed voice onsets and reduced range of pitch. If no signs of degradation were present, the rating value of 1 was expected; value of 10 reflected complete inability to produce soft voice. This task was specifically designed to

predict vocal fold swelling (a potential symptom of vocal fatigue) (Bastian et al., 1990); however, the original rating has not been rigorously examined. The IPSV could also be subjected to acoustic analysis and future modifications of the tasks are currently being tested.

### 3. The SAVRa Instructions

Teachers were initially trained in the ratings at the voice laboratory and were sent home with a CD or video that reviewed the ratings. Written instructions (below) and rating forms were also sent with the teachers along with a pre-metered return envelope. Teachers were instructed to perform a two-hour reading task at home. They were asked to complete the self-ratings every fifteen minutes during the reading task, and every two hours following the task for the next two days.

*The following written instructions were provided to each teacher:*

Please listen to the instructional cd/tape before completing the self-perception ratings and/or reading task. You will be asked to rate your voice in three ways: [1] ease of phonation (how much effort it takes to produce voice); [2] your discomfort level in your throat; [3] your soft phonation (quality, evenness, predictability).

The first thing you will rate is how effortful your loud voice is to produce. You will rate your voice on a scale from 1 to 10 (1 = no effort to make sound; 10 = I'm so tired it's taking all my effort just to keep reading). Rate your effort by circling the number that matches your effort level while speaking loudly.

The second rating is for discomfort in your throat. Rate your level of discomfort on a scale from 1-10, and then swallow and take note of where you feel discomfort, inside or outside your throat. Rate your effort by circling the number that matches your discomfort level while speaking loudly. Record where you feel that discomfort, inside or outside your throat, or both.

The third rating is for soft phonation. The following tasks will be used to make these assessments: Rate your soft voice production by circling the number that matches your soft phonation after performing the four tasks.

1. Sustain the vowel /i/ (ee) for 5 seconds as softly as possible on a comfortable pitch.
2. Glide on the vowel /i/ (ee) from low to high pitch as softly as possible.
3. Say ee-ee-ee-ee-ee-ee-ee (staccato—very short and high-pitched).
4. Sing a few bars of “Happy Birthday” extremely soft and high-pitched.

Here are some “undesirables” to keep in mind as you perform the above tasks and rate your soft phonation.

- Roughness/breathiness
- Times when your voice doesn't come out at all
- Only loud voice comes out
- It takes a while for your voice to activate

- Your sound cuts out as you use it

When the subjects used the Dosimeter several months later, subjects were retrained to do the SAVRa every two hours using the dosimeter as the input device, with instructions given by the Dosimetry as a reminder. Implementation of the SAVRa in the Dosimeter is shown in Figure 2.1.

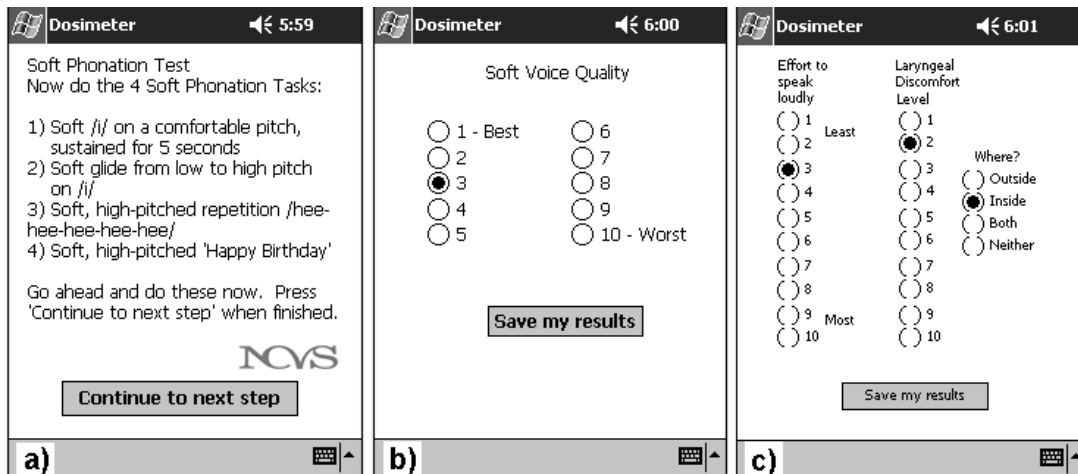


Figure 2.1 consists of three screenshots of the NCVS Vocal Dosimeter software interface, labeled a), b), and c).

**Screen a) (5:59):** Displays the 'Soft Phonation Test' instructions. It lists four tasks: 1) Soft /i/ on a comfortable pitch, sustained for 5 seconds; 2) Soft glide from low to high pitch on /i/; 3) Soft, high-pitched repetition /hee-hee-hee-hee-hee/; 4) Soft, high-pitched 'Happy Birthday'. It includes a 'Continue to next step' button and the NCVS logo.

**Screen b) (6:00):** Displays the 'Soft Voice Quality' rating scale. It features two columns of radio buttons numbered 1 to 10, with '1 - Best' on the left and '10 - Worst' on the right. Radio button 3 is selected. A 'Save my results' button is at the bottom.

**Screen c) (6:01):** Displays two rating scales. The left scale is 'Effort to speak loudly' with radio buttons 1 to 10, where 'Least' is at the top and 'Most' is at the bottom. Radio button 3 is selected. The right scale is 'Laryngeal Discomfort Level' with radio buttons 1 to 10. Below it is a 'Where?' section with radio buttons for 'Outside', 'Inside', 'Both', and 'Neither'. Radio button 'Inside' is selected. A 'Save my results' button is at the bottom.

Figure 2.1 Collecting the IPSV on the NCVS Vocal Dosimeter.

## 4. The Statistics of the SAVRa over 2 weeks

In order to use parametric statistics on these ratings, statistics on their distributions were needed. The ratings were normalized for each subject by taking a subject's median rating and adjusting that to zero. Then the overall distribution of all subjects was presented. This distribution was periodically updated as additional subjects were recorded.

The three ratings were used to track normal voice events in teachers over two weeks, the distribution of the ratings (52 teachers, 13,857 rating events). Table 3.1 lists the statistics while Figure 3.1 illustrates the distributions of the three parts of the SAVRa and also a combined distribution of all ratings (ALL).

Table 3.1. Statistics of the SAVRa.

	EFFT	IPSV	DISC	ALL
N	4619	4619	4619	13865
$\mu$ (mean)	2.4894	3.8357	2.4894	2.9387
Variance	2.3758	4.0234	2.5616	3.3906
$\sigma$ (std-dev)	1.5414	2.0058	1.6005	1.8413
Median	2	3	2	2
1 <sup>st</sup> Quartile	2	2	1	2
3 <sup>rd</sup> Quartile	3	4	3	3
Skewness	1.458	0.8065	1.4782	1.2087
Kurtosis	5.6221	3.1859	5.3732	4.281

Hunter & Titze: Dynamics of a three-dimensional model of vocal fold abduction/adduction

While the data in Figure 3.1 is not normal, it is possible that each individual teacher's response was normal. Thus, every teacher's individual responses in the raw data above was adjusted by their individual median value (basically moving the median to 0); the new statistics are shown in Table 3.2. The results were highly peaked (elevated kurtosis) with reduced skewing (nearly normal) for EFFT and DISC (Figure 3.2).

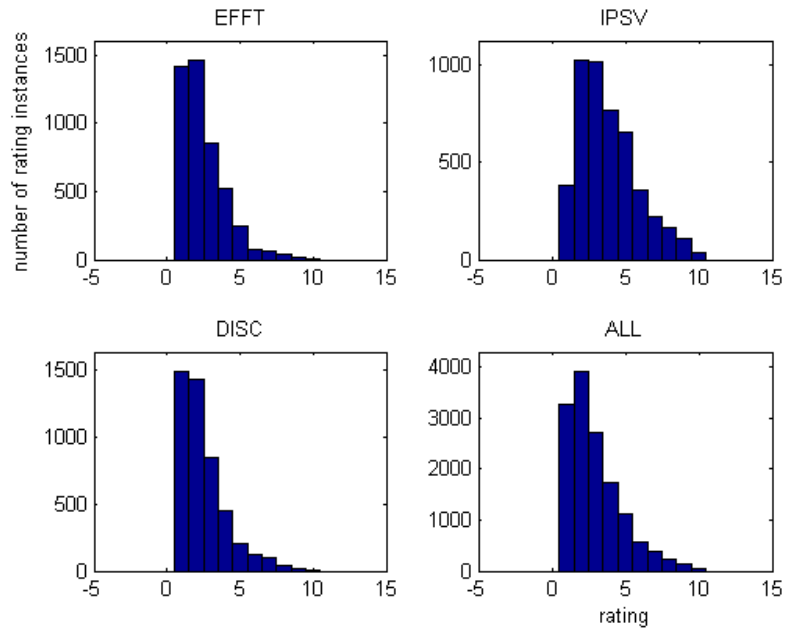


Figure 3.1. Distributions of perceptual ratings over two weeks during normal phonation.

Table 3.2. Statistics of individual teacher responses where the median was adjusted to 0 and then combined for all teachers.

	EFFT	DISC	IPSV	ALL
N	4619	4619	4619	13865
$\mu$ (mean)	0.2048	0.2221	0.113	0.1799
Variance	1.2172	1.366	2.6178	1.7365
$\sigma$ (std-dev)	1.1033	1.1688	1.6179	1.3177
Median	0	0	0	0
1 <sup>st</sup> Quartile	0	0	-0.5	0
3 <sup>rd</sup> Quartile	0	0	1	0
Skewness	1.1641	0.9765	0.0328	0.4366
Kurtosis	9.5879	7.8227	5.7023	7.5589

To see if there was learning or an increased awareness by the subjects over the two weeks, the overall change in the SAVRa ratings was tracked over the participation time. Since participation was generally two weeks, changes due to regular daily or weekly voicing events should be repetitive and have no effect on the general trend (slope) of the ratings. Slope (rating change over a day) for the two weeks was calculated for each subject. From these

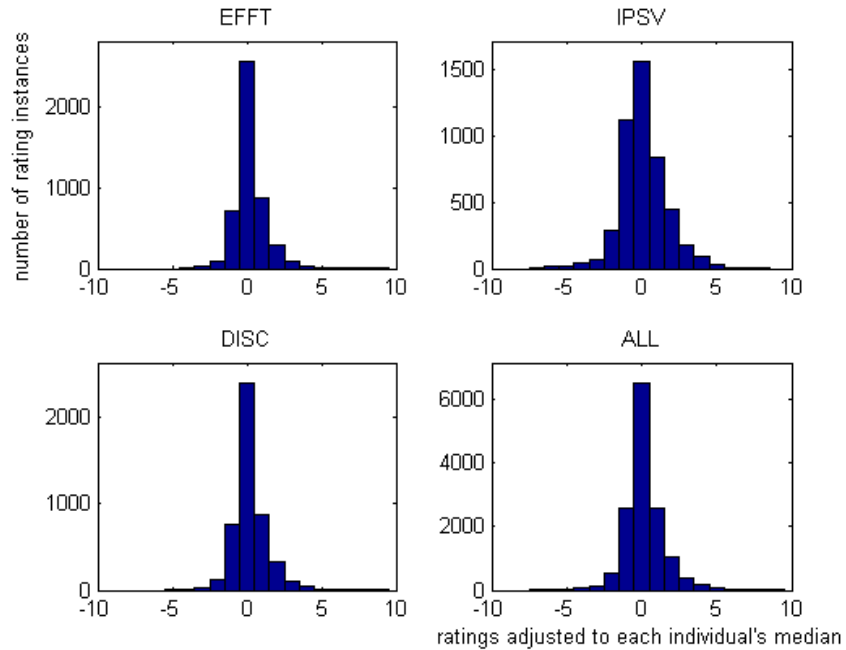


Figure 3.2. Distributions of perceptual ratings over two weeks during normal phonation.

average slopes the all subject slope statistics for the approximately 2 week participation time was calculated (Table 3.3). The median across the subjects for the three ratings were 0.014, 0.002, and 0.021 change per day respectively; meaning that it would take approximately 71 days, 500 days and 48 days respectively for an average change of 1 point to take place.

Table 3.3. Statistics of slope (change per day) of ratings for all subject's ratings over their approximately 2-week participation time.

	EFFT	IPSV	DISC	Days Participating
N	52	52	52	52
$\mu$ (mean)	0.042	0.0419	0.0426	502
Variance	0.0171	0.0279	0.0184	7.1672
$\sigma$ (std-dev)	0.1308	0.1671	0.1358	2.6772
Median	0.0142	0.002	0.0209	15
1 <sup>st</sup> Quartile	-0.0251	-0.0646	-0.0141	15
3 <sup>rd</sup> Quartile	0.0772	0.0793	0.0903	15
Mode	-0.3302	-0.28	-0.354	15
Skewness	1.1194	1.3333	0.5923	0.0912
Kurtosis	6.7039	5.1952	5.3469	7.1453

## Hunter &amp; Titze: Dynamics of a three-dimensional model of vocal fold abduction/adduction

It is possible that effects from learning or awareness may be larger the first week and more stable by week two. Therefore, Table 3.4 depicts the same statistics from the first half of participation time, while Table 3.5 shows the statistics from the second half of the participation time. The median slope values for Effort in Tables 3.4-3.5 are 0.0177 and 0.0096 change in rating per day. There is a smaller slope from the second half of the participation time than the first half (0.0096/day vs 0.0177/day) so there may be a small learning or awareness effect. However, at the rate of change seen here, it would take as few as 56 days (for the first half) or as many as 100 days (for the second half) for a rating to change by a single point. Therefore, any learning or awareness adjustments are small for EFFT. IPSV was slightly bigger possibly indicating a learning effect in the first week (20 days until change).

Table 3.4. Statistics of slope (change per day) of ratings for all subject's ratings over the first half of participation time.

	EFFT	DISC	IPSV	Days Participating
N	52	52	52	52
$\mu$ (mean)	0.0447	0.0573	-0.0514	4
Variance	0.0422	0.0507	0.1687	5.6283
$\sigma$ (std-dev)	0.2054	0.2251	0.4108	2.3724
Median	0.0177	0.0516	-0.0527	8
1 <sup>st</sup> Quartile	-0.0965	-0.0617	-0.1918	7
3 <sup>rd</sup> Quartile	0.2146	0.1628	0.1311	8
Mode	-0.3891	-0.5785	-2.4696	8
Skewness	0.3034	0.1041	-3.6867	1.7438
Kurtosis	2.7933	4.6506	23.4715	5.8836

Table 3.5. Statistics of slope (change per day) of ratings for all subject's ratings over the second half of participation time.

	EFFT	DISC	IPSV	Days Participating
N	52	52	52	52
$\mu$ (mean)	0.0539	0.067	0.0388	7.0192
Variance	0.0423	0.0457	0.0789	0.9419
$\sigma$ (std-dev)	0.2057	0.2138	0.2809	0.9705
Median	0.0096	9.89E-04	0.0391	7
1 <sup>st</sup> Quartile	-0.032	-0.0277	-0.0912	7
3 <sup>rd</sup> Quartile	0.109	0.1293	0.1666	7
Mode	-0.4336	-0.474	-0.6076	7
Skewness	0.9592	0.9089	0.66	-1.9317
Kurtosis	5.8521	4.6522	5.5909	8.2354

## Literature

- R. W. Bastian, A. Keidar, and Katherine Verdolini-Marston. Simple vocal tasks for detecting vocal fold swelling. *J.Voice* 4 (2):172-183, 1990.
- R. Buekers. Are voice endurance tests able to assess vocal fatigue? *Clin.Otolaryngol.Allied Sci.* 23 (6):533-538, 1998.
- A. Chang and M. P. Karnell. Perceived phonatory effort and phonation threshold pressure across a prolonged voice loading task: a study of vocal fatigue. *J.Voice* 18 (4):454-466, 2004.
- L. Czerwonka, J. J. Jiang, and C. Tao. Vocal nodules and edema may be due to vibration-induced rises in capillary pressure. *Laryngoscope* 118 (4):748-752, 2008.
- S. Gray and I. Titze. Histologic Investigation of Hyperphonated Canine Vocal Cords. *Ann.Otol.Rhinol.Laryngol.* 97 (4):381-388, 1988.
- E. J. Hunter and I. R. Titze. Refinements in modeling the passive properties of laryngeal soft tissue. *J.Appl.Physiol* 103 (1):206-219, 2007.
- A. M. Laukkanen, K. Jarvinen, M. Artkoski, T. Waaramaa-Maki-Kulmala, E. Kankare, S. Sippola, T. Syrja, and A. Salo. Changes in voice and subjective sensations during a 45-min vocal loading test in female subjects with vocal training. *Folia Phoniatr.Logop.* 56 (6):335-346, 2004.
- A. M. Laukkanen and E. Kankare. Vocal loading-related changes in male teachers' voices investigated before and after a working day. *Folia Phoniatr.Logop.* 58 (4):229-239, 2006.
- A. M. Laukkanen, I. Ilomaki, K. Leppanen, and E. Vilkman. Acoustic Measures and Self-reports of Vocal Fatigue by Female Teachers. *J.Voice* 22 (3):283-289, 2008.
- D. J. McCabe and I. R. Titze. Chant therapy for treating vocal fatigue among public school teachers: a preliminary study. *American Journal of Speech-Language Pathology* 11:356-369, 2002.
- I. R. Titze. Toward occupational safety criteria for vocalization. *Log.Phon.Vocol.* 24:49-54, 1999.
- I. R. Titze, E. J. Hunter, and J. G. Svec. Voicing and silence periods in daily and weekly vocalizations of teachers. *J.Acoust.Soc.Am.* 121 (1):469-478, 2007.
- N. V. Welham and M. A. Maclagan. Vocal fatigue: current knowledge and future directions. *J.Voice* 17 (1):21-30, 2003.
- J. W. Yates, J. T. Kearney, M. P. Noland, and W. M. Felts. Recovery of dynamic muscular endurance. *Eur.J.Appl.Physiol Occup.Physiol* 56 (6):662-667, 1987.
- Y. Zhang, L. Czerwonka, C. Tao, and J. J. Jiang. A biphasic theory for the viscoelastic behaviors of vocal fold lamina propria in stress relaxation. *J.Acoust.Soc.Am.* 123 (3):1627-1636, 2008.

## Acknowledgements

This work was supported by NIDCD grant No DC004224 from National Institutes of Health. Thanks to research team at the NCVS Denver site particularly Andrew Starr, Albert Worley, Jennifer Spielman, Angela Halpern, and Peter Popolo.

## Copyright

The images and text enclosed in this memo and accompanying this memo are open to use by the public as a service of the National Center for Voice and Speech. However, we ask the reader to respect the time and effort put into this manuscript and research. If the text or images are used, the user agrees to reference to this document and the National Center for Voice and Speech. We reserve the right of refusal to (1) be authors on papers using the information and (2) have the supporting project acknowledged. The user agrees to freely share with the NCVS any extension research based on the research contained.

## Revisions

- 1.0 Eric Hunter: Main document (July. 2008)
- 1.1 Eric Hunter: Minor typographical changes (August, 2008)