Original Paper

Folia Phoniatrica et Logopaedica

Folia Phoniatr Logop 2009;61:227–238
DOI: 10.1159/000228000

Published online: July 9, 2009

A Comparison of the Effects of Voice Massage™ and Voice Hygiene Lecture on Self-Reported Vocal Well-Being and Acoustic and Perceptual Speech Parameters in Female Teachers

Kirsti Leppänen[a] Anne-Maria Laukkanen[a] Irma Ilomäki[a] Erkki Vilkman[b]

[a]Department of Speech Communication and Voice Research, University of Tampere, Tampere, and
[b]Phoniatric Clinic, Helsinki University Hospital, Helsinki, Finland

Key Words

Voice professionals · Voice Massage™ · Voice hygiene · Vocal fatigue · Voice quality · Acoustic analysis

Abstract

This study compared the effects of Voice Massage™ (VM) and a voice hygiene lecture (VHL) on 60 female teachers. VM is a Finnish massage method which treats muscles related to voice production. All subjects attended the VHL (3 h). Half of them were randomly assigned to the VM group, the other half received only VHL. VM was given 5 times in 1-hour sessions at intervals of 1–2 weeks. At the beginning and end of the autumn school term, before and after a working day, (1) a 1-min reading sample was recorded at both the subject's habitual loudness and loudly, (2) a prolonged phonation on [a:] was recorded at habitual speaking pitch and (3) a questionnaire on voice quality, ease or difficulty of phonation, and tiredness of the throat was completed. The reading samples were analysed for the fundamental frequency (F_0), equivalent sound level (Leq) and α-ratio [Leq (1–5 kHz) – Leq (50–1,000 Hz)]. The vowel samples were analysed for the F_0 and Leq, as well as for jitter and shimmer. The reading samples were also evaluated perceptually by 3 speech trainers. The mean F_0 and difficulty of phonation increased from the beginning to the end of the term in the VHL group ($p = 0.026$, $p = 0.007$, respectively). In the VM group, the perceived firmness of loud reading decreased ($p = 0.026$). The results suggest that VM may help in sustaining vocal well-being during a school term.

Introduction

Voice problems have been reported to occur more frequently in teachers than in other occupational voice users [1–7]. According to Vilkman [8], a voice problem should be seen as an occupational disease rather than a private health problem. Although psychosocial factors may also contribute to the development of voice problems, the main focus here is on the physical loading related to voice production in the teaching profession.

The teaching profession requires prolonged voice use and often the use of a loud voice in order to be heard in the classroom, where the acoustic conditions are frequently less than ideal (i.e. a great deal of background noise, too long a reverberation time, etc.). A loud voice typically implies a higher sound pressure level (SPL), a higher fundamental frequency (F_0) and a more hyperfunctional voice production type (more adduction be-

Some results from this study were reported as a poster in Pevoc 7 (Pan European Voice Conference) in Groningen 2007.

KARGER

Fax +41 61 306 12 34
E-Mail karger@karger.ch
www.karger.com

1021–7762/09/0614–0227$26.00/0

Accessible online at:
www.karger.com/fpl

Kirsti Leppänen, MA
Department of Speech Communication and Voice Research
University of Tampere
FI–33014 Tampere (Finland)
Tel. +358 50 583 1965, E-Mail Kirsti.Leppanen@uta.fi

tween the vocal folds) resulting in less tilting of the spectrum and, thus, a more efficient acoustic message transfer. On the other hand, a higher F_0, SPL and tighter adduction imply more biomechanical loading on the vocal fold tissue [9]. Thus, teachers' and other voice professionals' voice problems may be seen as resulting from an overload of the vocal organ. Consequently, the prevention and cure of these problems should imply a reduction of vocal loading.

Different approaches have been used to reduce vocal loading. Voice hygiene education aims to improve vocal ergonomics by providing information on factors that increase vocal loading. These factors, in turn, are related to the environment, type of work and to the worker him/herself. After voice hygiene education, subjects have reported that they found the strategies taught to be helpful [10] and that the education increased the awareness of the warning symptoms of vocal fatigue [10, 11], resulting in a change of behaviour in terms of avoidance of vocal overuse and ensuring adequate vocal rest and hydration [11].

Massage methods have been used to reduce excessive tension in the laryngeal muscles and to improve awareness of the state of the muscles [12–14]. Hence, they may assist voice training and therapy processes [12, 13, 15]. There is a long tradition of using massage in phoniatric treatment in China [16]. Aronson [13] proposed using manual circumlaryngeal therapy, which aims at reducing tension in the muscles that raise the larynx. Voice Massage™ (VM) is a Finnish method (registered as a trademark by the National Board of Patents and Registration of Finland) developed by Leena Koskinen in collaboration with medical and voice specialists. It consists of the holistic manipulation of muscles related to voice production. Positive changes after massage treatment have been reported, e.g. a decrease in (excessive) perturbation (jitter and shimmer) values, an increased signal-to-noise ratio and better perceived voice quality [17, 18]. Ternström et al. [14] evaluated the immediate effects of a 30-min naprapathic treatment of the back, neck and face muscles for normal subjects with no voice complaints. According to their results, the F_0 and SPL decreased after treatment, but did not change in the control group which was only lying down for 30 min. This suggests that massage treatment has a relaxing effect which is not only related to resting in the supine position. After VM treatment, various positive subjective sensations have been reported [18].

Several studies have investigated the acoustic and self-reported consequences of prolonged loud voice use either in a voice professional's working day or through a vocal loading test carried out in a laboratory environment. An increase in the F_0 and SPL as well as a decrease in spectral tilt have been reported. Furthermore, subjects with a higher mean F_0 and SPL during vocal loading tasks have reported more symptoms of vocal fatigue [6, 19]. On the other hand, Rantala and Vilkman [20] found a greater working-day-related increase in the F_0 and SPL in teachers who reported fewer voice complaints. Additionally, according to Jónsdóttir [21], teachers reported less tiredness in the vocal mechanism after a working day when they used an electric amplifier in the classroom. However, a greater increase in the F_0 and SPL as well as a greater decrease in the spectral tilt of the long-term average spectrum was observed. A loading-related increase in various vocal parameters may thus be seen as a normal adaptation to vocal loading. The summarized findings suggest that positive changes in vocal behaviour do not only imply subjective reports of diminished symptoms of vocal fatigue, but that they could also be related to a somewhat lower mean F_0, SPL and spectral tilt. The findings also suggest a greater increase in these parameters during loading, e.g. a voice professional's working day. Since vocal fatigue typically implies a deterioration of voice quality [22], positive changes, in turn, are expected to be reflected in improved perceptual voice quality, e.g. in terms of decreased hoarseness, breathiness or strain [23].

The present study compares the effects of VM and a voice hygiene lecture (VHL) on acoustic and perceptual speech parameters as well as on voice-related sensations in teachers. The aim is to find treatments which improve the vocal well-being of teachers during school terms. The positive effects of VHL and VM are expected to be seen in (1) a lower mean F_0, equivalent sound level (Leq, i.e. an energy mean of the sound level averaged over the measurement period) and α-ratio, which reflects spectral energy distribution and thus phonation type, (2) an increase in these acoustic parameters during a working day, (3) fewer symptoms of vocal fatigue after loading and (4) a better perceived voice quality (not strained, breathy or hoarse; pitch is neither too high nor low) both before and after loading. The positive effects of both VHL and VM are expected to be greater than with VHL alone.

Subjects and Methods

Subjects

A total of 90 Finnish female primary school teachers in the Tampere region of Finland volunteered to be subjects by completing a questionnaire on the internet. The present study concerns

60 of them. The remaining 30 subjects were given voice training and the results have been reported elsewhere [23]. Some results for some of the same subjects, but with a different focus, have also been reported [23–25]. The questionnaire collected background variables like age, health (allergies, reflux etc.), vocal symptoms after voice use, teaching experience, group size in the classroom, hours taught per week and a self-evaluation of the person's voice (quality, endurance and audibility). They also filled in the Finnish translation [26] of the Voice Activity and Participation Profile [27].

The mean age of the subjects was 40.6 years (SD 7.8, range 26–57 years) and the average amount of teaching experience was 14.7 years (SD 7.9 years). The average amount of teaching per week was 24.2 h (SD 4.4 h) and the average size of the group taught was 18.5 pupils (SD 7.7). All subjects were functionally healthy. They regarded themselves as capable of the vocally demanding profession. The mean Voice Activity and Participation Profile total score was 40.9 (SD 35), while the results obtained by Ma and Yiu [27] were 114.12 (SD 55.28) for dysphonic subjects and 13.4 (SD 18.90) for the (healthy) controls.

The subjects were randomly assigned to 2 groups of 30 subjects. Both groups attended a VHL, and one of them also received VM. The group which only attended the VHL is called the VHL group and the other one is called the VM group. Table 1 presents data on the background variables for both groups. Randomization seems to have been successful and the groups can be regarded as comparable.

Table 1. Data of VHL group subjects (n = 29) and VM group subjects (n = 30)

	VHL	VM	p
Age, years	41.5 ± 7.6	39.6 ± 8.2	0.374
Years of teaching experience	16.2 ± 7.4	13.1 ± 8.6	0.157
Hours taught per week	23.3 ± 5.5	25.1 ± 2.7	0.528
Group size in the classroom	19.3 ± 7.6	17.7 ± 7.8	0.293
VAPP total score	38.6 ± 43	44.9 ± 35	0.43
Min.-max.	0 – 202	5 – 125	
Vocal symptoms total score	22.5 ± 13.2	23.9 ± 14.4	0.712
Min.-max.	2 – 59	0 – 53	
Self-evaluation total score			
(quality, endurance, audibility)	5.7 ± 1.3	5.9 ± 1.6	0.761
Min.-max.	3 – 9	2 – 9	

Significance of differences expressed as p values. Mann-Whitney U tests are used to determine the p values. VAPP = Voice activity and participation profile.

Interventions

All subjects received a 3-hour VHL. It covered the main factors that cause vocal loading in teachers, methods to avoid it and the basics of economic voice use. The lecturer holds a PhD and is a very experienced teacher of speech technique and vocology. VM is a special massage therapy performed by a trained VM therapist. It consists of manipulation of respiratory, laryngeal and articulatory muscles. In general, vocalization is not included in VM treatment. Classical massage grips (strokes, kneading, friction and intercostal pull) are used. Both the agonistic and antagonistic muscles are treated. While the terms 'agonistic' and 'antagonistic' do not reflect the 3-dimensional nature of movements as such, treatment of muscles with opposing effect aims at affecting the control of movements in the most holistic way possible. The goal is to increase the mobility of the ribcage when breathing and to avoid excessive tension in the various muscles used in voice production. Adhering to the traditionally used procedure, VM was given 5 times (1 h each time). The first 3 sessions were given at intervals of 1 week, while the last 2 sessions were given at intervals of 1 month. The treatment was given by a VM therapist with over 10 years of experience.

Recordings

At the beginning and end of the autumn term, before and after a working day, the teachers recorded a reading sample for 1 min at both habitual loudness and loudly, and a prolonged phonation on [a:] for 5 s. The recordings were made in a classroom before (at about 7.30 a.m.) and after (at about 4.00 p.m.) a vocally loading working day. Recordings were made using a portable digital recorder (Sony TCD-D8) and a microphone (AKG B29L) attached to a headset. The microphone was placed at a distance of 6 cm from the corner of the subject's mouth. The recordings were calibrated for Leq measurements using a sound generator (BOSS TU-120) and a sound level meter (Brüel and Kjær, 2206).

Acoustic Analyses

The reading and vowel samples were analysed for the average F_0, Leq and α-ratio [Leq (1–5 kHz) – Leq (50–1,000 Hz)], which describes sound energy distribution along frequencies and, thus, reflects voice quality. Additionally, the vowel samples on [a:] of about 5 s in duration were analysed for the mean and SD of jitter (in percent) and shimmer (in decibels). The acoustic analyses were made using the commercial program Intelligent Speech Analyser™ (ISA), a signal analysis system developed by Raimo Toivonen [28]. A blinded method was used to analyse the samples recorded before and after the treatments.

Subjective Sensations

After every recording, the subjects completed a questionnaire about voice quality, ease or difficulty of phonation, and tiredness of the throat. The visual analogue scale (VAS) was used. Voice quality and ease or difficulty of phonation were reported on a 200-mm VAS in order to also allow registration of possible warm-up effects (0 = very good/very easy, 100 mm = ordinary, 200 mm = very poor/very difficult). Tiredness of the throat was reported on a unipolar scale (0 = no tiredness at all, 100 mm = very tired). At the end of the autumn term, the teachers registered the degree of positive influence from the interventions by using a 100-mm VAS (0 = no influence at all, 100 mm = a lot of positive influence).

The type of influence was reported by choosing 1–3 of the alternatives given (voice quality, audibility and endurance) and by writing a free comment. The subjective results as an outcome measure have been considered in more detail in a previous article by Laukkanen et al. [25].

Perceptual Evaluation

Reading samples recorded at habitual loudness and loudly were perceptually evaluated by 3 speech trainers using the Judge

Table 2. Mean values ± SD for the average F_0, Leq and α-ratio [(Leq 1–5 kHz) – (Leq 50–1,000 Hz)] in reading samples at habitual loudness and the mean F_0, Leq, and jitter and shimmer (F_0 and amplitude perturbation) in vowel samples[1]

	Beginning of term			End of term			Beginning versus end	
	morning	afternoon	difference	morning	afternoon	difference	difference in mean values	difference in changes
Reading								
VHL group	n = 30	n = 30		n = 29	n = 29			
F_0, Hz	187.8 ± 15.8	193.9 ± 16.3	0.002	193.6 ± 19.1	197.2 ± 17.1	n.s.	0.026	n.s.
Leq, dB	76.4 ± 3.0	77.8 ± 3.5	0.028	76.6 ± 3.1	78.0 ± 2.6	0.006	n.s.	n.s.
α-Ratio, dB	–15.1 ± 2.8	–13.3 ± 3.1	0.000	–15.0 ± 2.9	–12.7 ± 2.6	0.000	n.s.	n.s.
VM group	n = 30	n = 30		n = 30	n = 30			
F_0, Hz	191.3 ± 13.4	195.7 ± 14.8	0.001	195.0 ± 16.8	200.6 ± 14.2	0.014	n.s.	n.s.
Leq, dB	75.5 ± 3.88	75.9 ± 5.04	n.s.	76.6 ± 3.5	76.5 ± 3.3	n.s.	n.s.	n.s.
α-Ratio, dB	–14.9 ± 2.9	–13.1 ± 2.9	0.000	–14.6 ± 2.9	–13.2 ± 3.1	0.001	n.s.	n.s.
[a:]								
VHL group	n = 28	n = 28		n = 28	n = 28			
F_0, Hz	191.1 ± 24.2	202.1 ± 31.3	0.031	193.1 ± 24.3	197.6 ± 28.8	n.s.	n.s.	n.s.
Leq, dB	81.9 ± 6.4	84.0 ± 5.2	0.045	83.0 ± 4.6	85.7 ± 5.4	0.010	n.s.	n.s.
SD of jitter, %	0.101 ± 0.088	0.067 ± 0.056	n.s.	0.074 ± 0.057	0.071 ± 0.102	n.s.	n.s.	n.s.
Mean jitter, %	0.414 ± 0.383	0.318 ± 0.282	n.s.	0.334 ± 0.280	0.287 ± 0.416	n.s.	n.s.	n.s.
Shimmer, dB	0.616 ± 0.320	0.456 ± 0.202	0.028	0.424 ± 0.186	0.369 ± 0.156	n.s.	n.s.	n.s.
VM group	n = 30	n = 30		n = 28	n = 28			
F_0, Hz	198.2 ± 31.1	203.7 ± 34.2	n.s.	197.5 ± 31.6	201.9 ± 29.7	n.s.	n.s.	n.s.
Leq, dB	82.9 ± 7.3	83.7 ± 6.4	n.s.	85.0 ± 6.5	84.8 ± 5.9	n.s.	n.s.	n.s.
SD of jitter, %	0.085 ± 0.115	0.046 ± 0.021	n.s.	0.079 ± 0.114	0.084 ± 0.150	n.s.	n.s.	n.s.
Mean jitter, %	0.346 ± 0.457	0.187 ± 0.090	0.049	0.321 ± 0.467	0.243 ± 0.215	n.s.	n.s.	n.s.
Shimmer, dB	0.486 ± 0.252	0.446 ± 0.353	n.s.	0.468 ± 0.271	0.429 ± 0.304	n.s.	n.s.	n.s.

All differences are expressed as p values. Significance of differences within groups determined by Wilcoxon signed-rank test; n.s. = non-significant, $p > 0.05$.

[1] Recorded before and after a teacher's working day at the beginning and end of the autumn term.

Program (developed by Svante Granqvist, Royal Institute of Technology) [29]. A VAS from 0 to 1,000 units was used to evaluate voice strain (0 = not at all, 1,000 = very strained), breathiness (0 = not at all, 1,000 = very breathy) and hoarseness (0 = not at all, 1,000 = very hoarse) in reading at habitual loudness, and firmness of phonation (0 = hypofunctional, 500 = adequate firmness, 1,000 = very hyperfunctional/strained) and pitch (0 = too low, 500 = suitable, 1,000 = too high for the speaker) in loud reading. The samples were presented in a randomized order through headphones (Sennheiser HD 530 II). The trainers could listen to each sample as many times as they wished in order to make their judgements.

Statistical Analyses

The significance of differences in the acoustic, subjective and perceptual parameters before and after a working day and between the beginning and end of the term were tested using a Wilcoxon signed-ranks test. Differences between the VHL and VM groups at the beginning of the term were studied with unpaired Student's t tests and Mann-Whitney U tests. Non-parametric tests were used for the parameters with a skewed distribution and when the number of cases studied was smaller (i.e. when samples before and after a working day were compared to each other within the groups). Mean and SD of jitter, mean shimmer, breathiness and hoarseness showed a skewed distribution in this material (skewing 3.8, 4.1, 2.2, 1.02 and 1.1, respectively). Analysis of variance (ANOVA) was used to study the effects of time and group on the parameters (repeated measures ANOVA) and to compare the groups at the end of term (general linear model univariate analysis) for the mean values obtained at the end of term [(values before working day + values after working day)/2]. Values obtained at the beginning of the term were set as covariates in the univariate analysis. Cronbach's α was used to calculate the inter-rater reliability of the perceptual evaluation. Pearson's and Spearman's correlations were used to investigate the relations between variables and for calculating intra-rater reliability in perceptual analysis. (Pearson's correlation was used for parameters with a normal distribution and for the perceptual ratings given on a continuous scale.) All computations were performed using the Statistical Program for Social Sciences (SPSS 15 Software; SPSS Inc., Chicago, Ill., USA).

Table 3. Mean values ± SD for the average F_0, Leq and α-ratio [(Leq 1–5 kHz) – (Leq 50–1,000 Hz)] in loud reading[1]

Load reading	Beginning of term			End of term			Beginning versus end	
	morning	afternoon	difference	morning	afternoon	difference	differences in mean values	differences in changes
VHL group	n = 28	n = 28		n = 28	n = 28			
F_0, Hz	207.9 ± 20.28	210.4 ± 19.13	n.s.	210.4 ± 20.74	215.9 ± 17.96	0.016	n.s.	n.s.
Leq, dB	83.82 ± 3.49	84.20 ± 4.33	n.s.	83.36 ± 3.77	83.88 ± 2.73	n.s.	n.s.	n.s.
α-Ratio, dB	–9.28 ± 3.32	–8.79 ± 3.68	n.s.	–9.26 ± 3.28	–8.21 ± 2.99	0.016	n.s.	n.s.
VM group	n = 30	n = 30		n = 30	n = 30			
F_0, Hz	208.7 ± 17.5	214.4 ± 19.09	0.000	213.5 ± 17.97	218.3 ± 16.57	0.014	n.s.	n.s.
Leq, dB	82.6 ± 4.82	82.9 ± 5.42	n.s.	83.01 ± 3.64	82.44 ± 4.19	n.s.	n.s.	n.s.
α-Ratio, dB	–9.32 ± 3.02	–8.09 ± 2.81	0.000	–9.14 ± 2.78	–8.41 ± 2.83	0.005	n.s.	n.s.

All differences are expressed as p values; n.s. = non-significant, p > 0.05. Significance of differences within groups determined by Wilcoxon signed-rank test.

[1] Recorded before and after a teacher's working day at the beginning and end of the autumn term.

Table 4. Differences between VHL and VM groups at the beginning of term

Habitual reading			Loud reading		
mean F_0	mean SPL	mean α-ratio	mean F_0	mean SPL	mean α-ratio
p = 0.362 t = –0.919	p = 0.292 t = 1.063	p = 0.837 t = –0.206	p = 0.884 t = –0.146	p = 0.285 t = 1.079	p = 0.964 t = 0.046
Sustained [a:]			**Self-evaluation**		
mean jitter	SD of jitter	mean shimmer	production	quality	throat
p = 0.059 Z = –1.885	p = 0.115 Z = –1.575	p = 0.073 Z = –1.790	p = 0.756 t = –0.312	p = 0.933 t = 0.084	p = 0.502 t = –0.676
Perceptual evaluation (habitual reading)			**Perceptual evaluation (loud reading)**		
strain	breathiness	hoarseness	firmness	pitch	
p = 0.600 t = 0.527	p = 0.585 Z = –0.546	p = 0.756 Z = –0.311	p = 0.375 t = –0.894	p = 0.063 t = –1.895	

Parameters have been compared with Student's unpaired t test for parameters with a normal distribution (degree of freedom 58, t values shown) and with Mann-Whitney U test (Z values shown) for parameters with a skewed distribution (i.e. skewing > ±1). Significance of differences indicated with p values. All differences were non-significant (p > 0.05).

Results

Acoustic Parameters

Tables 2 and 3 show the mean values and SD for the acoustic parameters in reading samples at habitual loudness and loudly and in the prolonged vowel phonation. The groups did not differ from each other in terms of acoustic variables at the beginning of the term (table 4). Various working-day-related changes were registered in the acoustic parameters for both groups (tables 2 and 3). At the end of the working day, the F_0, Leq and α-ratio tended to be higher in the reading samples read with both

Table 5. Repeated measures ANOVA analyses of the tests measuring the effects between subjects

		Mean squared	F	p
Habitual reading				
F_0	intercept	8,768,024.416	11,235.563	<0.001
	group effect	298.414	0.382	0.539
SPL	intercept	1,359,670.708	53,312.015	<0.001
	group effect	48.735	1.911	0.172
α-Ratio	intercept	45,744.591	1,838.627	<0.001
	group effect	1.979	0.08	0.779
Loud reading				
F_0	intercept	10,047,398.27	9,177.123	<0.001
	group effect	433.389	0.396	0.532
SPL	intercept	1,542,248.120	36,806.046	<0.001
	group effect	41.659	0.994	0.323
α-Ratio	intercept	17,592.833	599.925	<0.001
	group effect	4.554	0.155	0.695
Sustained [a:]				
Mean jitter	intercept	21.748	85.188	<0.001
	group effect	0.236	0.924	0.341
SD of jitter	intercept	1.335	84.121	<0.001
	group effect	0.001	0.088	0.768
Shimmer	intercept	49.432	396.445	<0.001
	group effect	0.005	0.037	0.849
Self-evaluation				
Production	intercept	1,909,236.817	481.125	<0.001
	group effect	13.067	0.003	0.954
Quality	intercept	2,261,070.938	674.316	<0.001
	group effect	203.504	0.061	0.806
Throat	intercept	399,024.150	401.898	<0.001
	group effect	48.600	0.049	0.826
Perceptual evaluation				
Strain	intercept	10,527,859.80	140.220	<0.001
	group effect	37,364.965	0.498	0.483
Breathiness	intercept	11,604,295.43	127.629	<0.001
	group effect	16,762.231	0.184	0.669
Hoarseness	intercept	9,862,108.940	101.793	<0.001
	group effect	2,367.092	0.024	0.876
Evaluation of loud reading				
Firmness	intercept	69,143,864.53	1,352.579	<0.001
	group effect	40,758.769	0.797	0.376
Pitch	intercept	72,277,931.32	2,029.304	<0.001
	group effect	168,363.304	4.727	0.034

Significance of differences shown with p values. Intercept = Effect of time; group effect = differences between VHL and VM groups.

Table 6. Differences between VHL and VM groups at the end of term for mean parameter values studied with univariate analyses (dependent variables)

		Mean squared	F	p
Habitual reading				
F_0	parameter	6,042.668	39.193	<0.001
	group effect	1.117	0.007	0.932
SPL	parameter	24.351	3.481	0.067
	group effect	0.696	0.100	0.754
α-Ratio	parameter	199.501	54.334	<0.001
	group effect	0.630	0.172	0.680
Loud reading				
F_0	parameter	8,052.281	47.648	<0.001
	group effect	48.028	0.284	0.596
SPL	parameter	96.983	10.300	<0.001
	group effect	0.509	0.054	0.817
α-Ratio	parameter	270.176	94.033	<0.001
	group effect	0.156	0.054	0.817
Sustained [a:]				
Mean jitter	parameter	2.115	31.563	<0.001
	group effect	0.037	0.556	0.459
SD of jitter	parameter	0.087	13.479	0.001
	group effect	0.006	0.999	0.322
Shimmer	parameter	0.359	9.999	0.003
	group effect	0.086	2.384	0.128
Self-evaluation				
Production	parameter	33,146.881	46.397	<0.001
	group effect	285.454	0.400	0.530
Quality	parameter	16,022.737	14.367	<0.001
	group effect	232.815	0.209	0.649
Throat	parameter	4,799.841	15.211	<0.001
	group effect	600.515	1.903	0.173
Perceptual evaluation				
Strain	parameter	546,127.931	35.651	<0.001
	group effect	3,572.557	0.233	0.631
Breathiness	parameter	981,370.976	87.996	<0.001
	group effect	2,509.283	0.225	0.637
Hoarseness	parameter	881,326.897	63.650	<0.001
	group effect	1,762.339	0.127	0.723
Evaluation of loud reading				
Firmness	parameter	320,591.502	36.024	<0.001
	group effect	96.050	0.011	0.918
Pitch	parameter	258,902.356	49.694	<0.001
	group effect	740.017	0.142	0.708

Values at the beginning of term set as covariates. Degree of freedom 1. Significance of differences shown with p values. All differences were non-significant (p > 0.05).

Table 7. Perceptual evaluation of reading samples (habitual loudness) by 3 speech trainers

Listening evaluations	Beginning of term			End of term			Beginning versus end	
	morning	afternoon	difference	morning	afternoon	difference	difference in mean values	difference in changes
Strain of phonation								
VHL group (n = 29)	173.9 ± 178.1	265.1 ± 197.8	0.013	190.1 ± 165.3	266.1 ± 198.9	0.004	n.s.	n.s.
VM group (n = 30)	152.2 ± 136.2	207.5 ± 157.1	0.039	174.6 ± 154.3	260.3 ± 170.7	0.006	n.s.	n.s.
Breathiness								
VHL group (n = 29)	242.1 ± 181.1	174.6 ± 135.9	0.010	259.6 ± 193.0	177.1 ± 126.0	0.002	n.s.	n.s.
VM group (n = 30)	264.3 ± 184.1	202.6 ± 144.3	0.003	252.8 ± 205.6	201.1 ± 173.8	0.006	n.s.	n.s.
Hoarseness								
VHL group (n = 29)	222.7 ± 169.7	170.6 ± 165.0	0.034	239.0 ± 181.5	177.8 ± 142.4	0.021	n.s.	n.s.
VM group (n = 30)	220.1 ± 192.1	175.4 ± 166.3	0.020	229.5 ± 205.1	205.7 ± 184.9	n.s.	n.s.	n.s.

Mean values ± SD for listening evaluations for strain (0 = not at all, 1,000 = very strained), breathiness (0 = not at all, 1,000 = very breathy) and hoarseness (0 = not at all, 1,000 = very hoarse). Differences are expressed as p values. Significance of differences within groups determined by Wilcoxon signed-rank test; n.s. = non-significant, p > 0.05.

Table 8. Mean ± SD for listening evaluations regarding voice quality, firmness of voice[1] and pitch[2] from the perceptual evaluations of loud reading samples evaluated in random order by 3 speech trainers

Listening evaluations	Beginning of term			End of term			Beginning versus end	
	morning	afternoon	difference	morning	afternoon	difference	diffence in mean values	difference in changes
Firmness of phonation								
VHL group (n = 29)	527.0 ± 156.0	572.8 ± 163.2	n.s.	509.9 ± 135.7	531.2 ± 111.3	n.s.	n.s.	n.s.
VM group (n = 30)	536.8 ± 134.4	598.5 ± 141.1	0.002	508.1 ± 125.8	547.6 ± 137.6	0.037	0.026	0.06
Perceived pitch								
VHL group (n = 29)	523.6 ± 114.2	529.3 ± 128.5	n.s.	507.4 ± 91.6	562.6 ± 77.1	0.001	n.s.	n.s.
VM group (n = 30)	560.7 ± 112.3	611.6 ± 199.4	0.001	557.5 ± 122.5	575.9 ± 123.3	n.s.	n.s.	n.s.

All differences are expressed as p values; n.s. = non-significant, p > 0.05. Significance of differences within groups determined by Wilcoxon signed-rank test.

[1] Scale for firmness of voice: 0 = breathy, 500 = ordinary, 1,000 = strained.

[2] Scale for pitch: 0 = too low, 500 = suitable, 1,000 = too high for the speaker.

habitual loudness and loudly. The F_0, Leq and α-ratio also tended to be higher at the end of the working day with prolonged vowel phonation. However, jitter and shimmer in the vowel samples tended to be lower at the end of the day. The changes in the perturbation values were not statistically significant, though. In the VHL group, the mean F_0 was higher at the end of the term than at the beginning. According to repeated measures ANOVA, the point of time during the term significantly affected ($p < 0.001$) all of the parameters (table 5), i.e. the acoustic parameters adopted a higher value after the working day and in the end of the term, except for shimmer which decreased (table 2). However, at the end of the term, the groups did not differ from each other in the analysis of variance (general linear univariate) when the values at the beginning were set as covariates (table 6).

Perceptual Results

Tables 7 and 8 present the results of the perceptual analyses of the reading samples. In general, the values for strain, breathiness and hoarseness were low, suggesting that the subjects had normal voices. Strain increased but

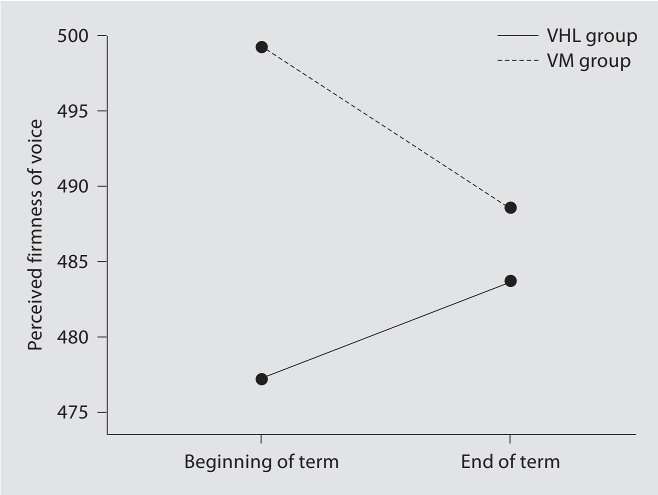

Fig. 1. Perceived firmness of voice production in reading samples at habitual loudness (scale: 0 = hypofunctional, 500 units = adequate firmness, 1,000 units = very hyperfunctional/strained).

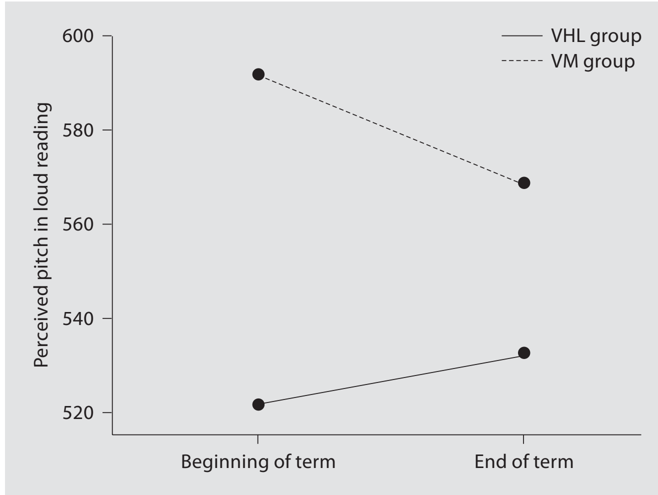

Fig. 2. Perceived pitch in loud reading (scale: 0 = too low, 500 units = suitable, 1,000 units = too high for the speaker).

breathiness and hoarseness decreased in the reading samples at habitual loudness after a working day. In loud reading, firmness and pitch increased. For the VM group, strain (firmness) during loud reading was lower at the end of the term than at the beginning of the term (p = 0.026; table 8). The groups did not differ from each other in the mean parameter values at the beginning or end of the term (tables 4 and 6). According to the repeated measures ANOVA results (table 5), the VM group differed from the VHL group in perceived pitch of loud reading. A further study of the material showed that the only difference between the groups was in perceived pitch of loud reading after the working day at the beginning of the term (higher in the VM group, p = 0.01). The trend was that on average, the firmness of the reading samples at habitual loudness and the pitch of loud reading increased in the VHL group, but decreased in the VM group, towards the end of the term (fig. 1, 2).

Self-Evaluation

Table 9 shows the results of the self-evaluation. In both groups, more tiredness of the throat, difficulty of phonation and somewhat poorer voice quality tended to be reported after the working day. However, the values were located very near 'ordinary'. On average, more tiredness of the throat, difficulty of phonation and somewhat poorer voice quality were reported at the end of the term after a working day. In the VHL group, the mean difficulty of phonation was significantly higher at the end of the term than at the beginning (p = 0.007, table 9). Figures 3 and 4 illustrate that, on average, the difficulty of phonation and tiredness of the throat increased more towards the end of the term in the VHL group. Self-perceived voice quality, in turn, deteriorated in the VM group (fig. 5). The changes in the parameters from the beginning to the end of the term were small, as were the differences between the groups. The groups did not differ significantly from each other at the end of term when the initial values were set as covariates (general linear univariate; table 6).

The VM group regarded the intervention as more positive (mean 66.6 mm VAS, SD 25.9, range 4–100) than the VHL group (mean 27.6 mm VAS, SD 19.2, range 0–73). The significance of the difference between the groups was $p < 0.001$. The positive effect of the treatment was enhanced knowledge (43.3%) in the VHL group, whereas the VM group reported enhanced relaxation (86.7%), voice endurance (66.7%), improved voice quality (56.7%) and awareness of the body (60%).

Self-Reports versus Acoustic and Perceptual Parameters

On average, there was no correlation between the acoustic parameters and self-evaluation. Only the Leq in habitual reading after a working day at the end of the term correlated weakly and negatively with self-reported voice quality (r = –0.39, p = 0.002). Self-evaluation and perceptual evaluation did not correlate either.

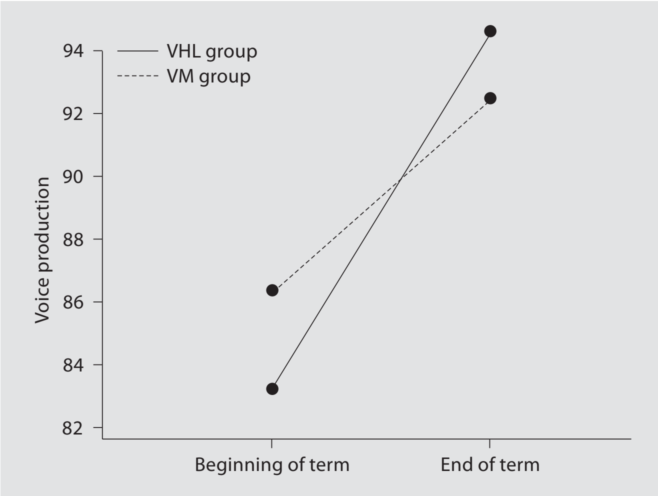

Fig. 3. Self-evaluation of voice production (scale in mm VAS: 0 = easy, 100 = ordinary, 200 = very difficult voice production).

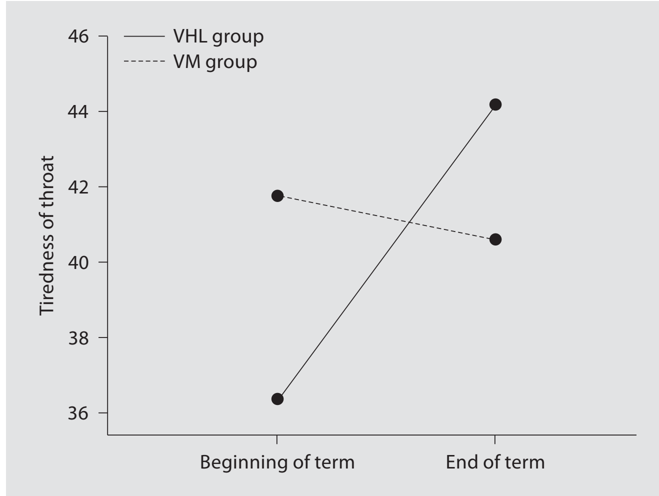

Fig. 4. Self-evaluation of tiredness of throat (scale in mm VAS: 0 = not tired, 100 = very tired).

Table 9. Mean values ± SD for self-evaluated difficulty of voice production ('phonation')[1], voice quality ('voice')[2] and tiredness of throat ('throat')[3]

Subjective evaluations	Beginning of term			End of term			Beginning versus end	
	morning	afternoon	difference	morning	afternoon	difference	difference in mean values	difference in changes
Phonation (200-mm VAS)	n = 30	n = 30		n = 30	n = 30			
VHL group	80.13 ± 36.1	86.50 ± 36.7	n.s.	91.63 ± 42.4	97.57 ± 41.4	n.s.	0.007	n.s.
VM group	83.00 ± 35.1	89.80 ± 34.8	n.s.	93.60 ± 36.2	91.30 ± 38.9	n.s.	n.s.	n.s.
Voice (200-mm VAS)	n = 30	n = 30		n = 30	n = 30			
VHL group	94.97 ± 36.1	96.90 ± 35.7	n.s.	95.37 ± 43.7	97.33 ± 41.2	n.s.	n.s.	n.s.
VM group	94.23 ± 30.9	97.30 ± 32.9	n.s.	102.07 ± 35.2	98.33 ± 41.7	n.s.	n.s.	n.s.
Throat (100-mm VAS)	n = 30	n = 30		n = 30	n = 30			
VHL group	32.00 ± 19.1	40.83 ± 22.1	0.033	40.63 ± 21.5	47.63 ± 23.9	0.056	n.s.	n.s.
VM group	35.33 ± 19.1	48.20 ± 19.8	0.002	34.67 ± 20.7	46.70 ± 24.3	0.013	n.s.	n.s.

Reported before and after a working day. All differences are expressed as p values. Significance of differences within groups determined by Wilcoxon signed-rank test; n.s. = non-significant, $p > 0.05$.

[1] Phonation: 0 = particularly easy, 100 = ordinary, 200 = very difficult.

[2] Voice: 0 = very good, 100 = ordinary quality, 200 = very poor quality.

[3] Throat: 0 = no tiredness at all, 100 = very tired.

Discussion

The main changes observed after a working day were increases in the F_0, Leq and α-ratio, and a tendency for a decrease in jitter and shimmer, even though the changes in perturbation values were not significant (see also Ilomäki et al. [23]). The observations on the F_0, Leq and spectral changes are in line with many studies on the effects of vocal loading [7, 21, 28, 30–33]. These changes could result from increased muscle activity in the vocal organs as an adaptation to vocal loading. The results for the effects of vocal loading on jitter and shimmer vary in the literature [34], most likely depending on the subjects, the amount of loading and the amount of consequent fa-

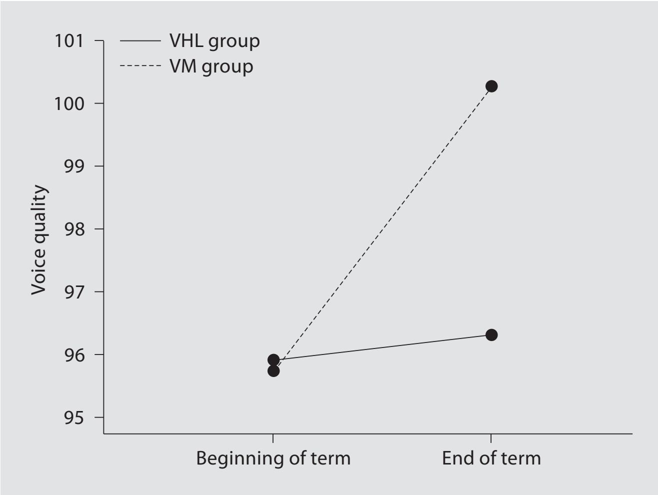

Fig. 5. Self-reported evaluation of voice quality (scale in mm VAS: 0 = good, 100 = ordinary, 200 = very poor voice quality).

tigue. A higher F_0 and Leq (SPL), and thus higher muscle activity in the voice production muscles would imply decreased perturbation, while loading changes in the vocal fold tissue and deterioration in the neural control could increase perturbation [35, 36].

Similar trends observed after a working day were also observed when the beginning and end of the term were compared. This suggests that the average activity in the vocal organ and symptoms of vocal fatigue increase towards the end of the term. This seems to imply a need for interventions to assist voice professionals in coping with cumulative loading.

In the present study, the mean F_0, Leq and α-ratio did not correlate with the self-reports. Furthermore, the changes within the mean F_0, Leq and α-ratio did not correlate with the self-reports either. This is in contrast to the findings by Rantala and Vilkman [20] and Mäki et al. [19]. This may suggest that, in experienced voice professionals, the type of voice production (reflected in these acoustic variables) is not so clearly related to symptoms of vocal fatigue. It is also plausible that other factors, e.g. the amount of voice use and the amount of vocal rest per day, affect the prevalence of symptoms of vocal fatigue. According to the results of Mäki et al. [19], a stronger correlation was found between symptoms and the relative F_0 and SPL (relative to the subjects' individual ranges) than with the absolute mean F_0 and SPL. Furthermore, individual sensitivity to loading changes should affect the self-evaluation and, thus, interfere with the results.

On average, perceived firmness of phonation in reading aloud text samples at habitual loudness and perceived pitch in loud reading increased in the VHL group, but decreased in the VM group. Since the values of firmness varied in a relatively narrow range and were near 'adequate', it is not possible to state whether the changes as such should be regarded as positive or negative. The pitch, in turn, was evaluated to be, on average, too high in loud reading in both groups. Hence, the lowering of pitch in the VM group may be regarded as positive. The VHL group reported significantly more tiredness of the throat and difficulty of phonation at the end of the term than at the beginning. Additionally, the mean F_0 was higher at the end of the term. Tiredness of the throat increased in the VHL group, but decreased in the VM group. The difficulty of phonation increased more in the VHL group. These results would suggest that VM may have helped the teachers to maintain their vocal well-being during the term. The results also concur with self-reports of the general benefits of the intervention. A significantly more positive effect was reported from the VM group than from the VHL group. It should be noted that all changes observed in the various parameters from the beginning to the end of the term and all differences between the groups were small (e.g. vertical scale in total was only 10 mm VAS in fig. 4 and 6 mm in fig. 5). However, they seem to indicate behavioural trends and, therefore, deserve reporting.

The VHL and VM groups did not differ from each other when the values at the beginning of the term were set as covariates. One may ask whether the intervention time was too short. On the other hand, the present study design lacks a clear neutral control group as both the VHL and VM groups may be expected to have positive effects. Mere adaptation to loading may also improve vocal well-being during the term. Obviously, there is a lot of individual variation in the response to any intervention. This variation may reflect differences in learning strategies, tolerance to such physical contact as massage (also whether or not a subject has had any earlier experience with VM), personal opinion of the teacher/therapist, etc.

A positive influence on vocal function was reported in both the VHL group and the VM group, but the degree of the effect was significantly higher in the VM group. Subjects in the VM group reported increased relaxation, improved body awareness and improved vocal endurance. The generally very positive experience reported after VM may reflect the fact that the participants of the VM group received personal attention from the therapist,

which is prone to build positive interaction and confidentiality typically related to massage treatment [37]. In line with the present results, an earlier study [25] reported many positive sensations after VM, while no significant acoustic or perceptual effects were found. One reason may be that the type of change in acoustic or perceptual parameters that can be regarded as positive are prone to differ depending on the starting point (i.e. an increase in the α-ratio and perceived firmness of voice production may be regarded as a positive change if the starting point is hypofunctional voice production, while contrastive changes may be aimed at when the subject has a hyperfunctional voice type).

The general evaluation of the interventions did not correlate with working-day-related changes in the self-reports. This may be related to the different time window and focus used in the evaluations. Another possible intervening factor is that both the VHL and the VM treatment may increase the subjects' awareness of the symptoms of vocal fatigue.

A combination of VHL or VM with voice training warrants further study. For instance, Roy and Leeper [13] concluded that massage combined with traditional voice therapy might result in more lasting positive changes. A further study will concern the long-term effects of the interventions after 6-month and 12-month follow-up periods.

Conclusions

In comparing the end of the term with the beginning, the mean F_0 (in reading samples) was higher and more difficulty of phonation was reported in the VHL group. Perceived pitch in loud reading increased in the VHL group, but decreased in the VM group. The groups did not differ from each other when parameter values at the beginning were set as covariates, although wide individual differences did exist. Significantly greater positive effects were reported after VM, and this form of treatment may assist teachers in maintaining vocal well-being during the term. A follow-up study will be conducted 6 and 12 months after the interventions. The combined effects of VHL/VM and voice training warrant further study. Investigation of optimal outcome measures for improvement of vocal well-being in subjects with normal voices is also warranted.

Acknowledgements

The present study was supported by the Finnish Work Environment Fund (grant No. 103309), the Scientific Foundation of the City of Tampere and the Voice Massage Therapist Association in Finland. We would especially like to thank laboratory technician Mr. Jussi Helin for assisting in the analyses of the material and Ms. Virginia Mattila, MA, for checking the language of the manuscript.

References

1 Titze IR, Lemke J, Montequin D: Populations in the US workforce who rely on voice as a primary tool of trade: a preliminary report. J Voice 1997;11:254–259.
2 Mattiske JA, Oates JM, Greenwood KM: Vocal problems among teachers: a review of prevalence, causes, prevention, and treatment. J Voice 1998;12:489–499.
3 Thomas G, de Jong FI, Cremers CW, Kooijman PG: Prevalence of voice complaints, risk factors and impact of voice problems in female student teachers. Folia Phoniatr Logop 2006;58:65–84.
4 Sliwinska-Kowalska M, Niebudek-Bogusz E, Fiszer M, Los-Spychalska T, Kotylo P, Sznurowska-Prxygocka B, Modrzeweska M: The prevalence and risk factors for occupational voice disorders in teachers. Folia Phoniatr Logop 2006;58:85–101.
5 Stemple JC, Stanley J, Lee L: Objective measures of voice production in normal subjects following prolonged voice use. J Voice 1995; 10:127–133.
6 Rantala L, Vilkman E, Bloigu R: Voice changes during work: subjective complaints and objective measurements for female primary and secondary schoolteachers. J Voice 2002;16:344–355.
7 Sala E, Laine A, Simberg S, Pentti J, Suonpää J: The prevalence of voice disorders among day care centre teachers compared with nurses: a questionnaire and clinical study. J Voice 2001;15:413–423.
8 Vilkman E: Voice problems at work: a challenge for occupational safety and health arrangement. Folia Phoniatr Logop 2000;52: 120–125.
9 Jiang J, Titze IR: Measurement of vocal fold intraglottal pressure and impact stress. J Voice 1994;2:132–144.
10 Chan RW: Does the voice improve with vocal hygiene education? A study of some instrumental voice measures in a group of kindergarten teachers. J Voice 1994;8:279–291.
11 Bovo R, Galceran M, Petrucelli J, Hatzopoulos S: Vocal problems among teachers: evaluation of a preventive voice program. J Voice 2007;21:705-722.
12 Roy N, Ford CN, Bless DM: Muscle tension dysphonia and spasmodic dysphonia: the role of manual laryngeal reduction in diagnosis and management. Ann Otol Rhinol Laryngol 1996;105:851–856.
13 Roy N, Leeper HA: Effects of the manual laryngeal musculoskeletal tension reduction technique as a treatment for functional voice disorders: perceptual and acoustic measures. J Voice 1993;7:242–249.

14 Ternström S, Andersson M, Bergman U: An effect of body massage on voice loudness and phonation frequency in reading. Logoped Phoniatr Vocol 2000;25:146–150.
15 Ylitalo R, Hammarberg B: Voice characteristics, effects of voice therapy, and long-term follow-up of contact granuloma patients. J Voice 2000;14:557–566.
16 Dharmananda S: Treatment of throat and voice disorders with Chinese medicine. 2002. http://www.itmonline.org/arts/throat.htm (accessed April 4, 2007).
17 Mathieson L, Hirani SP, Epstein R, Baken RJ, Wood G, Rubin JS: Laryngeal manual therapy: a preliminary study to examine its treatment effects in the management of muscle tension dysphonia. J Voice, E-pub ahead of print.
18 Laukkanen AM, Leppänen K, Tyrmi J, Vilkman E: Immediate effects of 'voice massage' treatment on the speaking voice of healthy subjects. Folia Phoniatr Logop 2005;57:163–172.
19 Mäki E, Niemi HM, Lundén S, Laukkanen AM: F_0, SPL and vocal fatigue in a vocally loading test. Proc 25th World Congr Int Assoc Logop Phoniatr, Montreal, August 2001.
20 Rantala L, Vilkman E: Relationship between subjective voice complaints and acoustic parameters in female teachers' voices. J Voice 1999;13:484–495.
21 Jónsdóttir VI: The Voice – An Occupational Tool. A Study of Teachers' Classroom Speech and the Effects of Amplification; dissertation, University of Tampere, 2003, No 969.
22 Scherer RC, Titze IR, Raphael BN, Wood RP, Ramig LA, Blager FB: Vocal fatigue in a trained and an untrained voice user; in Baer T, Sasaki C, Harris KS (eds): Vocal Fold Physiology: Laryngeal Function in Phonation and Respiration. Boston, College-Hill Press, 1987, pp 533–555.
23 Ilomäki I, Laukkanen AM, Leppänen K, Vilkman E: Effects of voice training and voice hygiene education on acoustic and perceptual speech parameters and self-reported vocal well-being in female teachers. Logoped Phoniatr Vocol 2008;33:83–92.
24 Laukkanen AM, Ilomäki I, Leppänen K, Vilkman E: Acoustic measures and self-reports of vocal fatigue by female teachers. J Voice 2008;22:283–289.
25 Laukkanen AM, Leppänen K, Ilomäki I: Self-evaluation of voice as a treatment outcome measure. Folia Phoniatr Logop 2009;61:57–65.
26 Sukanen O, Sihvo M, Rorarius E, Lehtihalmes M, Autio V, Kleemola L: Voice activity and participation profile (VAPP) in assessing the effects of voice disorders on patient's quality of life: validity and reliability of the Finnish version of VAPP. Logoped Phoniatr Vocol 2007;32:3–8.
27 Ma EP, Yiu EM: Voice activity and participation profile: assessing the impact of voice disorders on daily activities. J Speech Lang Hear Res 2001;44:511–524.
28 Toivonen R: Intelligent Speech Analyser™ (ISA) Software. 2007. http://www.sci.fi/~pitchsys/ (accessed February 13, 2008).
29 Granqvist S: Computer methods for voice analysis; doctoral dissertation, KTH, Stockholm, 2003.
30 Åkerlund L, Gramming P: Average loudness level, mean fundamental frequency and subglottal pressure; comparison between female singers and nonsingers. J Voice 1994;8:263–270.
31 Södersten M, Hammarberg B: Effects of voice training in normal-speaking women: videostroboscopic, perceptual, and acoustic characteristics. Scand J Logop Phoniatr 1993;18:33–42.
32 Vilkman E, Lauri ER, Alku P, Sala E, Sihvo M: Ergonomic conditions and voice. Logoped Phoniatr Vocol 1998;23:11–19.
33 Laukkanen AM, Järvinen K, Artkoski M, Waaramaa-Mäki-Kulmala T, Kankare E, Sippola S, Syrjä T, Salo A: Changes in voice subjective sensations during a 45-min vocal-loading test in female subjects with vocal training. Folia Phoniatr Logop 2004;56:335–346.
34 Rantala L: Ääni työssä. Naisopettajien äänenkäyttö ja äänen kuormittuminen (Voice at Work: Female Teachers' Use and Loading of Voice); dissertation, University of Oulu, 2000, Acta Universitatis Ouluensis, Humaniora, B 37.
35 Fant G: Acoustic Theory of Speech Production. The Hague, Mouton, 1960.
36 Titze IR: Principles of Voice Production. Englewood Cliffs, Prentice-Hall, 1994.
37 Svennevig H: Hyvän olon hoidot (Therapies of Well-Being); dissertation, University of Tampere, 2003, No 949.