Logopedics Phoniatrics Vocology. 2005; 30: 63–71

ORIGINAL ARTICLE

The effect of a 'voice course' on the voices of people with and without pathologies: Preliminary observations*

OFER AMIR, MICHAL DUKAS & RACHEL SHNAPS-BAUM

Department of Communication Disorders, Sackler Faculty of Medicine, Tel-Aviv University, Israel

Abstract
Several institutions provide voice practice for their teaching staff. These 'voice courses' provide a unique treatment situation, in which some participants have a diagnosed laryngeal pathology, while others have no vocal complaints. Therefore, these voice courses can be viewed as group voice therapy as well as preventive voice treatment. The present study was aimed at providing a preliminary voice evaluation of participants in such a group, prior to and following treatment, using perceptual and acoustic analyses. Sixteen male teachers, who enrolled in a voice course, participated in eight consecutive sessions. Of this group, seven teachers were diagnosed with laryngeal pathologies, while the others had none. All participants were recorded before and after treatment. Recordings were analyzed acoustically, and also evaluated perceptually by ten experienced speech pathologists. Results indicated that: (i) voice quality improved after the voice course, (ii) vocal improvement was more pronounced in the pathological group than in the nonpathological group, and (iii) the acoustic analysis paradigm yielded results that were not always readily related with those of the perceptual paradigm.

Key words: *Acoustic analysis, perceptual evaluation, voice assessment, voice course, voice treatment*

Introduction

The effectiveness of voice therapy has been demonstrated and discussed previously, using various experimental paradigms. Different therapy protocols were employed and compared, with a general agreement that voice therapy is effective for improving voice quality and laryngeal dysfunction, in many non-organic as well as organic based conditions (for review see, for example, Ramig & Verdolini (1)). Conventional voice therapy and preventive treatment differ, primarily, in their goals. While the former is aimed at recovering a vocal dysfunction, the latter is aimed at preventing it. The two therapeutic approaches also target different clients. Voice therapy clients typically have initial complains about vocal symptoms, and they are examined by a phoniatrician before therapy. Conversely, preventive treatment clients, generally, do not complain about vocal or laryngeal symptoms. Instead, they are interested in preserving their vocal health and learning to prevent such problems in the future. These clients' occupations (e.g., vocal performers, sales persons and teachers) usually require excessive vocal use. In addition, people who attend preventive voice treatments are typically not examined by a phoniatrician prior to therapy, since it is often administered in the work place or in other non-medical settings. Finally, while conventional voice therapy is mostly conducted on an individual basis, preventive voice treatments are commonly conducted in groups.

Teachers are more often sensitive to voice attrition than people in other professions. Therefore, many educational institutions provide voice practice in group settings for their teachers, even before they complain about their voice. This practice is referred to as a 'voice course'. These voice courses provide a unique treatment situation, in which participants are markedly heterogeneous: some could have a diagnosed laryngeal pathology, while others have no complaints about their voice. Consequently, these

* Was presented at PEVOC6, 2005.

Correspondence: Ofer Amir, PhD, Department of Communication Disorders, The Chaim Sheba Medical Center, Tel-Hashomer, Israel 52621. Fax: +972-3-5352868. E-mail: oferamir@post.tau.ac.il

(accepted 29 June 2005)

ISSN 1401-5439 print/ISSN 1651-2022 online © 2005 Taylor & Francis
DOI: 10.1080/14015430500233229

voice courses can be viewed as group voice therapy for some patients and as preventive voice treatment for others. A preliminary study, conducted in our department (2), has shown that teachers who underwent such 'voice courses' perceived them as advantageous, and in many cases reported subjective improvement in their voice. One goal of the present study was, therefore, to provide a preliminary evaluation of the voice of participants in such a group, prior to and following treatment, using subjective and acoustic analyses.

While the effect of voice therapy has been evaluated in numerous studies (3–7), the literature on the effect of preventive voice treatment is limited. Kindergarten teachers, for example, were reported to improve their voices after an hour-and-a-half workshop on vocal hygiene (8). Interestingly, these teachers did not practice any traditional voice or respiratory exercises. Instead, they attended a single lecture on vocal hygiene, and were then instructed to follow a set of recommendations for two months. Different results were reported in a study on a group of singing students (9). These students received four one-hour sessions on vocal hygiene. Based on self-evaluations of the participants, only minimal vocal changes were noticed after the course, although most of them reported a high degree of benefit and satisfaction with the additional knowledge they gained. In a different study (10), a group of call-center customer service advisors received a short voice training course. The two-day course included an informative lecture on basic principles of voice anatomy and voice production, in addition to vocal hygiene. The remaining time of the course was dedicated to vocal exercises. After the course, most participants reported voice improvement and reduction in vocal fatigue. Due to the limited data on the effectiveness of preventive voice treatment, the inconsistent findings, and the fact that most studies relied on subjective self-reports, one target of the present study was to provide additional information on the effect of a voice course, by combining acoustic and perceptual evaluation.

There is no general agreement with regard to which method is most appropriate for evaluating the outcome of voice therapy. Several studies evaluated therapy effect using *subjective* judgments performed by trained listeners (3,11), or self-evaluations made by the clients themselves (for review, see Hogikyan & Rosen (12)). Other studies employed more *objective* measures, such as acoustic analyses of voice quality (4,13), or combinations of acoustic and aerodynamic measures (5,14,15). Subjective evaluation is generally regarded as the gold standard for voice evaluation, chiefly because the preliminary identification of a voice problem, which is done initially by the patient, is perceptual and therefore subjective (16,17). However, subjective judgment has fluctuating and relatively low inter- and intra-judge reliability (11). In addition, the definitions of the perceptual vocal measures (e.g., hoarseness, strain, roughness) may vary considerably among different professionals (17). These methodological issues present an obstacle for comparing different studies of vocal quality (1). Although several attempts have been made to construct standardized scales for subjective evaluation of voice quality (11,18–20), the validity and reliability of these scales, as well as their clinical merit in different settings, are yet to be established.

Acoustic analysis has the potential benefit of measuring and quantifying subtle differences in voice quality more reliably than most perceptual measures. Normal and pathological voices were shown to differ in various acoustic measures, such as fundamental frequency, amplitude- and frequency-perturbation and different signal-to-noise indices (21). Although higher perturbation values, for example, are generally associated with pathological voices (22,23), the relation between specific *acoustic* and *perceptual* measures is not fully understood yet. Moreover, it is not clear whether acoustic measurements considered abnormal, indeed represent pathological or deviant voice quality consistently and reliably. Therefore, an additional research question of the present study was of a methodological nature. We were interested to learn whether sustained vowels recorded after a voice course would differ from those recorded by the same participants before the course, and whether an acoustic analysis paradigm would illustrate these differences similarly or differently than a perceptual paradigm.

Method

Participants

Twenty-five male teachers, who enrolled in a program that qualifies teachers to be special-education teachers, volunteered to participate in this study. After obtaining the approval from our institutional review board and written consent from all volunteers, all participants were referred to a phoniatrician, for a laryngeal exam. Of the 25 volunteers, only 16 who underwent the examination and had completed the experimental protocol were eventually included in the study. Mean age for this group was 38.1 years (SD = 10.7), mean height was 174.8 cm (SD = 8.2) and mean weight was 87 kg (SD = 13.7). Mean number of years of teaching-experience was 11.6 (SD = 9.8). All participants reported no remarkable medical history, and four of them reported prior experience with different forms of speech therapy.

As part of the qualifying requirements of the special-education teachers' program, all 16 teachers participated in a voice course. The purpose of this course is formally defined as providing the teachers with the knowledge and capability to preserve and improve their voice quality, in order to prevent vocal dysfunction in their future careers. For the purpose of this study, the 16 teachers were regarded as two experimental groups, based on the diagnoses obtained from their laryngeal exams. The first group, Pathological, consisted of seven teachers, of which three were diagnosed with vocal nodules, one with incomplete adduction of the glottis during phonation, and the remaining three were diagnosed with both vocal nodules and incomplete adduction of the glottis. The second group, Normal, consisted of nine teachers who had no laryngeal pathology. Group means of anamnestic variables, including age, height, weight, years of teaching experience, and subjective self-evaluation of voice quality (on a 1–10 scale, where 10 represents 'highly satisfied' and 1 represents 'extremely unsatisfied') are reported in Table I.

To verify that the two experimental groups were comparable, with regard to all reported demographics, a series of separate t-tests were conducted between the two groups, one for each demographic parameter. No significant differences were found between the two groups for any of the anamnestic parameters ($p > 0.05$).

Voice course program

All 16 teachers were randomly assigned to two equal-size training groups. Both training groups were treated by the same speech-language pathologist (SLP) once a week, over a period of approximately two months. To reduce possible bias effects, the SLP was not informed of the laryngeal findings of each participant. Nonetheless, he was aware that several of the participants in both groups were diagnosed with laryngeal pathologies.

The voice course consisted of eight 45-minute sessions. It incorporated elements from both direct and indirect therapeutic approaches. The application of a combined approach was chosen based on previous studies that suggested that treating teachers with a vocal hygiene approach alone might not be sufficiently effective (5). The first session consisted of a personal introduction of all participants, followed by a general overview of the vocal mechanism anatomy and physiology. Then, a vocal hygiene approach was advocated. The SLP and the participants identified and discussed sources of vocal abuse, specific to the daily activities of the group, and all participants were instructed to: (i) avoid or reduce excessive vocal use; (ii) reduce or preferably stop smoking (when applicable); (iii) avoid food that is more likely to provoke gastro-esophageal reflux (GER); (iv) increase water intake to two liters per day.

To monitor vocal-hygiene modified behaviors, each participant completed a daily chart of vocal use, on which he was also asked to identify risk factors that he encountered, and ways in which he attempted to deal with them. These charts were, then, submitted and discussed weekly by the participants during each session and were utilized for increasing self-awareness to vocal use and vocal hygiene.

The second and third sessions focused on respiration control. Specifically, each participant identified his type of breathing, and all participants were taught diaphragmatic and mid-support breathing (24). However, respiration practice did not focus directly on diaphragmatic breathing. Instead, it focused on reducing breathing effort in general. Then, these basic principles were gradually implemented to isolated words, sentences and eventually connected speech. During these exercises, special attention was given to reducing 'hard attacks', which are commonly used by Hebrew speakers, and instead facilitating 'soft' vocal onset of specific consonants (especially /ʕ, ʔ /). Breathing exercises were practiced repeatedly in the opening of each session, throughout the course.

Table I. Mean age, height, weight, years of teaching experience, and self-evaluation of voice-quality (on a 1–10 rating scale), in the pathological and normal rroups.

	Pathological ($n = 7$)		Normal ($n = 9$)	
Subject's variables	Mean	(SD)	Mean	(SD)
Age (years)	40.3	(12.9)	36.4	(9.1)
Height (cm)	172.3	(7.1)	176.7	(5.0)
Weight (kg)	80.4	(7.6)	92.1	(15.6)
Teaching experience (years)	12.1	(12.3)	11.2	(8.2)
Voice quality self-evaluation	6.7	(2.1)	7.2	(2.9)

The following five sessions focused on voice production. During this section of the course, a chant therapy approach (3,25) was practiced as the major therapy technique. The participants were informed that the technique would be used only briefly, as a voice training device. After the participants had learned how to chant, they were asked to read sentences at various lengths, in a chant and in normal voice. Gradually, voice productions were increased and varied, until all participants were able to produce chant talk, with relative ease. Then, they were asked to gradually reduce the chanting quality, while maintaining soft glottal attack, and vocal resonance. A number of participants experienced difficulties in producing chant voice, or during the generalization process. For these, additional facilitating approaches were used on an individual basis, such as yawn-sigh, open-mouth (25) and chewing (26).

While the five concluding sessions focused on voice production exercises, each of them opened with practicing respiratory exercises. In addition, each session included a group discussion on various topics related to vocal hygiene that were raised by the SLP or by members of the training groups.

Recording procedure and instrumentation

All participants were recorded twice, prior to the first session, and one week following the eighth session. Each participant was recorded while producing the three vowels /i/, /a/ and /u/ twice, in a random order that was changed among sessions. Each vowel was sustained for 3 seconds. These vowels were selected because they are commonly used for experimental and clinical evaluation and, in addition, they represent different articulatory gestures in many languages (21), as well as in Hebrew (27).

All recordings were performed individually while the subject was seated in a quiet room. A Sony ECM-T150 microphone (Sony, Tokyo, Japan), attached to a headset, was placed approximately 5 cm from the subject's mouth. The signal was recorded directly onto a computer, using a Goldwave© (version 4.23) software with a sampling rate for signal capturing of 48 kHz (16-bit).

Analysis of recordings

Each sustained vowel was fed to the MDVP (Multi Dimensional Voice Program) software (model 5105, version 2.0 (Kay Elemetrics, Lincoln Park, NJ)). For each production, five acoustic parameters were measured. These parameters included fundamental frequency (F0), a frequency perturbation measure (Jitter), an amplitude perturbation measure (Shimmer), and two noise indices NHR (Noise-to-Harmonic Ratio) and VTI (Voice Turbulence Index). Sampling rate for analysis was set at 44 kHz. To evaluate reliability, a random set of 24 voice samples were re-measured and analyzed by a second experimenter. A paired-sample t-test revealed no significant difference between the two sets of data for all acoustic measures ($p > 0.05$). In addition, Pearson correlation analyses yielded high correlation between the two measurements made by the two experimenters, for all acoustic parameters ($0.88 \leq r \leq 0.99$; $p < 0.001$), indicating high inter-judge reliability.

Listener judgment

For the purpose of subjective listeners' evaluation, the two first productions of the vowels /a/ and /i/ were selected from the voice samples of each teacher before and after therapy, providing four vowels from each speaker. To reduce the duration of the listening task, the recordings of the vowel /u/ were not included in this section of the study. In addition, one speaker's voice samples were entirely excluded from the perceptual analysis, due to background noise that was identified in the 'after therapy' conditions, which might have affected listeners' judgment. As a result, 60 vowels (15 speakers $\times 2$ vowels $\times 2$ experimental conditions) were arranged in a random order for presentation to the listeners. These vowels were digitally recorded onto a compact disk for presentation to the listeners, with 20-second intervals between every two successive vowels.

Perceptual evaluation of voice quality was performed by ten experienced SLPs from three different medical centers. Mean age of the listeners group was 48.2 years (SD = 10.6), and mean number of years of clinical experience was 23.0 (SD = 9.1). Each SLP listened to the recordings individually in a quiet room, through a Panasonic SL-CT480 disk player and Sony MDR-CD380 headphones. The SLPs were asked to listen to each vowel, and evaluate its quality on six rating scales, one for each voice characteristic: Pitch, Roughness, Strain, Breathiness, Resonance and Stability. Each scale consisted of seven 'X' marks, titled with the appropriate vocal quality; where one end of the scale was labeled 'Normal' and the other was labeled 'Severe', as illustrated in Figure 1.

The judges were asked to circle the appropriate 'X' that corresponded with their evaluation of voice characteristic for each voice sample. Total duration of the listening task was approximately 25 minutes.

	Normal						Severe
Pitch	X	X	X	X	X	X	X
Roughness	X	X	X	X	X	X	X
Strain	X	X	X	X	X	X	X
Breathiness	X	X	X	X	X	X	X
Resonance	X	X	X	X	X	X	X
Stability	X	X	X	X	X	X	X

Figure 1. Perceptual rating scales completed by listeners for each voice sample.

Results

Acoustic analysis

For each vowel (/a/, /i/ and /u/) and recording condition (before and after training), group means were calculated as follows. First, for each subject, the repeated recordings of every vowel, obtained in one session, were averaged. Then, group means were calculated for the three vowels separately, at the two recording conditions. These values are presented in Table II.

In general, the pathological group had higher values of most acoustic parameters than the normal group. These group differences were statistically significant for the jitter measure ($F(1, 14) = 6.64$, $p = 0.022$). A similar trend for group differences was observed for most other acoustic measures, yet these differences failed to reach statistical significance ($p > 0.05$).

Data show that most acoustic measures improved (that is to say were lowered) after the voice course in both groups. Statistical analyses revealed a significant main effect for training, for jitter ($F(1, 14) = 8.97$, $p = 0.010$), shimmer ($F(1, 14) = 22.78$, $p < 0.001$), and NHR ($F(1, 14) = 6.71$, $p = 0.021$). No significant training effect was found for the F0 and VTI measures ($p > 0.05$).

While a general trend for improvement was observed for both groups, the magnitude and consistency of improvement was greater in the pathological group than it was in the normal group. This was confirmed by a significant Group × Training interaction that was found for jitter ($F(1, 14) = 5.82$, $p = 0.030$) and shimmer ($F(1, 14) = 11.84$, $p = 0.004$). This interaction is illustrated in Figure 2. No significant Group × Treatment interaction was found for the F0, NHR and VTI measures ($p > 0.05$).

Vowel differences were not the interest of the present study. Yet, it was included in the analyses to improve validation of the results. A significant main effect for vowel was found for F0 ($F(1, 14) = 9.27$, $p = 0.009$), shimmer ($F(1, 14) = 7.11$, $p = 0.018$), NHR ($F(1, 14) = 11.48$, $p = 0.004$) and VTI ($F(1, 14) = 37.11$, $p < 0.001$). Post-hoc contrast analyses revealed that F0 values of the low vowel /a/ were significantly lower than that of both high vowels (/i/ and /u/) ($p < 0.05$). Shimmer values were significantly higher for the /a/ vowel than for the vowels /i/ and /u/ ($p < 0.05$). For the noise indices, all pairs of vowels yielded significant differences ($p < 0.05$). No significant vowel differences were observed for jitter ($p > 0.05$).

Listeners' judgment

To enable statistical analyses of the listeners' judgments, each response on the 7-point rating scales was first converted to a numerical scale such that '0' corresponds with 'normal' and '6' with 'severe'. Then, group means for the two vowels that were included in the analysis (/a/ and /i/) were calculated for each perceptual measure, before and after the voice course. These values are presented in Table III. Note that, in this table, higher values represent more pathological voice quality.

Table II. Mean values and standard deviations (in parentheses) of F0, jitter, shimmer, NHR and VTI of the normal and pathological groups for the vowels /a/, /i/ and /u/ obtained before and after the voice-course.

		Normal			Pathological		
Acoustic measure	Training	/a/	/i/	/u/	/a/	/i/	/u/
F0 (Hz)	Before	114.24 (9.62)	115.69 (9.62)	116.72 (9.40)	118.63 (19.59)	122.69 (20.60)	122.49 (21.36)
	After	116.38 (9.90)	118.58 (10.46)	118.16 (9.49)	116.16 (19.75)	121.91 (25.79)	122.63 (27.18)
Jitter (%)	Before	0.722 (0.27)	0.786 (0.26)	0.907 (0.46)	1.667 (1.64)	1.734 (0.69)	1.098 (0.31)
	After	0.800 (0.54)	0.755 (0.34)	0.644 (0.17)	0.691 (0.15)	1.178 (0.75)	0.631 (0.18)
Shimmer (%)	Before	4.076 (1.59)	3.101 (0.87)	3.683 (3.99)	7.792 (2.94)	4.356 (1.38)	4.518 (2.09)
	After	3.858 (1.69)	3.261 (1.38)	2.554 (1.46)	4.557 (2.33)	2.530 (0.77)	2.265 (0.95)
NHR	Before	0.142 (0.01)	0.134 (0.02)	0.136 (0.03)	0.191 (0.09)	0.140 (0.02)	.132 (0.01)
	After	0.146 (0.02)	0.125 (0.02)	0.130 (0.01)	0.145 (0.01)	0.122 (0.02)	0.121 (0.01)
VTI	Before	0.050 (0.01)	0.047 (0.01)	0.035 (0.01)	0.050 (0.02)	0.061 (0.02)	0.032 (0.01)
	After	0.044 (0.01)	0.063 (0.02)	0.028 (0.01)	0.040 (0.01)	0.042 (0.01)	0.029 (0.01)

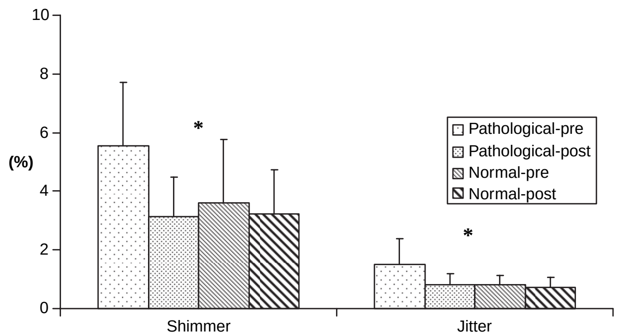

Figure 2. Group means (±1 standard deviation) for shimmer and jitter for the normal and pathological groups pre- and post-voice course. Each bar represents the average of the three vowels collapsed together for each recording condition. * $p < 0.05$.

In general, mean listeners' ratings on the 7-point scale for all participants ranged between 1.90 and 4.14. The pathological group was rated higher (that is to say more severe) than the normal group on all perceptual scales. Yet, significant overall group differences were found only for the 'pitch' and 'strain' scales ($F(1, 13) = 5.79$, $p = 0.032$ and $F(1, 13) = 6.31$, $p = 0.026$, respectively). Marginal group differences ($0.05 < p < 0.10$), however, were found for the 'breathiness' and 'resonance' scales ($F(1, 13) = 4.44$, $p = 0.055$ and $F(1, 13) = 3.81$, $p = 0.073$, respectively), whereas no significant group differences were found for the 'roughness' and 'stability' scales.

Although most subjective ratings decreased (that is to say were improved) after the voice course, no significant training effect was found for any of these scales. Marginally significant differences ($0.05 < p < 0.10$) were found only for the breathiness, resonance and stability scales ($F(1, 13) = 4.24$, $p = 0.060$, $F(1, 13) = 3.35$, $p = 0.090$ and $F(1, 13) = 4.62$, $p = 0.051$, respectively). No significant Training × Group interaction was found for any of the perceptual scales ($p > 0.05$). Yet, in general, the magnitude of the improvement following training was greater in the pathological group than it was in the normal group.

The vowel /i/ was rated higher than the vowel /a/ using the strain scale in both groups and recording conditions ($F(1, 13) = 9.76$, $p = 0.008$). No other significant vowel differences were observed using any other perceptual scale ($p > 0.05$), and no Vowel × Group interaction was found for any of the perceptual scales.

Discussion

The primary focus of this study was on the effect of a voice course on voice quality. The study was not intended to evaluate therapy efficacy, but to provide preliminary evaluation of the effect of such a voice training scheme on the voices of people who participate. Results suggested that voice quality improved after training. This finding was consistent using most acoustic parameters, but only marginal for several perceptual parameters. The adventitious effect of voice therapy on voice quality was previously documented using various methodologies, and thus, voice therapy is considered an effective approach for treating voice disorders (1,7,14,15). In that respect, our findings support earlier studies. Yet, the present findings provide preliminary indication that a voice course, which is conducted in a group setting, and is practiced by a heterogeneous group, has also a favorable effect on voice, similar to the effect of conventional voice therapy.

The present study included two groups of participants. The first group had neither vocal complaints nor laryngeal pathology, while all members of the

Table III. Mean values and standard deviations (in parentheses) of perceptual evaluation performed by the listeners group, for pitch, roughness, strain, breathiness, stability and resonance of the normal and pathological groups for the vowels /a/, /i/, obtained before and after the voice-course.

		Normal		Pathological	
Subjective measure	Training	/a/	/i/	/a/	/i/
Pitch	Before	2.09 (0.89)	2.18 (0.78)	2.87 (1.10)	3.01 (0.72)
	After	2.24 (0.74)	2.03 (0.54)	2.73 (0.90)	3.09 (1.24)
Roughness	Before	2.88 (1.01)	2.69 (0.58)	3.20 (1.04)	3.26 (0.55)
	After	2.59 (0.53)	2.85 (0.75)	2.89 (.99)	2.97 (1.02)
Strain	Before	2.43 (0.50)	2.59 (0.53)	2.99 (0.41)	3.57 (0.73)
	After	2.43 (0.58)	2.60 (0.34)	2.60 (0.34)	2.91 (0.79)
Breathiness	Before	2.29 (0.96)	2.31 (0.68)	2.71 (0.98)	2.99 (0.99)
	After	1.90 (0.43)	2.01 (0.49)	2.37 (0.59)	2.40 (0.51)
Stability	Before	3.11 (1.00)	2.89 (0.65)	3.29 (1.09)	3.61 (0.55)
	After	2.59 (0.48)	2.79 (0.63)	2.76 (0.90)	2.91 (0.67)
Resonance	Before	3.35 (1.07)	3.41 (0.60)	3.84 (1.09)	4.14 (0.57)
	After	2.94 (0.63)	3.10 (0.66)	3.31 (1.13)	3.56 (0.89)

pathological group had laryngeal abnormalities. Although both groups exhibited improvement in voice quality following the voice course, results indicated that the pathological group benefited from the voice course more than the normal group. This finding was demonstrated by the significant Training × Group interaction observed for the frequency- and amplitude-perturbation measures. These interactions are attributed mainly to fact that the pathological group (which showed high values prior to the voice course) improved markedly after training, approaching acoustic characteristics typical of healthy voice. Conversely, the normal group, which also improved their voices after therapy, presented smaller magnitudes of change, because their voices were within normal range even at the initial recording. Hence, only limited vocal improvement could be achieved by this group. On the other hand, the pathological group had voice quality that was more deviant from normal characteristics, providing more potential for improvement. Evidently, both groups reached similar acoustic values after therapy, which were well within normal range for voice quality.

Interestingly, while the two perturbation measures presented a consistent trend for group differences and for training effect, the two noise indices reacted differently. In both experimental groups, the NHR measure lowered after therapy, thus no significant Group × Training interaction was found. In contrast, the VTI measure revealed a significant Group × Training interaction, due to a decrease in values in the pathological group and an increase in the normal group. Yet, no significant overall group difference was observed for this parameter. These results can be interpreted to show that, within the context of the present study, the two noise indices were not effective in demonstrating group or training effects. Furthermore, in the pathological group, acoustic values of the two noise indices were within normal range, even at the first recording.

Three possible explanations may be given for this lack of significant improvement in the values of the noise indices within the pathological group. First, none of the participants in the pathological group had a severe voice disorder. Thus, it is possible that, within this group, a 'floor effect' was shown by the noise indices. Hence, since both NHR and VTI measures were within normal range before training, these measurements had only limited potential for improvement following therapy. A second possible explanation is that the voice training that was practiced here, improved mainly vocal stability, as reflected by the lowered perturbation values. It should be noted here, though, that while jitter and shimmer values improved after treatment, the perceptual measure of stability did not show a significant training effect. While acoustic analyses of frequency- and amplitude-perturbation relate to cycle-to-cycle variability in the acoustic signal, it should be clarified that the perception of stability probably relates to a different acoustic feature not tested here. It is conceivable, then, that the therapeutic approach advocated here had only little effect on glottal closure efficiency, as reflected by the noise indices. A third explanation for this result can be attributed to the fact that our acoustic analyses were based on vowels produced in isolation. It is possible that noise indices would demonstrate group or training effect more readily when analyzing longer and more natural voice samples, such as sentences or continuous speech (28). Our current data are not sufficient for providing a definite answer to this question. Due to the preliminary nature of this study, it was decided to analyze only isolated vowels. Obviously, further research on the effect of voice treatment in general and voice course in specific, is required to address this issue more thoroughly.

The second question addressed in this study was whether acoustic and perceptual evaluation of voice would illustrate the effect of a voice course similarly. To that end, voice samples that were recorded before and after training were, first, analyzed acoustically and, then, rated by a group of experienced listeners. Results indicated that the acoustic analysis paradigm identified group differences as well as training effect more consistently than the perceptual paradigm. Specifically, within the context of this study, group differences as well as training effect were observed in both evaluation paradigms. However, whereas several acoustic measures showed significant effects, only marginal effects were observed for a limited set of the perceptual measures. Furthermore, Training × Group interaction was found using both perturbation measures, but no such interaction was found using any of the subjective measures. This result stresses the value of integrating acoustic analyses of voice into the routine clinical voice evaluation, prior, during and after voice therapy. This conclusion is reminiscent of previous studies which demonstrated the validity of acoustic analysis of voice quality (4,13,21). Yet, our findings suggest that in some cases, acoustic analysis could reveal clinical changes in voice, before they are noticed perceptually. This asserts acoustic analysis as a valuable clinical tool and not merely as support for perceptual evaluation. The refined sensitivity of the acoustic analyses could be clinically valuable during the process of voice therapy, where a positive feedback of success or improvement could enhance treatment effectiveness.

Finally, it was noted that overall group differences were observed, in the acoustic paradigm, only using the jitter measure. A similar group difference was observed, in the perceptual paradigm, using the 'pitch' and 'strain' scales. Although it is possible that the added impact of the acoustic correlates of pitch and strain could result in frequency-perturbation increase, it seems, currently, premature to speculate on this relationship. Suffice it to say that this similarity between the acoustic and perceptual findings enhances the incentive for further exploration of this relationship, in an endeavor to improve our understanding of voice production and quality.

Two caveats of this study should be noted. First, due to the preliminary nature of this study, all participants underwent a laryngeal exam prior to enrolling in the voice course. However, because of administrative limitations, most of them were not re-examined after conclusion of the voice course. Consequently, it was impossible to compare laryngeal findings observed prior to therapy with later findings, and to correlate these observations with our acoustic and perceptual evaluations. Future studies should include laryngeal exams before and after voice courses. Second, treatment effect was evaluated in this study using two sets of measures: acoustic and perceptual. Additional information on treatment effect could be obtained from self-evaluation of the participants themselves pre- and post-therapy, and possibly several months post-therapy. Such a study that would also increase sample size could provide additional valuable information on the effect and efficacy of voice courses.

Conclusion

Results of the present study suggest three major findings. First, this study provides preliminary indications that a voice course could improve voice quality of people with and without voice disorders, similar to the effect of conventional voice therapy. Second, participants of the voice course, who were diagnosed with laryngeal pathologies, were shown to benefit more from the training than those with the healthy larynges. Third, acoustic analysis of voice quality revealed group differences and training effect even before these differences were identified by the listeners. It is suggested that while perceptual evaluation of voice quality is rightfully considered the gold standard for voice evaluation, the inclusion of acoustic analyses could provide important supplemental information on voice quality, in clinical as well as in research settings.

References

1. Ramig LO, Verdolini K. Treatment efficacy: Voice disorders. J Speech Lang Hear Res. 1988;41:S101–16.
2. Muchnik C, Hertzberg O, Mekai E, Shabtai E, Yedwab D, Ezrati R, et al. A community project for prevention of voice disorders among teachers: 1996–1999 [Hebrew]. Seeing Voices. 2001;1:66–72.
3. McCabe DJ, Titze IR. Chant therapy for treating vocal fatigue among public school teachers: A preliminary study. Am J Speech Lang Pathol. 2002;11:356–69.
4. Fex B, Fex S, Shiromoto O, Hirano M. Acoustic analysis of functional dysphonia: Before and after voice therapy (accent method). J voice. 1994;8:163–7.
5. Roy N, Gray SD, Simon M, Dove H, Corbin-Lewis K, Stemple JC. An evaluation of the effects of two treatment approaches for teachers with voice disorders: A prospective randomized clinical trial. J Speech Lang Hear Res. 2001;44: 286–96.
6. Roy N, Weinrich B, Gray SD, Tanner K, Toledo SW, Dove H, et al. Voice Amplification versus vocal hygiene instruction for teachers with voice disorders: A treatment outcome study. J Speech Lang Hear Res. 2002;45:625–38.
7. Carding PN, Horsley IA, Docherty GJ. A study of the effectiveness of voice therapy in the treatment of 45 patients with nonorganic dysphonia. J Voice. 1999;13:72–104.
8. Chan RWK. Does the voice improve with vocal hygiene education? A study of some instrumental voice measures in a group of kindergarten teachers. J Voice. 1994;8:279–91.
9. Broaddus-Lawrence PL, Treole K, McCabe, Allen RL, Toppin L. The effect of preventive vocal hygiene education on the vocal hygiene habits and perceptual vocal characteristics of training singers. J Voice. 2000;14:58–71.
10. Lehto L, Rantala L, Vilkman E, Alku P, Bäckström T. Experiences of short vocal training course for call-centre customer service advisors. Folia Phoniatr Logop. 2003;55: 163–76.
11. De Bodt MS, Wuyts FL, Van de Heyning PH, Croux C. Test-retest study of the GRBAS scale: Influence of experience and professional background on perceptual rating of voice quality. J Voice. 1997;11:74–80.
12. Hogikyan ND, Rosen CA. A review of outcome measurements for voice disorders. Otolaryngol Head Neck Surg. 2002;126:562–72.
13. Doyle P, Perkell JS, Hammarberg B, Hillman RE. Aerodynamic and acoustic voice measurements of patients with vocal nodules: Variations in baseline and changes across voice therapy. J Voice. 2003;17:269–82.
14. MacKenzie K, Millar A, Wilson JA, Sellars C, Deary I. Is voice therapy an effective treatment for dysphonia? A randomized controlled trial. BMJ. 2001;323:1–6.
15. Holmberg EB, Hillman RE, Hammarberg B, Södersten M, Doyle P. Efficacy of a behaviorally based voice therapy protocol for vocal nodules. J Voice. 2001;15:395–412.
16. Kreiman J, Gerratt BR, Kempster GB, Erman A, Berke GS. Perceptual evaluation of voice quality: Review, tutorial and a framework for future research. J Speech Hear Res. 1993;36: 21–40.
17. Kreiman J, Gerratt BR. Sources of listener disagreement in voice quality assessment. J Acoust Soc Am. 2000;108: 1867–76.
18. Wuyts FL, De Bodt MS, Van de Heyning PH. Is the reliability of a visual analog scale higher than an ordinal scale? An experiment with the GRBAS scale for the perceptual evaluation of dysphonia. J Voice. 1999;13:508–17.

19. Jacobsen BH, Johnson A, Geywalski C. The voice handicap Index (VHI): Development and validations. Am J Speech Lang Pathol. 1997;6:66–70.
20. Hogikyan ND, Sethuraman G. Validation of an instrument to measure voice-related quality of life (V-RQOL). J Voice. 1999;13:557–69.
21. Baken RJ. Clinical Measurement of Speech and Voice. London: Singular Publishing Group, Inc; 1996.
22. Sorenson D, Horii Y. Frequency and amplitude perturbation in the voice of female speakers. J Commun Dis. 1983;16: 57–61.
23. Sorenson D, Horii Y. Frequency characteristics of male and female speakers in the pulse register. J Commun Dis. 1984; 17:65–73.
24. Hixon TJ, Goldman MD, Mead J. Kinematics of the chest wall during speech production; Volume displacement of the rib cage, abdomen and lung. J Speech Hear Res. 1973;16: 78–115.
25. Boone DR, McFarlane SC. The Voice and Voice Therapy. Englewood Cliffs, New Jersey: Pranctice-Hall, Inc; 1988.
26. Weiss D, Beebe H. The Chewing Approach in Speech and Voice Therapy. Basel, Switzerland: S. Karger; 1951.
27. Most T, Amir O, Tobin Y. The Hebrew vowel system: Raw and normalized acoustic data. Lang Speech. 2000;43:295–308.
28. Halberstam B. Acoustic and perceptual parameters relating to connected speech are more reliable measures of hoarseness than parameters relating to sustained vowels. J Oto Rhino Laryngol. 2004;66:70–3.