Massed versus spaced practice in vocology: effect of a short-term intensive voice training versus a longer-term traditional voice training

Iris Meerschman[1], Kristiane Van Lierde[1,2], Caro Van Puyvelde[1], Astrid Bostyn[1], Sofie Claeys[3], Evelien D'haeseleer[1]

[1] Department of Speech Language and Hearing Sciences, Ghent University, Belgium
[2] Department of Speech-Language Pathology and Audiology, University of Pretoria, South Africa
[3] Department of Otorhinolaryngology, Ghent University, Belgium

ABSTRACT

Aims. The aim of this study was to compare the effect of a short-term intensive voice training with a longer-term traditional voice training on the vocal quality and vocal capacities of vocally healthy non-professional voice users.

Design. A pretest-posttest randomized control group design with follow-up measurements was used.

Methods & Procedures. Twenty healthy female non-professional voice users with a mean age of 21.7 years (range: 20-24 yrs.) were randomly assigned into a short-term intensive voice training group (IVT, n=10) or a longer-term traditional voice training group (TVT, n=10). Both groups received an identical 6-hour lasting voice training. Only the distribution of practice varied between groups: two hours a day for three consecutive days for the IVT group versus two 30-minute sessions a week for six weeks for the TVT group. In both groups, a voice assessment protocol consisting of subjective (questionnaire, participant's self-report, auditory-perceptual evaluation) and objective (maximum performance task, acoustic analysis, voice range profile, dysphonia severity index) measurements and determinations was used to evaluate the participants' voice pre- and post-training and at 6 weeks follow-up. Groups were compared over time using linear mixed models and generalized linear mixed models. Within-group effects of time were determined using post-hoc pairwise comparisons with Bonferroni corrections.

Outcomes & Results. No significant time-by-group interactions were found for any of the outcome measures, indicating no significant differences in evolution over time between the groups. Significant time effects were found for maximum phonation time, lowest intensity, lowest frequency, highest frequency, and dysphonia severity index, all improving over time in both groups. More in-depth within-group analyses indicate a preference for the IVT group regarding the evolution of maximum phonation time, lowest frequency and dysphonia severity index, and a preference for the TVT group regarding the evolution of lowest intensity.

Conclusions & Implications. A short-term intensive voice training may be equally, or even more, effective in training vocally healthy non-professional voice users compared with a longer-term traditional voice training.

INTRODUCTION

Voice therapy and voice training are processes of behavioral change (Behrman, 2006; Van Leer, Hapner and Connor, 2008; Mcllwaine, Madill and McCabe, 2010; Patel, Bless and Thibeault, 2011; Vinney and Turkstra, 2013; Wenke et al., 2014; Behlau, Madazio and Oliveira, 2015; Iwarsson, 2015; Fu, Theodoros and Ward, 2015a; Fu, Theodoros and Ward, 2015b). They involve the acquisition, optimization and maintenance of healthy and efficient vocal behaviors through (re)learning cognitive and motor skills (Mcllwaine, Madill and McCabe, 2010; Patel, Bless and Thibeault, 2011; Fu, Theodoros and Ward, 2015a). Principles inherent to behavioral change (learning) are well known from the fields of neurobiology, exercise physiology, motor learning, psychology, and language therapy (Patel, Bless and Thibeault, 2011). Nevertheless, limited research has been devoted to explore how these principles apply to voice therapy or training (Mcllwaine, Madill and McCabe, 2010; Wenke et al., 2014).

Recently, increased attention has been paid to the principle “distribution” of practice (Patel, Bless and Thibeault, 2011; Wenke et al., 2014; Fu, Theodoros and Ward, 2015a). In motor learning, practice distribution may be categorized as “massed” versus “spaced” practice. In massed practice, all practice sessions occur very closely together with little or no rest time between sessions. In spaced practice, the time interval between practice sessions is larger (Bergan, 2010). Practice sessions in vocology are traditionally organized according a spaced practice schedule with weekly sessions spread over several weeks to months (Carding, Horsley and Docherty, 1999; Chen et al., 2007; Fischer, Gutenbrunnera and Ptokb, 2009; Bergan, 2010; Demmink-Geertman and Dejonckere, 2010). A literature overview of De Bodt, Patteeuw and Versele (2015) between 1975 and May 2013 showed that voice therapy lasts an average of 9.25 weeks distributed over 10.87 sessions of mostly 30 or 60 minutes and occurs once or twice a week, although substantial geographical differences were observed.

In contrast with most medical and pharmaceutical therapies, the optimal dosage for voice therapy or training is unknown (De Bodt, Patteeuw and Versele, 2015; Roy, 2012). The exact frequency and duration used today depends on several factors, such as the medical prescription, rules of reimbursement, the specific vocal pathology and its severity, the type of therapy or training, the client's limitations and expectations, and upcoming vocal performances (Mueller and Larson, 1992; De Bodt et al., 2008; Van Lierde et al., 2007; Van Lierde et al., 2010a). Having standardized guidelines in terms of the ideal frequency and duration for voice therapy and training could be a merit for both the patient/client, the voice therapist/coach and the health care system (Patel, Bless and Thibeault, 2011; Wenke et al., 2014; De Bodt, Patteeuw and Versele, 2015).

Returning to the fields of neurobiology, exercise physiology, motor learning, psychology, and language therapy, there seems to be a general preference for high-intensity training (i.e. massed practice) to obtain desirable learning and behavioral changes (Patel, Bless and Thibeault, 2011). To date, evidence for a high-intensity approach in vocology is limited to few specific programs, such as the Lee Silverman Voice Treatment (LSVT®, Ramig et al., 1994) and Vocal Function Exercises (VFE, Stemple et al., 1994).

Although the preference for high-intensity training has not yet broken through our field, it recently gained interest by the concept article of Patel, Bless and Thibeault (2011). The authors developed a “boot camp” voice therapy, which is an innovative approach of concentrated practice, performed in a time frame of 1-4 consecutive days with 4-7 hours of therapy a day. In addition to the high-intensity principle, the “boot camp” therapy is also based on principles of “variability” and “specificity” of training, which may positively contribute to transfer and carryover. A variety of voice therapy techniques are given by a large number of clinicians (3-7) and therapy is tailored to the nature of the voice disturbance and the individual’s specific needs. It is designed for people who have pressing needs to improve their voice (e.g. upcoming vocal performances), who failed traditional voice therapy (e.g. recalcitrant dysphonia), and/or have an inability to schedule weekly appointments (e.g. living at geographical distances far from a voice center). Behlau et al. (2014) mentioned the use of a similar intensive short-term voice therapy in Brazil for a variety of cases, including patients with iatrogenic dysphonia and professional voice users suffering from acute dysphonia. The therapy lasts 3 days to 2 weeks, with 3 to 4 sessions a day, and 2 to 4 speech-language pathologists.

Clinical trials comparing the effect of an intensive versus a traditional voice therapy are still in its infancy. Fu, Theodoros and Ward (2015a) found comparable positive perceptual, physiological, and acoustic outcomes for both models in patients with vocal nodules (eight 45-minute sessions over 3 weeks versus eight 45-minute sessions over 8 weeks). Limitations of this study are the lack of long-term follow-up and self-rating questionnaires. Furthermore, a pragmatic randomized controlled trial (RCT) was used instead of an explanatory RCT in which subjects were assigned to either of two treatment groups according to their availability. Wenke et al. (2014) found high satisfaction and a significantly reduced Voice Handicap Index (VHI) after an intensive treatment (four 1-hour treatment sessions a week over 2 weeks) in patients with functional dysphonia. A general trend of improved mean VHI ratings was found in the standard group (one 1-hour treatment session a week over 8 weeks) as well, although this improvement was not significant. Moreover, significantly higher attendance rates were found in the intensive group compared to the group receiving the standard therapy. A major limitation of the study is that the therapy program was not standardized (i.e. subjects received different treatment techniques depending on the individual's profile), which means that it is not clear whether the treatment success was related to the type of techniques or the distribution of practice. Furthermore, perceptual and objective vocal measures were missing. Fischer, Gutenbrunnera and Ptokb (2009) investigated the effect of a 2-week intensive voice therapy combined with elements of physical medicine (physiotherapy, manual therapy, inhalations, vibration massage etc.) in patients with chronic functional or organic dysphonia. The authors found a significantly reduced overall voice handicap in patients with moderate baseline voice handicap values, whereas no significant changes could be detected in patients with severe handicap. Because voice therapy was combined with physical therapies, the effect of intensive voice therapy alone cannot be concluded. Furthermore, the superiority of a more intensive schedule was postulated without an actual comparison with the traditional model. To our knowledge, no studies compared an intensive with a traditional voice training in healthy subjects.

Possible advantages of a high-intensity approach in vocology are creating a greater opportunity to practice, giving the ability to focus entirely on improving vocal behavior, and obtaining a better simulation of cognitive, motor, and physiological requirements of daily communication (Patel, Bless and Thibeault, 2011). These factors may in turn improve transfer of learned skills, and increase or regain client's motivation and compliance (Patel, Bless and Thibeault, 2011; Wenke et al., 2014; Fu, Theodoros and Ward, 2015a). Motivation and compliance are essential for behavioral change and are often poor in the traditional model of voice therapy (Behrman, 2006; Patel, Bless and Thibeault, 2011; Wenke et al., 2014; Fu, Theodoros and Ward, 2015b), which may lead to emotional frustration for clinicians, a negative impact on the client's vocal outcome and reduced cost efficiency for health care services (Wenke et al., 2014).

Estimating the optimal dosage for therapy and training is an unsolved challenge in the field of vocology, particularly due to several influencing factors such as severity of the voice disturbance, and motivation and expectations of the patient or client. Fact remains that a general picture of the most effective and efficient frequency and duration of voice therapy and training is essential (Patel, Bless and Thibeault, 2011; Wenke et al., 2014; De Bodt, Patteeuw and Versele, 2015). This study aims to explore the motor learning principle “distribution of practice” in our field. Therefore, two extreme “dosages” of voice training were compared using a study group of vocally healthy non-professional voice users. Every voice user, also a vocally healthy individual, is able to change his or her vocal behavior, and learn efficient and healthy voice use. Furthermore, the exact same vocal techniques can be used for both training as therapy, which makes this study population suitable for a preliminary exploration. At last, a stronger study design with randomization of the groups, a better control of influencing factors, and standardization of the training program is possible in a healthy study group.

The aim of this study was to compare the effect of a short-term intensive voice training (two hours a day for three consecutive days) with a longer-term traditional voice training (two 30-minute sessions a week for six weeks) on the vocal quality and vocal capacities of vocally healthy non-professional voice users. Based on the principles of behavioral change and the previously mentioned possible advantages of high-intensity training, it was hypothesized that a short-term intensive voice training may be equally, or even more, effective than a longer-term traditional voice training.

METHODS

This study was approved by the Ethics Committee of Ghent University Hospital (registration number: B670201422095).

Participants

Twenty young and healthy female participants with a mean age of 21.7 years (SD: 0.8 yrs., range: 20-24 yrs.) participated in the study. Recruitment was based on convenience sampling. None of the participants reported hearing problems or voice problems. Fifteen subjects were students (studies: social work and social welfare, political sciences, international relations and diplomacy, law school (2), nursing, medicine (3), rehabilitation sciences and physiotherapy, educational sciences, linguistics and literature, multilingual professional communication, sociology, applied economic sciences) and 5 subjects were employed (nurse, midwife, process operator, pedagogue, sales manager). None of them were professional voice users. All participants provided written informed consent at an initial briefing. They were randomly assigned into two groups: an experimental group (n = 10) receiving the intensive short-term voice training (IVT, two hours a day for three consecutive days), and a control group (n = 10) receiving a longer-term traditional voice training (TVT, two 30-minute sessions a week for six weeks). There were no differences between the two groups in mean age (Mann–Whitney U test: $p = 0.108$). Only women were recruited to avoid an unequal distribution of sex due to the small sample size and randomization procedure.

Voice assessment

An identical voice assessment protocol was used to evaluate the participants' voice pre- and post-training and at 6 weeks follow-up. Data were collected in a sound-treated room at Ghent University Hospital. The voice assessment protocol included both subjective (questionnaire, participant's self-report and auditory-perceptual evaluations) and objective (maximum performance task, acoustic analysis, Voice Range Profile, Dysphonia Severity Index) vocal measurements and determinations.

Questionnaire voice-related symptoms, risk factors, vocal load, and lifestyle habits. A questionnaire based on the checklists of Russel, Oates, & Permberton (2000), De Bodt, Mertens, & Heylen (2008), and Van Lierde et al. (2010b, 2010c) was presented at the pretest to explore voice-related symptoms, risk factors, vocal load, and lifestyle habits and to confirm the success of randomization. The presence of vocal complaints and upper respiratory tract infections was rechecked at the posttest and at 6 weeks follow-up.

Participant's self-report. The Voice Handicap Index (Jacobson et al., 1997; Dutch version: Belgian Study Group on Voice Disorders, De Bodt et al., 2000) was used to evaluate the psychosocial impact of potential voice problems. The VHI is a self-administered questionnaire consisting of 30 statements, evaluating functional (10 statements, F-scale), physical (10 statements, P-scale), and emotional (10 statements, E-scale) restrictions. Each statement was scored on a 5-

point scale (0: never, 1: almost never, 2: sometimes; 3: almost always; 4: always). The total VHI-score varies between 0 and 120; the higher the score, the more severe is the psychosocial impact.

Auditory-perceptual evaluation. Voice samples of a sustained vowel /a/ and connected speech (reading aloud the phonetically balanced text “De noordenwind en de zon”) were recorded for the auditory perceptual evaluation using a digital camera with high quality microphone (Sanyo VPC-HD200). The parameters Grade, Roughness, Breathiness, Asthenia, Strain, and Instability were evaluated using the 0-3 intensity score (0: absent, 1: mild, 2: moderate, 3: severe) of the GRBASI-scale (Hirano, 1981; completed with an "I" parameter by Dejonckere et al., 1996). Samples were randomized and rated blinded by the same voice therapist (I.M.). To assure inter-rater reliability, twenty samples (33.3%) were randomly selected and rated blinded and independently by another voice therapist (E.D.).

Maximum performance task. To measure the maximum phonation time (MPT, in s), participants were asked to sustain the vowel /a/ at their habitual pitch and loudness after a maximal inspiration, in free field while seated. The MPT was modeled by the experimenters and the participants received visual and verbal encouragements to produce the longest possible sample. The length of the sustained vowel was measured with a chronometer. The best trial of three attempts was retained for further analysis.

Acoustic analysis. The fundamental frequency (f_o, in Hz), jitter (in %), shimmer (in %), variation in f_0 (vf_o, in %) and noise-to-harmonic ratio (NHR) were obtained by the Multi Dimensional Voice Program of the Computerized Speech Lab (CSL; model 4500, KayPENTAX. Montvale, NY), using a Shure SM-48 microphone located at a distance of 15 cm from the mouth and angled at 45°. The subjects were instructed to produce the vowel /a/ at their habitual pitch and loudness. A midvowel segment of 3 seconds registered with a sampling rate of 50 kHz was used.

Voice Range Profile (VRP). The VRP was determined using the CSL, following the procedure outlined by Heylen et al. (1998). This assessment includes determination of the highest and the lowest fundamental frequency (F-high, F-low) and intensity (I-high, I-low). Participants were instructed to produce the vowel /a/ for at least 2 seconds using, respectively, a habitual pitch and loudness, a minimal pitch, a minimal intensity, a maximal pitch, and a maximal intensity. Each production was modeled by the experimenters and the participants received visual and verbal encouragement.

Dysphonia Severity Index (DSI). The DSI is a multiparameter approach designed to establish an objective and quantitative correlate of the perceived vocal quality (Wuyts et al., 2000). The DSI is based on a weighted combination of the following parameters: MPT (in s), highest frequency (F-high, in Hz), lowest intensity (I-low, in dB) and jitter (in %). The DSI is constructed as 0.13 MPT + 0.0053 F-high − 0.26 I-low − 1.18 jitter + 12.4. The index ranges from -5 to +5 for severely dysphonic to normal voices. The more negative the index, the worse is the vocal quality. Values higher than 5 are possible in subjects with very good vocal capacities. A DSI of 1.6 is the threshold separating normal voices from dysphonic voices (Raes et al., 2002).

Voice training

Both the IVT group and the TVT group received an identical 6-hours lasting voice training. Only the distribution of practice varied between groups: two hours a day for three consecutive days for the IVT group versus two 30-minute sessions a week for six weeks for the TVT group. The training program included counseling and vocal hygiene (30 minutes), posture and relaxation (30 minutes), respiration (1 hour), humming and resonant voice (1 hour), voice placing and forward focus (30 minutes), pitch and loudness control (30 minutes), vocal function exercises (30 minutes), voice onset (30 minutes), generalization and transfer (1 hour). Details of the training program are provided in Table 1 (De Bodt, Mertens and Heylen, 2008; De Bodt et al., 2008; Timmermans, 2008; Verdolini-Marston et al., 1995; Verdolini, 2000; Stemple, 1994).

Table 1: Content of the voice training program

Counseling and vocal hygiene (30 minutes)	Counseling: - explaining the anatomy and functioning of the larynx using simple educational images - clarifying the distinction between normal and pathological voices Vocal hygiene: - checking vocal abuse, vocal load, and influencing lifestyle habits using a questionnaire - discussing feasible solutions and general advice concerning vocal hygiene
Posture and relaxation (30 minutes)	Posture: - highlighting the importance of a correct posture for phonation - demonstrating a correct posture while standing and sitting

	- applying specific exercises to stimulate a correct posture (e.g. standing upright with your feet slightly apart, the knees relaxed, the pelvis balanced, lift one arm and then the other with the palm facing upward, pull the arms pretending to push the sky above with alternating hands) Relaxation: - performing localized relaxation techniques: head, neck, shoulders, larynx, and pharynx (e.g. moving the head sideways as much as possible so that the ear almost touches the shoulder; lifting the shoulders as high as possible without movement of the back or trunk for a few seconds, then slowly lower the shoulders; pretending to drink out of cupped hands with deep inhalations; introducing a yawn while feeling a slight tension in the palate, lowering of the larynx and widening of the pharynx)
Respiration (1 hour)	- highlighting the importance of an efficient respiration type for phonation - discussing and demonstrating the different respiration types (clavicular, costal, costo-abdominal, abdominal) - advancing awareness of the subject's habitual respiration type and adjusting to a costo-abdominal type while laying, sitting and standing; using tactile-kinesthetic and visual feedback - practicing the costo-abdominal type and respiratory control on different hierarchical levels: inhaling through the nose and exhaling while producing voiceless fricatives ([f] and [s]), voiced fricatives ([v] and [z]), other consonants and vowels, words, automatic sequences, sentences, and texts
Humming and resonant voice (1 hour)	- explaining the physiology and the purpose of resonant voice exercises - sensing "easy" phonation and vibrations in the midfacial region while humming on [m], [n], [ŋ] - practicing resonant voice exercises on different hierarchical levels (isolated, syllable, word, phrase, sentence, text) using tactile-kinesthetic and auditory feedback - reducing the degree of resonance while maintaining the "easy phonation" with forward focus
Voice placing and forward focus (30 minutes)	- highlighting the importance of removing the energy and muscle tension away from the larynx and bringing it to the mouth ("mask resonance") - highlighting the importance of transferring the message to the listener ("forward focus") - specific exercises using visual, auditory and tactile-kinesthetic feedback: gawking to reduce muscle tension in the cheeks and neck, humming to place the voice, using an imaginary megaphone to stimulate forward focus, "bringing" the voice to the nose, sighing, speaking while "throwing" away words like darts to a dartboard, using open and exaggerated articulation etc. (selection was adjusted to the participant, avoiding excessive muscle tension)
Pitch and loudness control (30 minutes)	- ascending and descending pitch glides - crescendo and decrescendo
Vocal function exercises (30 minutes)	Vocal Function Exercises - warm-up: sustaining the vowel [i] as long as possible on the musical note F above middle C - stretching: upward pitch glide on [o] - contracting: downward pitch glide on [o] - adductory power: sustaining the vowel [o] as long as possible on the musical notes C-D-E-F-G
Voice onset (30 minutes)	- discussing and demonstrating the different types of voice onset (hard, aspirated/soft, balanced) - practicing a balanced voice onset starting from an aspirated/soft onset: a) blowing air through pursed lips, followed by a rounded vowel or diphthong, gradually reducing the blowing

	b) producing words with a vowel or diphthong at medial position, inserting a [h] sound between the vowel/diphthong, gradually reducing the [h] production c) producing words with a vowel or diphthong at initial position, adding a [h] sound before the vowel/diphthong, gradually reducing the [h] production d) practicing sentence and text level
Generalization and transfer (1 hour)	generalization of the learned techniques during reading aloud and spontaneous speech; using auditory, visual and tactile-kinesthetic feedback

Statistical analysis

SPSS version 24 (SPSS Corporation, Chicago, IL, USA) was used for the statistical analysis of the data. Analyses were conducted at $\alpha = 0.05$.

Fisher's Exact tests were used to compare the groups regarding self-reported voice-related symptoms, risk factors, vocal abuse and lifestyle habits, and to confirm the success of randomization.

Cohen's κ was run to determine the inter-rater reliability for the auditory-perceptual evaluation (GRBASI).

Linear mixed models were used to compare groups over time on each continuous outcome measure, using the restricted maximum likelihood estimation and scaled identity covariance structure. Time, group, and time-by-group interaction were specified as fixed factors. A random intercept for subjects was included. Model assumptions were checked by inspecting whether residuals were normally distributed. Generalized linear mixed models were used for the categorical outcome measures. If a significant main (time*group, time, or group) effect was found, within-group effects of time were determined using pairwise comparisons with Bonferroni corrections (pretraining versus posttraining, posttraining versus 6 weeks follow-up, pretraining versus 6 weeks follow-up).

RESULTS

Questionnaire voice-related symptoms, risk factors, vocal abuse, and lifestyle habits

Results on the questionnaire regarding voice-related symptoms, risk factors, vocal abuse and lifestyle habits are presented in Table 2. Fischer's exact tests showed no significant baseline differences between the two groups. The presence of vocal complaints and upper respiratory tract infections did not differ between groups when rechecked at the posttest and at 6 weeks follow-up.

Table 2: Presence of voice-related symptoms, risk factors, vocal abuse, and lifestyle habits in the IVT group and the TVT group.

		IVT (n = 10)	**TVT (n = 10)**	**p-value**
Vocal complaints	pretest	2	1	> 0.999
	posttest	3	0	0.211
	6 weeks follow-up	1	4	0.303
Upper respiratory tract infection	pretest	2	4	0.628
	posttest	5	4	> 0.999
	6 weeks follow-up	4	7	0.370
Allergy	pretest	4	3	> 0.999
Reflux	pretest	3	1	0.582
Vocal abuse	pretest	4	8	0.170
Smoking	pretest	4	0	0.087
Alcohol use	pretest	10	10	> 0.999
Coffee	pretest	6	4	0.656

Inter-rater reliability auditory-perceptual analysis

Cohen's κ showed moderate to excellent degrees of inter-rater reliability for the GRBASI parameters. An excellent degree of reliability was found for the parameters G, B, A and I with κ

= 0.77, κ = 0.86, κ = 1.00, and κ = 1.00 respectively. A moderate degree of reliability was found for the parameters R and S, with κ = 0.50.

Evolution outcome measures

Evolution of the outcome measures in both groups are presented in Tables 3 and 4. (Generalized) linear mixed models showed no significant time-by-group interactions for any of the outcome measures, indicating no significant differences in evolution over time between both groups. A significant group effect was found for MPT ($F(1,18) = 5.423$, $p = 0.032$), indicating a significant difference among groups independent of time. Significant time effects were found for MPT ($F(2,36) = 11.990$, $p < 0.001$), I-low ($F(2,36) = 6.091$, $p = 0.005$), F-low ($F(2,36) = 5.667$, $p = 0.007$), F-high ($F(2,36) = 14,456$, $p < 0.001$), and DSI ($F(2,36) = 11.785$, $p < 0.001$), indicating significant changes over time in the sample as a whole, independent of group assignment. All these measures improved over time (MPT: Figure 1, I-low: Figure 2, F-low: Figure 3, F-high: Figure 4, DSI: Figure 5).

Within-group effects of time showed a significant improvement in MPT pre- to post-training in the IVT group (+5.3s, $p = 0.005$); MPT also improved pre- to post-training in the TVT group although not significantly (+3.5s, $p = 0.090$). MPT did however significantly improve pretraining to 6 weeks follow-up in both groups (IVT: +4.4s, $p = 0.022$, TVT: +5.3s, $p = 0.005$), and improved MPTs posttraining remained until 6 weeks follow-up in both groups (post – 6 weeks follow-up, $p > 0.05$). I-low significantly improved pretraining to 6 weeks follow-up in the TVT group (- 2.1dB, $p = 0.023$). F-low significantly improved pre- to post-training in the IVT group (-10.4Hz, $p = 0.015$), and improvement remained until 6 weeks follow-up (post – 6 weeks follow-up, $p > 0.05$). F-high significantly improved pre- to post-training in both groups (IVT: +194.3 Hz, $p = 0.015$; TVT: +212.7 Hz, $p = 0.007$), and improvements remained until 6 weeks follow-up (post – 6 weeks follow-up, $p > 0.05$). DSI significantly improved pre- to post-training in the IVT group (+2.1, $p = 0.025$); DSI also improved pre- to post-training in the TVT group although not significantly (+ 1.8, $p = 0.055$). DSI did however significantly improve pretraining to 6 weeks follow- up in both groups (IVT: +2.3, $p = 0.016$; TVT: +2.5, $p = 0.004$), and improved DSI scores posttraining remained until 6 weeks follow-up in both groups (post – 6 weeks follow-up, $p > 0.05$).

Table 3: Evolution of the categorical outcome measures in the IVT group and the TVT group.

		Pretraining	Posttraining	6 weeks follow-up	Time*Group	Group	Time
Parameters	Group	Median (IQR)	Median (IQR)	Median (IQR)	p-value	p-value	p-value
Auditory-perceptual evaluation							
G	**IVT**	0 (0 – 0.75)	0 (0 – 0)	0 (0 – 0)	0.905	0.856	0.597
	TVT	0 (0 – 1)	0 (0 – 0.5)	0 (0 – 0.25)			
R	**IVT**	0.5 (0 – 1)	0 (0 – 1)	1 (0 – 1)	0.434	0.466	0.782
	TVT	0 (0 – 1)	0 (0 – 1)	0 (0 – 1)			
B	**IVT**	0 (0 – 0)	0 (0 – 0)	0 (0 – 0)	>0.999	0.995	>0.999
	TVT	0 (0 – 0.25)	0 (0 – 1)	0 (0 – 1)			
A	**IVT**	0 (0 – 0)	0 (0 – 0)	0 (0 – 0)	>0.999	0.997	>0.999
	TVT	0 (0 – 0)	0 (0 – 0)	0 (0 – 0)			
S	**IVT**	0 (0 – 0.75)	0 (0 – 0)	0 (0 – 0)	0.798	0.993	0.836
	TVT	0 (0 – 0.25)	0 (0 – 0.5)	0 (0 – 0.25)			
I	**IVT**	0 (0 – 0)	0 (0 – 0)	0 (0 – 0)	0.769	0.658	0.769
	TVT	0 (0 – 0)	0 (0 – 0.5)	0 (0 – 0)			

Abbreviations: IVT, intensive voice training; TVT, traditional voice training; IQR, interquartile range; G, grade; R, roughness; B, breathiness; A, asthenia; S, strain; I, instability.

Table 4: Evolution of the continuous outcome measures in the IVT group and the TVT group.

					Time*Group	Group	Time	Comparison Time within groups		
		Pretraining	**Posttraining**	**6 weeks follow-up**				Pretraining - Posttraining	Posttraining – 6 weeks follow-up	Pretraining – 6 weeks follow-up
Parameters	Group	Estimated Mean (95% CI)	Estimated Mean (95% CI)	Estimated Mean (95% CI)	p-value	p-value	p-value	p-value	p-value	p-value
Maximum performance task										
MPT (s)	**IVT**	23.9 (19.9 – 28.0)	29.2 (25.2 – 33.3)	28.3 (24.3 – 32.4)	0.462	0.032*	<0.001*	0.005*	>0.999	0.022*
	TVT	18.4 (14.4 – 22.5)	21.9 (17.9 – 26.0)	23.7 (19.7 – 27.8)				0.090	0.757	0.005*
Acoustic analysis										
f_o (Hz)	**IVT**	191.5 (175.7 – 207.2)	193.8 (178.0 – 209.5)	190.7 (174.9 – 206.4)	0.197	0.264	0.508	-	-	-
	TVT	202.3 (186.5 – 218.0)	196.8 (181.1 – 212.5)	210.0 (194.3 – 225.8)				-	-	-
jitter (%)	**IVT**	1.423 (0.834 – 2.013)	1.493 (0.903 – 2.083)	1.479 (0.889 – 2.069)	0.720	0.631	0.752	-	-	-
	TVT	1.693 (1.103 – 2.283)	1.748 (1.158 – 2.338)	1.426 (0.836 – 2.015)				-	-	-
shimmer (%)	**IVT**	4.929 (4.085 – 5.772)	4.884 (4.040 – 5.728)	4.004 (3.160 – 4.848)	0.534	0.908	0.120	-	-	-
	TVT	4.609 (3.765 – 5.452)	4.918 (4.075 – 5.762)	4.446 (3.602 – 5.290)				-	-	-
vf_o (%)	**IVT**	1.503 (1.075 – 1.931)	1.583 (1.155 – 2.011)	1.504 (1.076 – 1.932)	0.611	0.960	0.530	-	-	-
	TVT	1.684 (1.256 – 2.112)	1.595 (1.167 – 2.023)	1.340 (0.912 – 1.768)				-	-	-
NHR	**IVT**	0.131 (0.117 – 0.144)	0.143 (0.129 – 0.156)	0.126 (0.112 – 0.139)	0.231	0.942	0.303	-	-	-
	TVT	0.127 (0.113 – 0.140)	0.134 (0.120 – 0.147)	0.137 (0.124 – 0.151)				-	-	-
Voice Range Profile										
I-low (dB)	**IVT**	59.6 (57.8 – 61.4)	58.0 (56.2 – 59.9)	58.2 (56.4 – 60.0)	0.524	0.561	0.005*	0.114	>0.999	0.203
	TVT	59.0 (57.2 – 60.9)	57.9 (56.1 – 59.8)	56.9 (55.1 – 58.8)				0.443	0.560	0.023*
I-high (dB)	**IVT**	104.8 (99.9 – 109.8)	107.6 (102.7 – 112.5)	108.3 (103.4 – 113.3)	0.877	0.496	0.077	-	-	-
	TVT	107.6 (102.7 – 112.6)	109.0 (104.0 – 114.0)	110.4 (105.5 – 115.4)				-	-	-
F-low (Hz)	**IVT**	147.6 (134.6 – 160.7)	137.2 (124.2 – 150.3)	138.7 (125.6 – 151.8)	0.232	0.274	0.007*	0.015*	>0.999	0.043*
	TVT	153.7 (140.6 – 166.7)	151.5 (138.5 – 164.6)	146.9 (133.9 – 160.0)				>0.999	0.585	0.186
F-high (Hz)	**IVT**	664.4 (496.8 – 832.1)	858.7 (691.1 – 1026.4)	915.6 (748.0 – 1083.3)	0.696	0.651	<0.001*	0.015*	>0.999	0.001*
	TVT	630.0 (462.4 – 797.64)	842.7 (675.0 – 1010.3)	824.2 (656.6 – 991.8)				0.007*	>0.999	0.015*
Dysphonia Severity Index										
DSI	**IVT**	1.9 (0.0 – 3.9)	4.0 (2.1 – 5.9)	4.2 (2.2 – 6.0)	0.838	0.385	<0.001*	0.025*	>0.999	0.016*
	TVT	0.9 (-1.0 – 2.8)	2.7 (0.8 – 4.6)	3.4 (1.5 – 5.3)				0.055	0.955	0.004*

Voice Handicap Index										
F-scale	**IVT**	3.9 (1.4 – 6.4)	2.9 (0.4 – 5.4)	2.7 (0.2 – 5.2)	0.483	0.385	0.054	-	-	-
	TVT	5.0 (2.5 – 7.5)	5.0 (2.5 – 7.4)	3.7 (1.2 – 6.2)				-	-	-
P-scale	**IVT**	5.7 (1.4 – 10.0)	5.2 (0.9 – 9.5)	4.7 (0.4 – 9.0)	0.988	0.236	0.478	-	-	-
	TVT	9.1 (4.8 – 13.4)	8.6 (4.3 – 12.9)	8.3 (4.0 – 12.6)				-	-	-
E-scale	**IVT**	0.9 (0 – 3.7)	1.1 (0 – 3.9)	1.1 (0 – 3.9)	0.491	0.090	0.601	-	-	-
	TVT	4.6 (1.8 – 7.4)	4.7 (1.9 – 7.5)	3.7 (0.9 – 6.5)				-	-	-
Total VHI	**IVT**	10.5 (1.9 – 19.1)	9.2 (0.6 – 17.8)	8.5 (0 – 17.1)	0.784	0.169	0.160	-	-	-
	TVT	18.7 (10.1 – 27.3)	18.2 (9.6 – 26.8)	15.7 (7.1 – 24.3)				-	-	-

Abbreviations: IVT, intensive voice training; TVT, traditional voice training; CI, confidence interval; MPT, maximum phonation time; I-low, lowest intensity; I-high, highest intensity; F-low, lowest frequency; F-high, highest frequency; f_o, fundamental frequency; vf_o, variation in fundamental frequency; NHR, noise-to-harmonic ratio; DSI, Dysphonia Severity Index; F-scale, functional scale; P-scale, physical scale; E-scale, emotional scale; VHI, Voice Handicap Index.

** indicates a significant effect.*

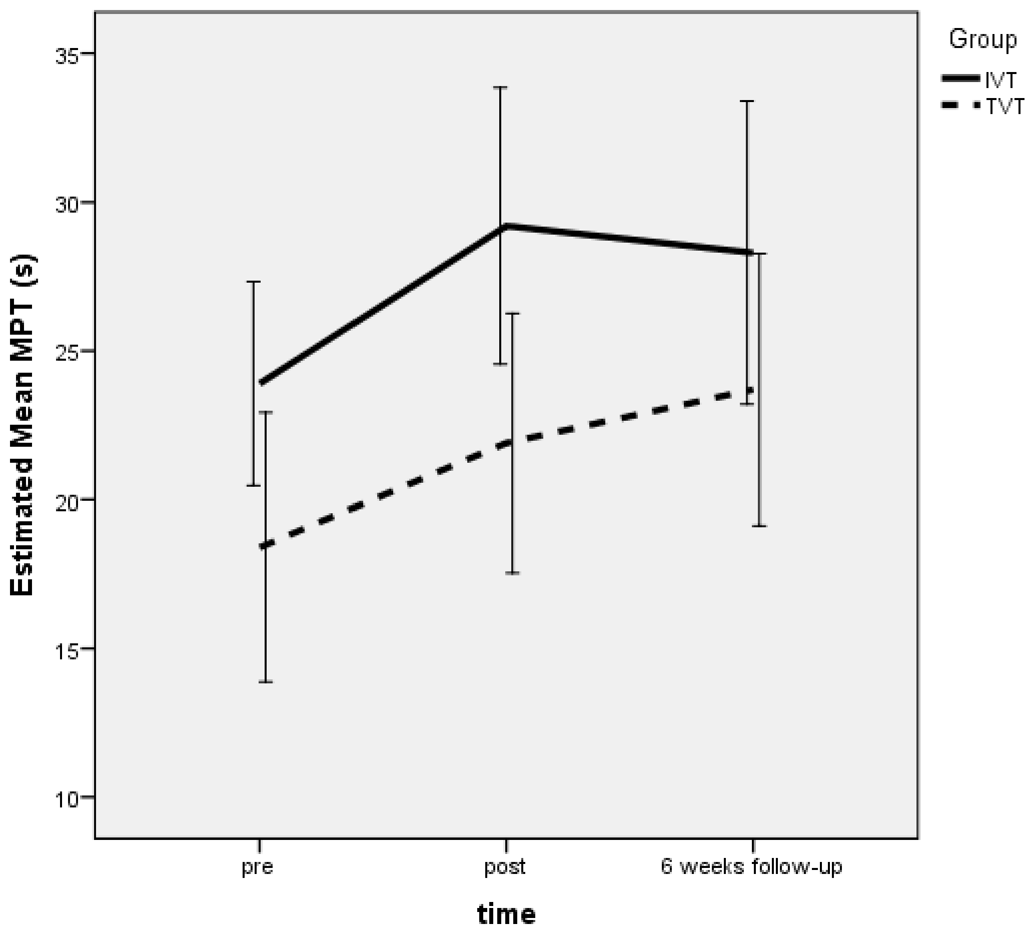

Figure 1. Evolution MPT (s) over time in the IVT and TVT groups.

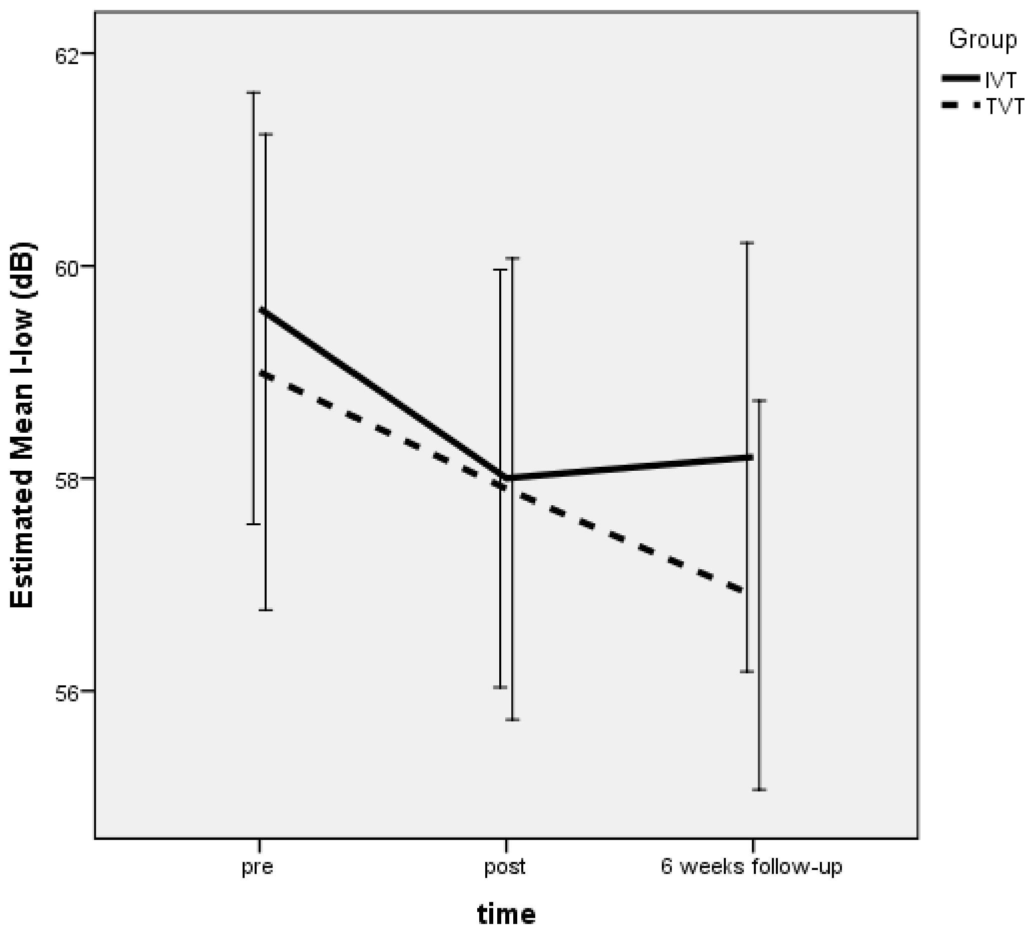

Figure 2. Evolution I-low (dB) over time in the IVT and TVT groups.

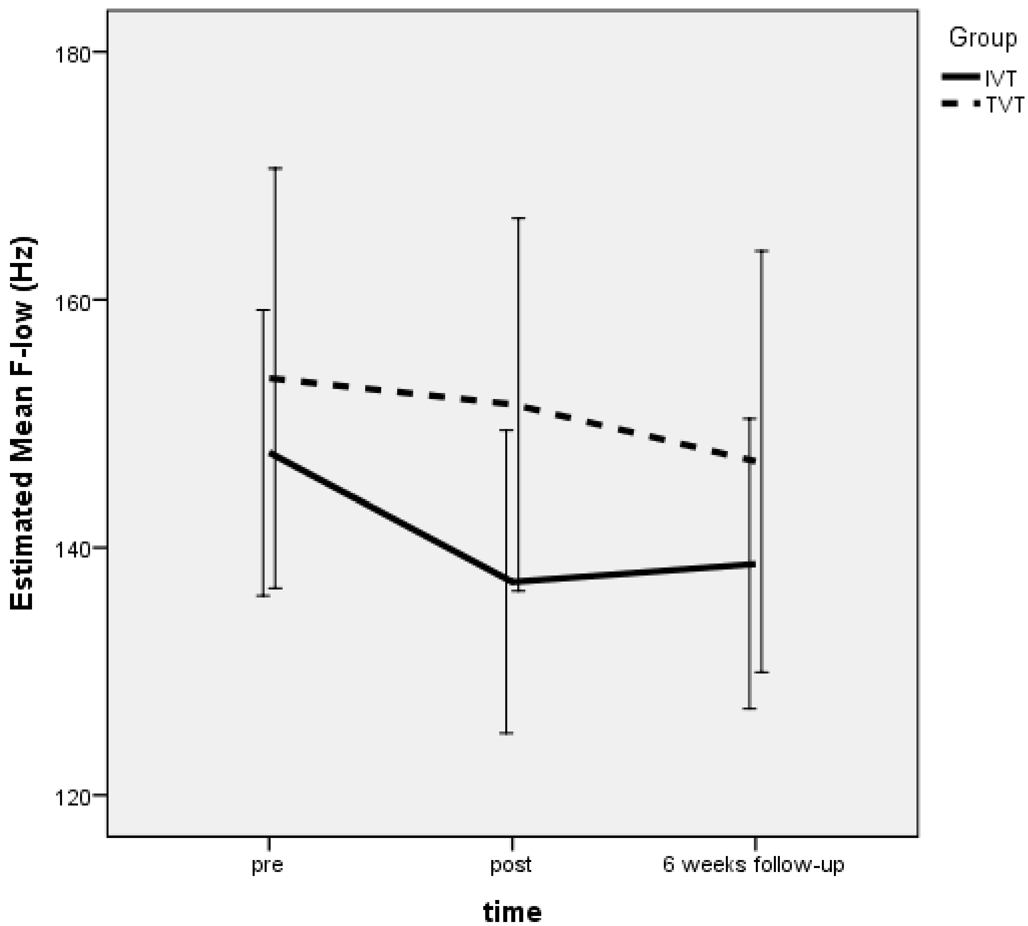

Figure 3. Evolution F-low (Hz) over time in the IVT and TVT groups.

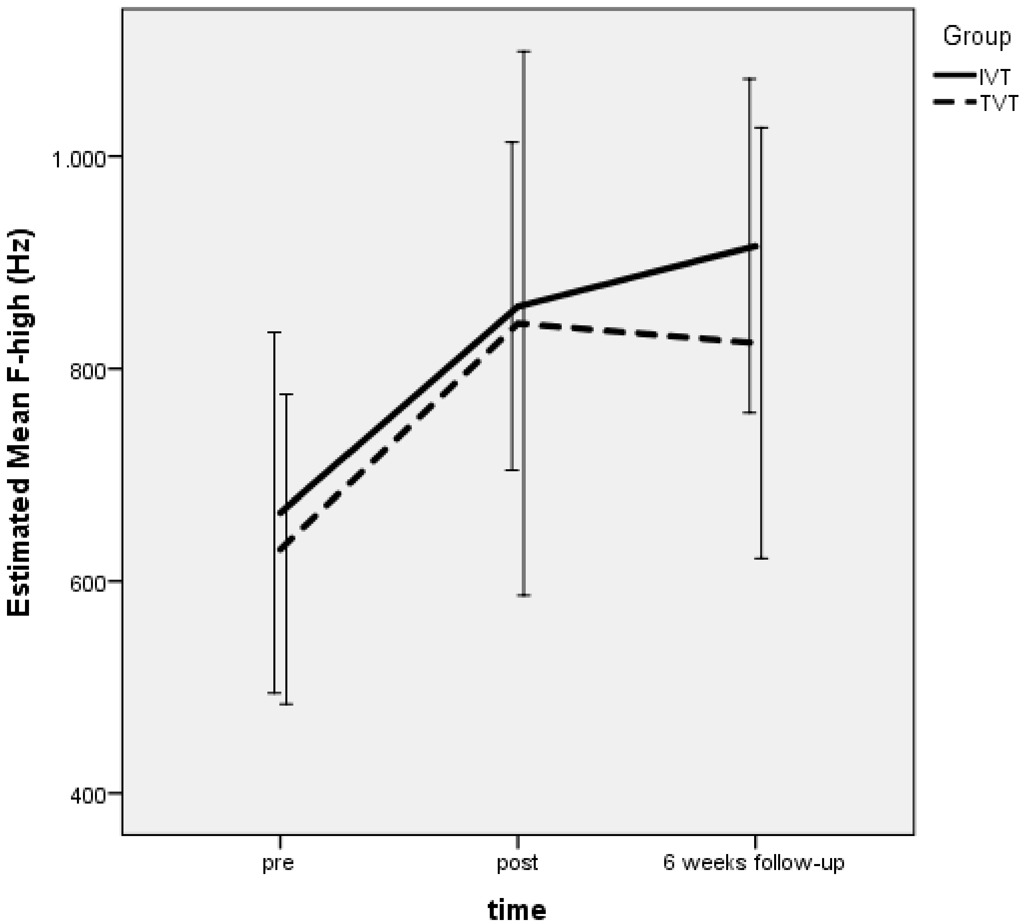

Figure 4. Evolution F-high (Hz) over time in the IVT and TVT groups.

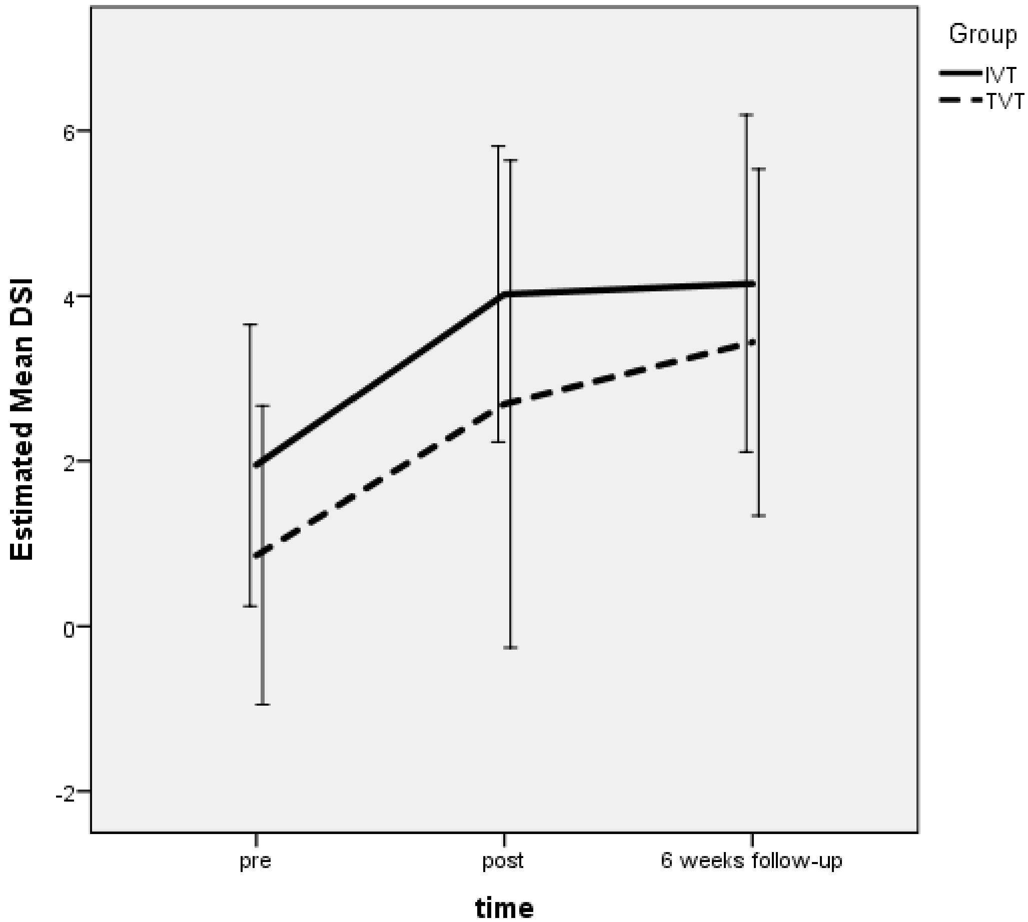

Figure 5. Evolution DSI over time in the IVT and TVT groups.

DISCUSSION

The aim of this study was to compare the effect of a short-term intensive voice training with a longer-term traditional voice training on the vocal quality and vocal capacities of vocally healthy non-professional voice users. Earlier shortcomings (Fu, Theodoros and Ward, 2015a; Wenke et al., 2014; Fischer, Gutenbrunnera and Ptokb, 2009) were met by using a pretest-posttest randomized control group design with follow-up measurements, a voice assessment including both objective measures, auditory-perceptual evaluations and a subjects' self-report, and a standardized and equal training program for both groups. The experiment started with a group of 20 healthy female non-professional voice-users. No significant differences were found regarding age, voice-related symptoms, risk factors, vocal abuse and lifestyle habits between the IVT group and the TVT group.

Based on the principles of behavioral change and the potential advantages of high-intensity training, the authors hypothesized that a short-term intensive voice training may be equally, or even more, effective than a longer-term traditional voice training. This hypothesis has been supported by the results of the current study. (Generalized) linear mixed models showed no significant time-by-group interactions for any of the outcome measures, indicating no significant differences in evolution over time between the groups. Significant time effects were found for the parameters MPT, I-low, F-low, F-high and DSI, all evolving in the desired directions in both groups. More in-depth within-group analyses indicate a preference for the IVT group regarding the evolution of MPT, F-low and DSI, and a preference for the TVT group regarding the evolution of I-low. In contrast to the study of Fu, Theodoros and Ward (2015a), auditory-perceptual evaluations and acoustic perturbation and noise measures showed no significant evolution, probably due to the fact that participants were vocally healthy in this study allowing less significant progress. Visual analog scales may be more sensitive to measure auditory-perceptual differences in this population. The same applies for the self-reported VHI scores, which, in contrast to the study of Wenke et al. (2014), did not significantly improve in the current study.

Vocally healthy participants were selected for this exploratory study. At first, this selection provided more options for a stronger methodological design with less bias. A randomization procedure and better control of influencing factors can easier be achieved in healthy participants than in dysphonic patients. Second, this is a well-considered study group for the aim of exploring motor learning principles of behavioral change that are totally new in vocology. Every voice user, also a vocally healthy individual, is able to improve his or her vocal quality and vocal capacities. Therefore, learning principles will probably apply to any type of voice user. This may cautiously be compared with a typical motor learning task, such as learning how to play tennis. An intensive tennis program will probably lead to more effective and efficient learning than a less intensive one, regardless of the type of player (age, sex, physical fitness, experience etc.). Of course, it is plausible that a younger player with a higher level of physical fitness and experience will learn even more and faster than an older player with less physical fitness and experience. However, a general trend of more effective and efficient learning in the intensive program will likely exist for both individuals. With this idea in mind, we may hypothesize that the current results in healthy participants will give a first general idea of what the most effective distribution of practice might be in vocology. Of course, further research in the whole field (dysphonic patients, elite vocal performers etc.) is needed to make more profound conclusions.

Suppose that this hypothesis is correct and that a short-term intensive model is indeed equally, or even more, effective than a longer-term traditional one, then this will have its consequences for both the patient/client, the voice therapist/coach and the health care system. Time efficiency would be the first advantage for both parties as busy work schedules are no exception these days. Occupational voice users and elite vocal performers are sometimes hindered to work because of their voice problems and want to resume work as soon as possible (Fischer, Gutenbrunnera and Ptokb, 2009; Fu, Theodoros and Ward, 2015a). People who live far from the voice center will experience benefits of a short-term intensive model as they do not have to schedule weekly appointments spread overall several weeks to months (Patel, Bless and Thibeault, 2011). Motivation may increase or be regained as more progress will be noted in a short time frame (Patel, Bless and Thibeault, 2011; Wenke et al., 2014; Fu, Theodoros and Ward, 2015a). Although not shown in the current study, Wenke et al. (2014) found higher attendance for the intensive model, which may reduce frustrations for clinicians associated with cancellations and no-shows. Furthermore, more time efficiency and less drop-out will obviously lead to less financial burden on the client and the health care system (Patel, Bless and Thibeault, 2011; Wenke et al., 2014; Fu, Theodoros and Ward, 2015b).

Besides the many benefits a short-term intensive model has to offer, certain aspects should be kept in mind. At first, the practicality and complexity of scheduling a short-term intensive voice training or therapy should not be underestimated (Bergan, 2010). As said before, time efficiency will eventually overcome, but in the short term it requires a strict scheduling for both the patient/client and the voice therapist/coach. Secondly, the potential risk of overdosing laryngeal tissues cannot be excluded (Bergan, 2010; Roy, 2012; Behlau et al., 2014). Compared with most medical and pharmaceutical therapies, little is known about the moment or threshold at which vocal training transitions from being beneficial to harmful (Roy, 2012). Extreme vigilance by the voice therapist/coach and otorhinolaryngologist will be indispensable in this trajectory (Roy, 2012). However, earlier findings by Fu, Theodoros and Ward (2015a) are promising as patients with vocal nodules showed comparable positive physiological results evaluated with laryngovideostroboscopy (improved ratings of mucosal wave, vocal fold edge smoothness, regularity of vocal fold movement, and glottal closure) postintensive treatment and posttraditional treatment. This indicates no overdose, even for patients with organic voice disorders. Of course, variability will be a key component in the balance between beneficial and harmful dosages (Roy, 2012; Behlau et al., 2014). It is quite possible that the ideal frequency and intensity for one individual may be insufficient or harmful for another (Roy, 2012; Behlau et al., 2014). Despite this variability, we are convinced that a general picture of the most effective and efficient frequency and duration of voice therapy and training is essential. Individualization will be a logical next step.

Limitations of this study are that subjects were not blinded to the purpose of the study and that objective measures were only based on sustained vowel samples. Including voice assessments based on both sustained vowels and continuous speech (e.g. Acoustic Voice Quality Index, Maryn et al., 2010) would be a merit in approximating daily speech and voice use patterns. Another possible limitation is that the (although not significantly) higher proportion smokers in the IVT group may have influenced the results. Stricter exclusion criteria and larger sample sizes with a greater success of randomization may be of value in further research. Besides, convenience sampling as a recruitment procedure has its shortcomings. Implementation of a longer-term follow-up and analysis of the subjects' opining regarding the administered frequency and duration can provide valuable information in future. Investigating the role of telepractice in intensive short-term service delivery models may be an interesting goal for further studies. In general, the principle distribution of practice should be further explored over the whole domain of vocology, which will give us an idea of the optimal dosage for different types of voice users (patients with a variety of voice disorders, professional voice users, elite vocal performers), and undoubtedly be a step forward for both patient/client, voice therapist/coach and the health care system.

CONCLUSIONS

Results suggest that a short-term intensive voice training may be equally, or even more, effective in training vocally healthy non-professional voice users compared to a longer-term traditional voice training. Whether similar results may be expected in different types of voice users and patients with a variety of voice disorders is subject for further research.

REFERENCES

Behlau, M., Madazio, G., Pacheco, C. and Gielow, I., 2014, Intensive short-term voice therapy: the Brazilian experience. *Perspectives on voice and voice disorders*, 24, 98-103.

Behlau, M., Madazio, G. and Oliveira, G., 2015, Functional dysphonia: strategies to improve patient outcomes. *Journal of Patient Related Outcome Measures,* 6, 243-253.

Behrman, A., 2006, Facilitating behavioral change in voice therapy: the relevance of motivational interviewing. *American Journal of Speech-Language Pathology,* 15, 215-225.

Bergan, C., 2010, Motor learning principles and voice pedagogy: theory and practice. *Journal of Singing,* 66(4), 457-468.

Carding, P. N., Horsley, I. A. and Docherty, G. J., 1999, A study of the effectiveness of voice therapy in the treatment of 45 patients with nonorganic dysphonia. *Journal of Voice,* 13, 72-104.

Chen, S. H., Hsiao, T-Y., Hsiao, L. C., Chung, Y-M. and Chiang, S-C., 2007, Outcome of resonant voice therapy for female teachers with voice disorders; perceptual, physiological, acoustic, aerodynamic, and functional measurements. *Journal of Voice,* 21, 415-425.

De Bodt, M. S., Jacobson, B. H., Musschoot, S., Zaman, S., Heylen, L. and Mertens, F., 2000, De Voice Handicap Index. Een instrument voor het kwantificeren van de psychosociale consequenties van stemstoornissen. *Logopedie, 13, 29-33.*

De Bodt, M., Heylen, L., Mertens, F., Vanderwegen, J. and Van de Heyning, P., 2008, *Stemstoornissen: handleiding voor de klinische praktijk* (Antwerpen: Garant).

De Bodt, M., Mertens, D. and Heylen, L., 2008, *Werken aan stem (*Antwerpen: Garant).

De Bodt, M., Patteeuw, T. and Versele, A., 2015, Temporal variables in voice therapy. *Journal of Voice*, 29(5), 611-617.

Dejonckere, P. H., Remacle, M., Fresnel-Elbaz, M., Woisard, V., Crevier- Buchman, L. and Millet, B., 1996, Differential perceptual evaluation of pathological voice quality: reliability and correlations with acoustic measurements. *Revue De Laryngologie - Otologie – Rhinologie,* 117(3), 219-224.

Demmink-Geertman, L. and Dejonckere, P. H., 2010, Differential effects of voice therapies on neurovegetative symptoms and complaints. *Journal of Voice,* 24, 585-591.

Fischer, M. J., Gutenbrunnera, C. and Ptokb, M., 2009, Intensified voice therapy: a new model for the rehabilitation of patients suffering from functional dysphonias. *International Journal of Rehabilitation Research,* 32(4), 348-355.

Fu, S., Theodoros, D. G. and Ward, E.C., 2015a, Intensive versus traditional voice therapy for vocal nodules: perceptual, physiological, acoustic and aerodynamic changes. *Journal of Voice,* 29(2), 260.e31-260.e44.

Fu, S., Theodoros, D. G. and Ward, E.C., 2015b, Delivery of intensive voice therapy for vocal fold nodules via telepractice: a pilot feasibility and efficacy study. *Journal of Voice,* 29(6), 696-706.

Heylen, L., Wuyts, F., Mertens, F., De Bodt, M., Pattyn, J., Croux, C. and Van de Heyning, P. H., 1998, Evaluation of the vocal performance of children using a voice range profile index. *Journal of Speech, Language, and Hearing Research,* 41, 232–238.

Hirano, M., 1981, *Clinical Examination of Voice* (New York: Springer Verlag).

Iwarsson, J., 2015, Facilitating behavioral learning and habit change in voice therapy – theoretic premises and practical strategies. *Logopedics, Phoniatrics, Vocology,* 40(4), 179-186.

Jacobson, B. H., Johnson, A., Grywalski, C., Sirbergleit, A., Jacobson, G. and Benninger, M. S., 1997, The Voice Handicap Index (VHI): development and validation. *American Journal of Speech-Language Pathology,* 6, 66-70.

Maryn, Y., Corthals, P., Van Cauwenberghe, P., Roy, N. and De Bodt, M., 2010, Toward Improved Ecological Validity in the Acoustic Measurement of Overall Voice Quality: Combining Continuous Speech and Sustained Vowels. *Journal of Voice,* 24(5), 540-555.

Mcllwaine, A., Madill, C. and McCabe, P., 2010, Voice therapy prepractice and the principles of motor learning. *Australian Communication Quarterly,* 12(1), 29-32.

Mueller, P. B. and Larson, G. W., 1992, Voice therapy practices and techniques: a survey of voice clinicians. *Journal of Communication Disorders,* 25, 251-260.

Patel, R. R., Bless, D. M. and Thibeault, L., 2011, Boot camp: a novel intensive approach to voice therapy. *Journal of Voice*, 25(5), 562-569.

Raes, J., Wuyts, F. L., De Bodt, M. and Clement, P., 2002, The Dysphonia Severity Index used with a percentage scale. *Stem-, Spraak- en Taalpathologie,* 11(1), 30-37.

Russell, A., Oates, J. and Pemberton, C., 2000, *Voice care for teachers* (Flinders University).

Ramig, L., Bonitati, C., Lemke, J. and Horii, Y., 1994, Voice treatment for patients with Parkinson disease: Development of an approach and preliminary efficacy data. *Journal of Medical Speech-Language Pathology,* 2, 191-209.

Roy, N., 2012, Optimal dose-response relationships in voice therapy. *Internatinonal Journal of Speech-Language Pathology*, 14(5), 419-423.

Stemple, J. C., Lee, L., D'Amico, B. and Pickup, B., 1994, Efficacy of vocal function exercises as a method of improving voice production. *Journal of Voice,* 8, 271-278.

Timmermans, B., 2008, *Klink klaar. Uitspraak- en intonatiegids voor het Nederlands* (Leuven: Davidsfonds).

Van Leer, E., Hapner, E. R. and Connor, N. P., 2008, Transtheoretical model of health behavior change applied to voice therapy. *Journal of Voice,* 22(6), 688–698.

Van Lierde, K. M., Claeys, S., De Bodt, M. and Van Cauwenberge, P., 2007, Long-term outcome of hyperfunctional voice disorders based on a multiparameter approach. *Journal of Voice,* 21(2), 179-188.

Van Lierde, K. M., De Bodt, M., D'haeseleer, E., Wuyts, F. and Claeys, S., 2010a, The treatment of muscle tension dysphonia: a comparison of two treatment techniques by means of an objective multiparameter approach. *Journal of Voice*, 24(3), 294-301.

Van Lierde, K. M., D'haeseleer, E., Wuyts, F. L., De Ley, S., Geldof, R., De Vuyst, J. and Claeys, S., 2010b, The objective vocal quality, vocal risk factors, vocal complaints, and corporal pain in Dutch female students training to be speech-language pathologists during the 4 years of study. *Journal of Voice, 24(5), 592-598.*

Van Lierde, K. M., Claeys, S., D'haeseleer, E., Deley, S., Derde, K., Herregods, I., Strybol, I. and Wuyts, F., 2010c, The vocal quality in female student teachers during the 3 years of study. *Journal of Voice, 24(5), 599-605.*

Verdolini-Marston, K., Burke, M. K., Lessac, A., Glaze, L. and Caldwell, E., 1995, Preliminary study of two methods of treatment for laryngeal nodules. *Journal of Voice*, 9(1), 74-85.

Verdolini, K., 2000, Resonant voice therapy. In: Stemple JC, ed. *Voice Therapy: Clinical Case Studies.* 2nd ed. San Diego, CA: Singular Publishing Group: 46-61.

Vinney, L. A. and Turkstra, L. S., 2013, The role of self-regulation in voice therapy. *Journal of Voice*, 27(3), 390.e1-390.e11.

Wenke, R. J., Stabler, P., Walton, C., Coman, L., Lawrie, M., O'Neill, J., Theodoros, D. and Cardell, E., 2014, Is more intensive better? Client and service provider outcomes for intensive versus standard therapy schedules for functional voice disorders. *Journal of Voice,* 28(5), 652.e31-653.e43.

Wuyts, F. L., De Bodt, M. S., Molenberghs, G., Remacle, M., Heylen, L., Millet, B., Van Lierde, K., Raes, J. and Van de Heyning, P. H., 2000, The Dysphonia Severity Index: an objective measure of vocal quality based on a multi-parameter approach. *Journal of Speech, Language, and Hearing Research,* 43, 796–809.