Vocal Loading and Recovery

Whitling, Susanna

2016

Document Version:
Publisher's PDF, also known as Version of record

[Link to publication](#)

Citation for published version (APA):
Whitling, S. (2016). *Vocal Loading and Recovery*. [Doctoral Thesis (compilation), Department of Clinical Sciences, Lund]. Lund University: Faculty of Medicine.

Total number of authors:
1

Vocal Loading and Recovery

Vocal Loading and Recovery

Susanna Whitling

DOCTORAL DISSERTATION
by due permission of the Faculty of Medicine, Lund University, Sweden.
To be defended at LUX Lower Hall C116A. 2 December, 2016, 9.15.

Faculty opponent
Associate Professor Jennifer Oates, La Trobe University, Australia

Organization	Document name
LUND UNIVERSITY	DOCTORAL DISSERTATION
Faculty of Medicine, Clinical Sciences Lund,	Date of issue 2 December, 2016
Author SUSANNA WHITLING	Sponsoring organization AFA Insurance

Title and subtitle
Vocal Loading and Recovery

Abstract

It is important to refine clinical instruments in order to adequately diagnose and relieve patients suffering from voice disorders. There has of yet been little evidence for how to best care for people who struggle to make themselves heard in noisy environments. There has been no evidence to show what benefits vocal recovery following phonomicrosurgery, when absolute voice rest is compared to relative voice rest with immediate onset following surgery. This dissertation focuses on vocal behaviour during heavy vocal loading and vocal recovery, in patients diagnosed with functional dysphonia and in patients with benign organic lesions in the vocal fold mucosa.

Paper 1 presents successful methodological development of a clinical vocal loading task, setup in order to cause vocal fatigue in voice healthy participants through enhanced vocal effort. The task applies heavy vocal loading by loud reading in a noisy environment and is terminated when participants sense vocal fatigue. Paper 2 examines vocal behaviour in female patients with functional dysphonia compared to women on a spectrum of functional voice problems, comparing vocal behaviour in three conditions: (1) a vocal loading task, (2) work, (3) leisure. The study showed that heavy vocal loading may be associated with high pitch phonation and that patients with functional dysphonia self-report voice problems all through the day, not only during work. Paper 3 compares vocal loading effects and vocal recovery from heavy vocal loading in the same groups as Paper 2. Results show greater detrimental impact from vocal loading and slower short-time recovery for patients with functional dysphonia. Paper 4 is a blind, randomised study, comparing absolute and relative voice rest following phonomicrosurgery in patients with benign lesions in the vocal fold mucosa. This study applies the vocal loading task to assess vocal stamina. Results show that relative voice rest is more beneficial than absolute voice rest, especially regarding compliance, and self-assessments of vocal function. Papers 1, 2 and 4 apply long-time voice accumulation by voice dosimetry and a voice health questionnaire, to track vocal behaviour. In Paper 4 voice accumulation used to track patient compliance with voice rest recommendations.

Conclusions: (1) Patients with functional dysphonia are more vulnerable than others in a context of heavy vocal loading. Their vocal function is worse affected and they recover more slowly compared to others. Care should be given to arm these patients with coping skills during loud speech. (2) Absolute voice rest is not complied with by patients following phonomicrosurgery. (3) Relative voice rest may be beneficial for long-time vocal recovery following phonomicrosurgery of benign lesions in the vocal fold mucosa. Absolute voice rest, i.e. silence, is not recommended. (4) The concept of *vocal loading* is reinforced by measurements of reaction to heavy vocal loading and of recovery time following said loading.

Key words: vocal loading; vocal recovery; functional dysphonia; voice rest; voice accumulation

Classification system and/or index terms (if any)

Supplementary bibliographical information	Language: English

ISSN and key title: 1652-8220, Lund University, Faculty of Medicine, Doctoral Dissertation Series 2016:143	ISBN 978-91-7619-370-9
Recipient's notes	Number of pages 177	Price
	Security classification	

Signature _____ Date 2 December 2016

Vocal Loading and Recovery

Susanna Whitling

LUND
UNIVERSITY

You is never doing anything unless you try
Roald Dahl's BFG

Contents

Preface

What constitutes vocal loading? The whole adventure of writing this dissertation stems from an interest in the concept of vocal loading from a clinical and semantic point of view. The answer to the questions will vary depending on which voice professional is being asked. A phonetician or engineer may explain forces caused by increased muscular tension, adduction and high airflow through the glottis that unfavourably affect the source of the voice signal. The acoustician may explain resonance and audibility, causing people to strain their voices because they cannot make themselves heard to themselves or others properly. The phoniatrician or laryngologist may explain morphological changes in the vocal folds caused by overuse or pathological changes in the vocal folds, affecting their physiology. The singing teacher may describe strain, being out of practice and damage to the vocal instrument. The speech and language pathologist may explain vocal fatigue caused partly by unfavourable vocal behaviour and might also include contextual variation, such as room acoustics or communicative intent to the mix. Vocal loading seems to be a part of vocal warm-up, vocal effort, vocal fatigue, and of vocal recovery. However, how it all comes together and becomes clinically relevant when trying to help patients with voice problems, is not entirely clear, but may be illustrated by the questions: How much phonation is too much and to whom?

Voice professionals seem to agree that vocal load increases as speech and phonation grows louder. As a speech and language pathologist I have wondered how to assimilate these different aspects of vocal loading, to find out if and how they matter in clinical practice. How does vocal loading affect vocal function, and by extension the communicative function of a human being?

Unsurprisingly I have not been able to investigate, all the potential questions branching out from this medical, phonetic, psychological – even philosophical – problem. I have aimed at determining traits of vocal loading and recovery in different parts of the population with clinical voice disorders.

I have attempted to keep a holistic hawk-eye perspective on vocal loading, in order to keep as close as possible to clinical management of voice pathology, which at its core is very complex. As time-consuming and tiring this effort has been, it has been very fruitful. Through my experiments I have striven to provide voice clinicians with hands-on take-home messages, often sought after and rarely provided in research. Hopefully my work can lay a foundation or awaken new questions regarding what constitutes vocal loading and recovery.

Acknowledgements

There is a great deal within the process of writing my dissertation for which I am very thankful. The details of it all cannot be explicitly put into words here. But I shall try to explain the essence of my gratitude, here goes:

A number of people have been especially important to me and my work during the last four years. "Two" people stand out first and foremost. I use quotation marks, because calling them Rolika or Vivand comes more naturally to me than trying to figure out where the great **Roland Rydell** ends and where the remarkable **Viveka Lyberg-Åhlander** begins. As my supervisors, you have trusted me with this awesome task from the word go. When I have not been able to trust myself, you have guided me and I am forever thankful to you both for helping me do this work. I am a different person (in a good way) for it. For the last time "för nu": Tack för hjälpen!

I am deeply thankful to my collaborators within the research project: **Jonas Brunskog** and **Alba Granados Corsellas**: thank you for taking me on board! Thanks to **Adam Vogel** for hosting me at Melbourne University last winter, your input has been essential to my analyses!

Big thanks to **all participants** who took part in my trials, not forgetting the n=20 of you doubling up, allowing me to follow your vocal recovery for up to six months. Thank you for your time, your effort, your patience and your willingness to contribute to clinical knowledge of vocal loading and recovery. I promise to now leave you alone.

During these four years I have been blessed to co-exist with great fellow PhD students, some at home, and some away. Thanks to **Karol**(ina) **Löwgren**, for teaching me the hearing screening process, for your quirky notes and interesting discussions over a cup of steaming hot… water! To **Emily Gre**(mlin)**nner** for your ability to see me on a deeply personal level in our very first meeting and for staying at it! To **Ketty Holmström** for always being up for an interesting talk on statistics and for patiently following my progress from ? to ! To **Olof Sandgren** for being a match to my cynical self and for being a brother in arms through many a concentrated office hour. To **Samuel Sonning** for your ability to quickly adapt hardware and software of "our" voice accumulator to fit my research purposes. To **Greta Wistbacka**, my colleague and friend, for always being a great support and for your ability to confirm, challenge and deepen my knowledge. Thanks also for great backing from **Heike von Lochow, Suvi Karjalainen, Pontus Wiegert, Sofia**

Last, but contrary to least, to my beloved **Thomas**: completing this work would *n o t* have been possible without your breathtaking, life-affirming support through thick and thin. Among many other things you have helped me listen to my body and to my heart, where you now live. Tack älskling!

Dissertation papers

Paper 1

Whitling S, Rydell R, Lyberg-Åhlander V. *Design of a clinical vocal loading test with long-time measurement of voice.* Journal of Voice 2015; 29:261 e13–27

Paper 2

Whitling S, Lyberg-Åhlander V, Rydell R. *Long-time voice accumulation during work, leisure and a vocal loading task in groups with different levels of functional voice problems.* Journal of Voice 2016; article in press, doi:10.1016/j.jvoice.2016.08.008

Paper 3

Whitling S, Lyberg-Åhlander V, Rydell R. *Recovery from heavy vocal loading in women with different degrees of functional voice problems*, 2016; manuscript under review

Paper 4

Whitling S, Lyberg-Åhlander V, Rydell R. *Absolute versus relative voice rest after vocal fold surgery, a randomized clinical trial*, 2016; manuscript

Abbreviations and concepts

Abbreviation or concept	Explanation
ANL	Acceptable Noise Level
AVR	Participants in group randomized for Absolute Voice Rest
C	Voice healthy control participants
F0	Fundamental frequency of the voice (Hz)
FD	Patients with functional dysphonia (all women)
HL	Hearing level
HLC	Participants with High everyday vocal Loading who experience self-assessed voice Complaints
HLNC	Participants with High everyday vocal Loading who do Not experience self-assessed voice Complaints
LTAS	Long Time Average Spectrum of the speech signal
NL	Noise Level (dB)
PAS	Phonatory Aerodymanic System
PTP	Phonation Threshold Pressure
Relative Phonation Time	Percentage of total measurement time spent phonating according to accelerometer signal, based on moving average of the signal in 60 second speech frames.
RVR	Participants in group randomized for Relative Voice Rest
SLP	SLP
SPL	Sound Pressure Level (dB)
SRP	Speech Range Profile during habitual phonatory speech range profile
VAS	100 mm Visual Analogue Scale (0= no voice problems, 100= maximal voice problems)
VHI-T	Voice Handicap Index Throat: self-assessment form
VL	Voice Level (dB)
VLT	Vocal Loading Task
VRP	Voice Range Profile in maximum vocal range phonetogram
8VQ=UVQ	8 voice health questions. Same as Underlying Voice health Questions.
10VQ=SVQ	10 voice health questions. Same as Underlying Voice health Questions.

Introduction

This dissertation deals with two major questions:

1. What clinical implication does vocal loading have for the voice function of patients afflicted with functional dysphonia?

2. What do recovery processes that follow vocal loading and phonomicrosurgery entail for patients with functional dysphonia (former) and patients with benign organic lesions in the vocal fold mucosa (latter)?

The questions have no clear answers as of yet, as very little voice research dealing with vocal loading and recovery processes has examined patients diagnosed with different types of voice pathology. A surprising fact, as vocal impairment afflicts more people than any other communication disorder during a lifespan (Ramig and Verdolini, 1998). Most field studies investigating vocal loading behaviour have explored teachers' voices, because teachers typically experience vocal loading every day (e.g. Lyberg-Åhlander, Pelegrín García, Whitling, Rydell and Löfqvist, 2014). Consequently, the angle in this dissertation is to examine women on a spectrum of muscle tension based functional voice problems in vocally demanding situations and to track ensuing recovery patterns. In order to look deeper into processes of vocal recovery, the vocal function and observable wound healing in patients undergoing phonomicrosurgery for benign lesions in the vocal fold mucosa have been investigated.

A desirable outcome when carrying out the work presented in papers 1–4, was to learn more about vocal behaviour during vocal loading and recovery, and to discuss how the newfound knowledge would be clinically applicable.

The following novel aspects in clinical voice research are put forth in this dissertation:

- o Exploring vocal loading behaviour and recovery from vocal loading in patients with functional dysphonia as well as patients with organic dysphonia.

- o Using voice accumulation technology to track controlled vocal loading, compared to real-life vocal loading and to track compliance with voice rest advice.

- o Letting participants set the time limit for a VLT.

- o Randomising vocal recovery in two groups following vocal fold surgery: assigned to either absolute voice rest or relative voice rest (i.e. speaking straight after surgery).

A principle problem is the relative and complex nature of phonation. Medical voice professionals need to keep a holistic point of view (Behlau, Madazio, and Oliveira, 2015), while at the same time, keeping a detailed record of phonatory physiology, mechanics and acoustics as well as the voice function's connection to emotion. Without the combined, detailed holistic view there will be no consensus base for intervention. Nor will evidence for successful short and long term rehabilitation of vocal function following vocal loading or other phonotrauma be forthcoming.

The four dissertation papers were designed to give as much detailed information as possible on vocal function, and any dysfunction caused by high vocal loading in different parts of the voice pathologic population. Emphasis has been put on participants' self-assessment of vocal function. An important aspect of the dissertation is tracking patients' vocal behaviour in great detail in the voice clinic as well as in the field, thus striving to reach a holistic patient perspective of vocal dysfunction.

In Paper 1 a VLT was developed and applied in n=11 voice healthy participants. The VLT was designed to cause immediate vocal fatigue, but no long-term damage. In Paper 2 vocal behaviour during long-time voice accumulation was tracked in n=50 women in four subgroups, based on degree of functional voice problems. Vocal loading in everyday situations was explored over 7 days, and compared to the controlled VLT. In Paper 3 the same participants as in Paper 2, recorded immediate and gradual self-assessed recovery from controlled vocal loading. Paper 4 also explored vocal recovery,

with a different scope. This study examined n=20 patients undergoing surgery for benign organic lesions in the vocal fold tissue. These patients were randomly assigned to absolute or relative voice rest, challenging the prevailing routine of absolute voice rest for full vocal recovery. Vocal stamina (i.e. vocal loading capacity) was examined in the VLT pre-surgery and at a 3–6-month follow-up. Vocal behaviour was also examined through voice accumulation for seven days at both time points.

Vocal loading

Ever since the 1960s it has been hypothesised that prolonged phonation at increased pitch and intensity levels, i.e. vocal loading, would be detrimental to the vocal function. It has in fact been called vocal abuse. The notion of vocal loading is difficult to contain within one definition, and not many have examined the phenomenon based on a clear definition.

Empirical evidence for vocal loading was relatively uncharted until the mid-1990's. In 1995, Linville expressed surprise at the lack of studies examining laryngeal changes and vocal performance changes due to loud phonation (Linville, 1995). Vocal loading had been explored earlier, but without proper definition. Rather, the already affected/pathological voice had been in focus before then (Gelfer, Andrews, and Schmidt, 1991; Neils and Yairi, 1987; Sherman and Jensen, 1962; Stone and Sharf, 1973). More knowledge on the transition from unaffected to affected vocal function was called for.

Definitions of vocal load started to emerge, e.g. when Titze (2001) equated vocal load to vocal dose, i.e. vocal intensity over total measured time. Vilkman (2004) gave a more faceted definition of vocal loading, basing the concept on prolonged phonation and adding loading factors, which would make phonation, while making oneself heard, more difficult. Figure 1 shows Vilkman's definition of vocal loading, which incorporates individual variation of vocal loading. Definitions by Titze (2001) and Vilkman (2004) have made up the basis for understanding vocal loading throughout this dissertation.

Field studies of vocal behaviour have not yet mapped vocal behaviour in clinical voice patients. Changes in phonatory function due to high vocal loading have been examined using different types of laboratory set vocal loading exercises, often based on loud, prolonged reading. The laboratory

setting is easy to control, which is important in a multifaceted research area as clinical voice research, however it may not represent real life voice use. For example, F0 is higher in real life phonation than in a laboratory setting (Lehto, Alku, Vilkman, and Laaksonen, 2006). Often a set time limit of 15–45 minutes has been applied (Buekers, 1998; Caraty and Montacié, 2014; Chang and Karnell, 2004; De Bodt, Wuyts, Van de Heyning, Lambrechts, and Vanden Abeele, 1998; Doellinger, Lohscheller, McWhorter, and Kunduk, 2009; Gelfer, Andrews, and Schmidt, 1996; Hanschmann, Gaipl, and Berger, 2011; Hunter and Titze, 2009; Kelchner, Toner, and Lee, 2006; Laukkanen, Järvinen, Artkoski and colleagues, 2004; Lohscheller, Doellinger, McWhorter, and Kunduk, 2008; Niebudek-Bogusz, Kotylo, and Sliwinska-Kowalska, 2007; Remacle, Morsomme, Berrué, and Finck, 2012; Sherman and Jensen, 1962; Stemple, Stanley, and Lee, 1995; Södersten, Ternström, and Bohman, 2005; Vilkman, Lauri, Alku, Sala, and Sihvo, 1999; Vintturi, Alku, Lauri, Sala, Sihvo and Vilkman, 2001; Vintturi, Alku, Sala, Sihvo and Vilkman, 2003).

The setup of vocal loading in the current study resembles that of De Bodt and colleagues (1998), with clinical and acoustic evaluations carried out before and promptly after a VLT involving loud reading (details in the methods section). An important difference between De Bodts' and the current setup is the latter's incorporation of Vilkman's additional loading factors (see Figure 1). Additional loading factors increases the level of complexity necessary for keeping a patient perspective, in line with the International Classification of Functioning, Disability and Health (WHO, 2001, see Figure 3). Prolonged vocal loading is expected to cause strain to inner and outer laryngeal structures, changing the conditions for effective vibration in the vocal folds, thus altering the voice source itself and the resonating vocal tract. Evidence of vocal loading in a noisy environment is a speaker's habitual raising of phonatory SPL according to the Lombard effect (Lane and Tranel, 1971), which leads to involuntary increase of phonatory F0, as merely a physiological response to increased SPL and as spectral adaptation, in order to be heard in the background noise (Lane and Tranel, 1971; Södersten and colleagues, 2005). Others have found detrimental, subjective changes of the voice function (Buekers, 1998; De Bodt and colleagues, 1998). One subjective change is a sense of increased vocal fatigue during vocal loading.

Loading factors	Contributing individual factors
Duration of voice use	Gender
Background noise	Endurance
Room acoustics	Health condition
Quality of air	Life habits
Other ergonomics	Vocal skills
Psychosocial	Psychosocial/personality
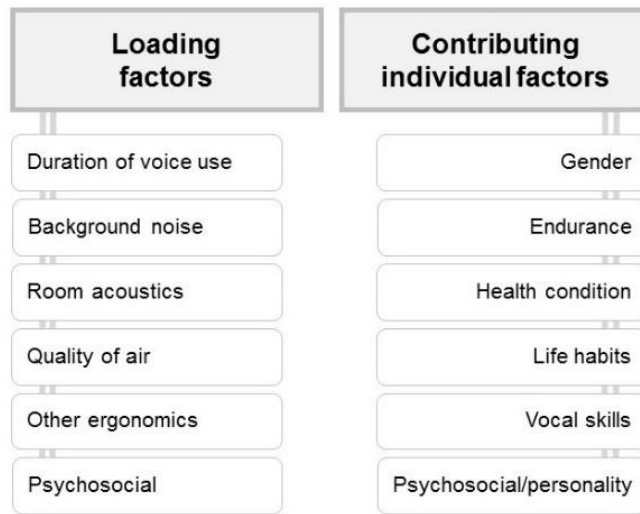
Vilkman, 2004

Figure 1: Intra and extrapersonal factors adding to vocal loading according to Vilkman, 2004.

There is no uniform definition of vocal fatigue (Kitch and Oates, 1994), but Welham and colleagues propose focusing on short-term functional voice changes, as opposed to other pathologic voice conditions, with the following definition:

> *Vocal fatigue is used to denote negative vocal adaptation that occurs as a consequence of prolonged voice use. Negative vocal adaptation is viewed as a perceptual, acoustic, or physiologic concept, indicating undesirable or unexpected changes in the functional status of the laryngeal mechanism.* Welham and Maclagan (2003, p. 22)

When it comes to objectively measurable changes to the voice signal there is little conformity and more knowledge is needed (Solomon, 2008). For this reason, many researchers have called for vocal loading to be explored in field studies instead of in the controlled situation of the laboratory (Buekers, 1998; Buekers, Bierens, Kingma and Marres, 1995; Oates and Winkworth, 2008), which many since have attempted. In the current dissertation short-term implications of vocal loading are investigated.

Voice accumulation

Around the same time as Titze defined vocal loading (Titze, 2001), Rantala and Vilkman had started exploring the vocal loading index – i.e. exploring an estimate of the number of oscillations the vocal folds perform during a set time of phonation, which correlated well with participants' subjective voice complaints (Rantala and Vilkman, 1999). Titze and colleagues developed this idea from the point of view that vocal dose is related to the vocal folds' exposure to vibration, and that excessive vibration can be detrimental to the lamina propria. They broke down vocal dose into three separate units: (1) time dose (total phonation time), (2) energy dissipation dose (total amount of heat dissipated over a unit volume of vocal fold tissues) and (3) distance dose (Titze, Švec and Popolo, 2003). This approach was important ground work for field measurements of vocal behaviour using voice dosimetry, i.e. voice accumulation.

Voice accumulation has been used in numerous studies to track participants' vocal behaviour in the field (e.g. Buckley, O'Halloran, and Oates, 2015; Hunter and Titze, 2010; Lyberg-Åhlander and colleagues, 2014; Szabo Portela, Hammarberg and Södersten, 2014), however none have thus far compared vocal behaviour of clinical voice patients. There are currently three commercially available voice dosimeters for the voice clinician: Ambulatory Phonation Monitor (APM, KayPENTAX, NJ, USA), VocaLog (Griffin Laboratories, CA, USA) and VoxLog (Sonvox AB, Umeå, Sweden). See Van Stan and colleagues for comparisons of these devices (Van Stan, Gustafsson, Schalling and Hillman, 2014). The last of these is used in this dissertation (details in the Methods section). VoxLog is portable and provide measurements of phonatory SPL, ambient SPL by measurements from a microphone and F0 and relative phonation time (the time spent phonating during the total recording time) by measurements from an accelerometer. An important feature of voice accumulation is that it gives the ability to track vocal behaviour without recording the speech signal, ensuring the integrity of the participant. As there is to date no voice accumulation system which includes subjective rating of vocal function, e.g. using a mobile phone app, it is important to track participants' self-assessments along with their carrying the voice accumulator.

Clinical voice measurements

In the clinic, objective vocal function is mainly evaluated through perceptual analysis of voice recordings and either laryngostroboscopy or high speed digital imaging of the vocal folds, during phonation and rest (Dejonckere, 2000; Dejonckere, Bradley, Clemente and colleagues, 2001). Recordings need to be standardized both in producing and in analysing, e.g. according to the G (grade) R (roughness) B (breathiness) A (asthenia) S (strain) scale (Hirano, 1981) or the SVEA (Stockholm Voice Evaluation Approach) using a 100 mm visual analogue scale (VAS) to evaluate hyperfunction, breathiness, instability, roughness, glottal fry, sonority and grade of overall voice pathology (Hammarberg, 1986). SVEA was chosen in this dissertation.

Subjective vocal function is also examined in the voice clinic, through patients' self-assessments, aimed at assessing the impact of vocal dysfunction on the individual. Voice Handicap Index (VHI) is often used in its original format by (Jacobson, Johnson, Grywalski and colleagues, 1997), or the shorter version, VHI-10 (Rosen, Lee, Osborne, Zullo and Murry, 2004). VHI addresses three areas: voice function, physical and emotional aspects of voice function. There are other self-assessment tools which deal more with the quality of life, such as the Voice Activity Participation Profile/VAPP (Ma and Yiu, 2001), and the Voice related Quality of Life/VrQoL (Hogikyan, Wodchis, Terrell, Bradford and Esclamado, 2000). Very recently Nanjundeswaran and colleagues also developed the vocal fatigue index (VFI), aimed to highlight vocal fatigue (Nanjundeswaran, Jacobson, Gartner-Schmidt and Verdolini Abbott, 2015).

Vocal Recovery

Voice rest affects the condition of the entire vocal instrument. No studies have thus far explored recovery from vocal loading in functional dysphonia, although it is said to be a fundamental part of vocal health. Hunter and Titze have explored vocal recovery along the lines of wound healing trajectories, comparing it to dermal wound tissue, which can be of an acute or a chronic nature (Hunter and Titze, 2009). McCabe and Titze proposed examining recovery from vocal loading in 15-minute intervals during one hour following

vocal loading. This measure has been chosen for this dissertation (McCabe and Titze, 2002).

When Titze discussed mechanical stress in phonation. Tensile stress is caused by cricothyroid contraction, i.e. applies during increased pitch of phonation. It is the largest mechanical strain put upon the vocal ligament and the hypothesis was that collagen fibres within the vocal ligament may rupture during very high stress. In his discussion, Titze drew an important parallel to the longitudinal load put upon the anterior cruciate ligament of the knee during maximum stress, which is roughly twice the size of the vocal ligament. Titze stated that if they behave in the same manner there is little risk of the vocal ligament rupturing during phonation, as it is impossible to stretch the ligament to such a required length:

> *We assume, therefore, that a well-developed and healthy vocal ligament provides a "safety-valve" for other, perhaps more injury-prone tissues in the vocal folds. The ligament limits elongation and assumes most of the tensile stress at high pitches.* Titze (1994a, p. 105)

Based on the assumption that vocal fatigue is a short-term, acute phonotrauma caused by vocal loading it can be assumed that patients with functional dysphonia would be worse afflicted by such loading than others. The parallel between orthopaedic strain and strain put upon the vocal fold tissue has been further explored. In a review, Ishikawa and Thibeault showed that one week's voice rest is the most common recommendation following phonomicrosurgery. What predicted optimal outcome was patient compliance to voice care advice. In line with Hunter and Titze (2009) they found vocal fold healing to resemble general wound healing, which turned them to orthopaedic literature. This literature suggests that early, controlled mobilization of ligaments, is beneficial for recovery of functional movement. What possible effects this would have of effect for the quite different fibroblasts in the lamina propria were thus far unknown. Orthopaedics need increased stiffness to improve scarring sustainability. Vocal folds need viscosity and soft tissue for the complexity of muscular and ligament activity combined with the supple waveform vibration on top. However, Ishikawa and Thibeault stated that early mobilization may improve collagen synthesis and encourage fibres to heal while arranging themselves in the direction of movement. The question is whether aerodynamics can heal the subtle and complex architecture of phonation (Ishikawa and Thibeault, 2010). Wrona and colleagues recently reviewed research on wound healing by replacement of tissue in the vocal folds. They present hope for future possibilities of using

native tissue replacement of the extracellular matrix in order to regain functional tissue movement in the lamina propria. The complexity of the vocal fold function is probably what hampers this treatment research (Wrona, Peng, Amin, Branski and Freytes, 2016).

To date the most common voice rest advice patients are recommended following phonomicrosurgery of benign lesions in the vocal fold mucosa is 7 days' absolute voice rest. Relative voice rest is less commonly recommended, and less well defined (Behrman and Sulica, 2003; Coombs, Carswell, and Tierney, 2013). When studying simulations of inflammation in laryngeal secretion, it has been shown that there is predicted inflammation for up to 24 hours post injury to the lamina propria (Li, Verdolini, Clermont and colleagues, 2008), suggesting short-term implications of vocal loading and scarring following phonomicrosurgery.

Absolute voice rest recommended for a differing number of days has very recently been compared in randomised clinical studies, examining patients undergoing phonomicrosurgery (Kaneko, Shiromoto, Fuji-Kurachi, Kishimoto, Tateya and Hirano, 2016; Kiagiadaki, Remacle, Lawson, Bachy, and Van der Vorst, 2015). Both studies show evidence that shorter periods of silence (3–5 days) is more efficacious than longer periods of silence (7–10 days). These findings are well in line with evidence from Rousseau, Cohen, Zeller, Scearce, Tritter and Garret (2011), who showed compliance with absolute voice rest to be difficult for patients. There are no published long-term studies that compare vocal fold function in patients following phonomicrosurgery.

Summary of introduction

Vocal loading and recovery from vocal loading and phonomicrosurgery is a complex research area, wherein many factors need to be controlled. Little is known about voice patients' reactions to vocal loading or about their recovery processes following (1) vocal loading or (2) phonomicrosurgery of benign lesions. Nor has their relevance been thoroughly evaluated for the voice clinic. These processes can be tracked using traditional tools from the voice clinic, combined with voice accumulation, that includes voice dosimetry and repeated measures of patients' self-assessed voice function.

Aims and hypotheses

Aims

The first aim of this dissertation was to investigate different aspects of vocal behaviour during vocal loading in parts of the population suffering from functional and benign organic vocal pathology. The second aim was to explore how participants react to vocal loading and recover from vocal loading and the strain of surgery of benign lesions in the vocal fold mucosa. See specific aims of each Paper in Table 1 and general aims Figure 2.

Table 1: Overview of main aims of dissertation studies.

Paper	Title	Main aim
1	*Design of a clinical vocal loading test with long-time measurement of voice*	Design and test of a clinical vocal loading task. Main goal: apply methods of vocal loading to mimic everyday vocal loading during a vocal loading task and tracking processes involved in loading and recovery
2	*Long-time voice accumulation during work, leisure and a vocal loading task in groups with different levels of functional voice problems*	Use long-time voice accumulation to examine how vocal behaviour in women with diagnosed functional dysphonia differs from that of women who also experience voice problems, but who do not seek therapy, in three conditions: a vocal loading task, work and leisure
3	*Recovery from heavy vocal loading in women with different degrees of functional voice problems*	Examine whether patients with diagnosed functional dysphonia take longer than others to recover from a vocal loading task over an immediate and/or a gradual course. Also examine if the patients react differently than others to a vocal loading task.
4	*Absolute versus relative voice rest after vocal fold surgery, a randomized clinical trial*	Explore patient's compliance with absolute and relative voice rest advice during seven days in two randomized groups following surgery of benign lesions in the vocal folds and also compare short and long-term vocal recovery in the two groups
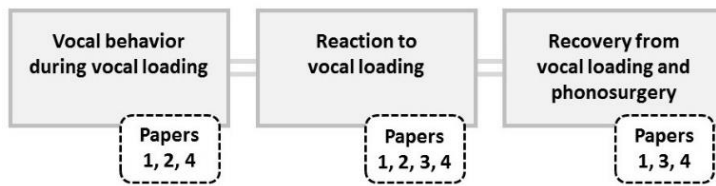
Figure 2: General aims of the dissertation, and their distribution across papers

Hypotheses

Paper 1

It is possible to set up a vocal loading task (VLT) that will cause vocal fatigue in voice healthy participants, without causing permanent damage to the vocal function.

Paper 2

1. Female patients with functional dysphonia use their voice to a greater extent than women who also experience voice problems, but who do not seek medical voice treatment.
2. Patients with functional dysphonia report self-assessed voice problems to a greater extent than women in three groups on a spectrum of functional voice problems

Paper 3

1. Female cisgender patients with functional dysphonia will be worse affected by a VLT than women in three groups on a spectrum of functional voice problems.
2. Patients with functional dysphonia will terminate the VLT sooner than other groups.
3. Recovery from vocal loading will take longer for patients with functional dysphonia compared to other groups.

Paper 4

1. It is difficult to comply with a recommendation of absolute voice rest following phonomicrosurgery.
2. It will benefit healing processes in the vocal folds to let patients comply with relative voice rest advice instead of absolute voice rest for one week following phonomicrosurgery.

Methods

The four papers included in this dissertation are all based on the same notion of testing controlled vocal loading and recovery of vocal function, either from a vocal loading task (VLT) or following vocal fold surgery. The research protocols were greatly inspired by the International Classification of Functioning, Disability and Health, which promotes salutogenesis, i.e. a patient perspective driven focus on health and well-being instead of merely focusing on dysfunction (Antonovsky, 1987; WHO, 2001). This entailed an interest in finding positive coping skills, as well as mapping dysfunctional vocal behaviour. The boundaries, where vocal ability and disability meet, have been of great interest when planning the current clinical trials. Figure 3 shows a model of how ICF has been applied in the current investigation.

Application of ICF: levels of functioning, disability and health
from WHO, 2001

health condition
vocal status

body structure & function
lungs, larynx, vocal tract

activity
phonation

participation
am I heard?

context

environment
room acoustics, noise,
communication partner(s)

emotion
personality traits
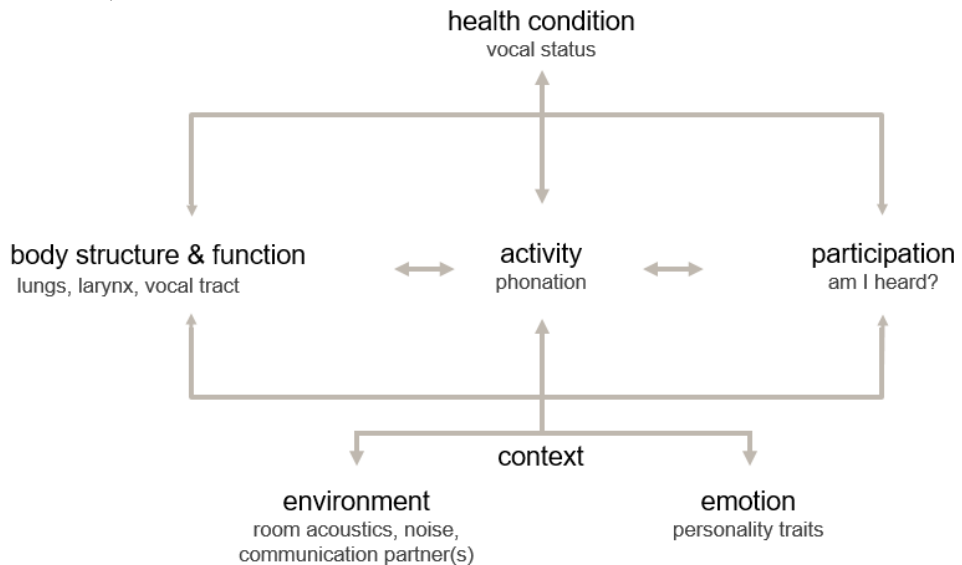
Figure 3: WHO's International Calssification of Functioning, Disability and Health, with applicable examples from the current investigation.

Voice ergonomics have also been important when planning trials. The notion raises awareness about risk factors for work-related voice disorders and how to prevent and avoid them by taking e.g. voice production and ambient, disturbing noise into account (Rantala, Hakala, Holmqvist and Sala, 2012; Sala, Ketola, Laine, Olkinuora, Rantala, Sihvo, 2009). The hypothesis, that a VLT will affect participants with functional voice problems worse than others, was tested with voice ergonomics in mind. The setup of the VLT was chosen to ensure as high ecological validity as possible, i.e. keeping laboratory trials close to how genuine, heavy vocal loading would manifest in the field (Brewer, 2000). Vocal loading was tested in a VLT, the outcome of and recovery from which has been compared in groups with differing vocal function. The VLT was meant to track vocal recovery of participants undergoing surgery for benign lesions in the vocal fold mucosa. Most of the methods of Papers 1–3 are based in and around the VLT. It is shown in Figure 4 and described in detail below. Another method used and examined in papers 1, 2 and 4 was the use of voice accumulation through vocal dosimetry and a voice health questionnaire.

Set-up of vocal loading task and voice accumulation
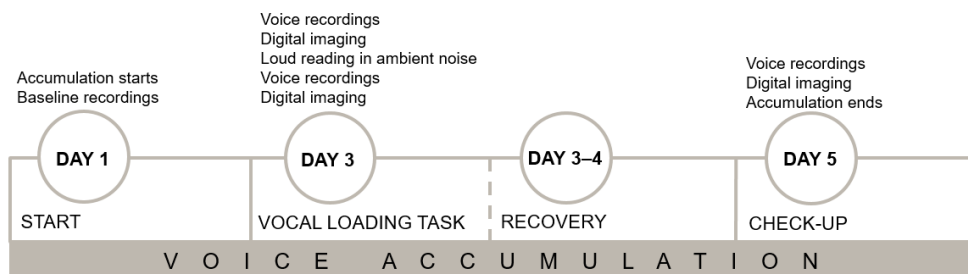
Figure 4: Flow chart of vocal loading task setup during voice accumulation.

Participants

The people examined in the dissertation covered groups of participants with typically functioning voices (papers 1, 2, 3), participants with functional voice problems of differing severity (papers 2, 3) and participants with organic changes to their vocal folds, which would possibly be resolved by one session of phonomicrosurgery and recovered within a couple of months (Paper 4), see Table 2. Participants taking part in the studies of papers 2 and 3 took part in one data collection only, not two separate sets. In no study were participants informed about the absolute study rationale.

Table 2: Distribution of participants in all papers

Paper	1	2 and 3	4
N	11	50	20
n drop-out at check-up	-	-	2
f/m	f: 6, m:5	f: 50	f: 14, m: 6
voice pathology	voice healthy	• functional dysphonia • undiagnosed functional voice problems • high everyday vocal loading • voice healthy	organic voice disorders with benign lesions in the vocal fold mucosa pre and post phonosurgery

Participants in Paper 1

This pilot study aimed to develop and test conditions for a clinical VLT. A range of voice healthy individuals was needed to test the VLT, in order to confirm its ability to attain vocal fatigue without causing any long-term damage to the vocal function. Inclusion criteria were ≥18 years of age, voice healthy, non-smokers with or without speaking voice training. Exclusion criteria were any laryngeal pathology, smokers and professional singers. Eleven (f=6, m=5) voice healthy, cisgender, non-smoking, participants with low to moderate vocal loading occupations were recruited among colleagues and friends of the test leader through verbal inquiry and agreement. No one turned down participation. Mean age 36 (28–55, SD: 8). n=4 participants were choir singers, none of whom had professional training. n=5 had trained speaking voices. All participants in Paper 1 underwent pure tone audiological screening at ≥20 dB SPL at 250 Hz, 500 Hz, 1 kHz, 4 kHz and 8 kHz, to ensure participants' reactions in the VLT were due to vocal loading, and not

caused by any apparent hearing impairment. All participants passed with better ear hearing level of ≥ 20 dB SPL. All participants were screened with laryngeal examination before taking part in the study, none showed any organic pathology. Specific demography is shown in Table 3.

Table 3: Demographics of n=11 participants in Paper 1.

Parti-cipant	f/m	Age	Better ear HL (dB SPL)	Vocal load work	Allergies	Medication	Stress level
1	f	33	25	low	pollen, nutritional, fur	antihista-mines	moderate, fluctuates
2	f	36	20	moderate	-	-	high, constant
3	f	36	20	low	-	-	high, fluctuates
4	f	34	20	low	pollen, cat	-	moderate, fluctuates
5	f	47	20	moderate	asthma, pollen, fur	omeprazole	high, slow fluctuation
6	f	28	20	low	penicillin V	-	high, fluctuates
7	m	33	20	some	-	-	moderate, fluctuates
8	m	55	20	moderate	-	-	high, constant
9	m	34	20	moderate	-	-	moderate, constant
10	m	28	20	moderate	-	-	high, fluctuates
11	m	34	20	moderate	undiagnosed cat, horse	-	high, constant

Participants in papers 2 and 3

Paper 2 examined the same participants as Paper 3. From different perspectives the two studies aimed to explore real life everyday vocal loading and controlled vocal loading in parts of the population suffering from functional voice problems of different magnitude. As functional voice problems occur more often in women than in men, the main inclusion criterion for these studies was that participants had to be cisgender women. In the two studies recruitment was made for four groups: (1) Patients with functional dysphonia=FD, (2) Women with high everyday vocal loading occupations with self-perceived voice complaints=HLC, (3) Women with

high everyday vocal loading occupations with no self-perceived voice complaints=HLNC and (4) Voice healthy women, controls=C.

Recruitment took place in various ways. It is difficult to account for the number of prospective patients who decided not to take part in the studies, as recruitment was partly outsourced to voice clinicians in different voice clinics in and around Lund in southern Sweden. These voice clinicians (Speech and Language Pathologists (SLP's) and phoniatricians) asked women diagnosed with functional dysphonia to take part and n=20 female voice patient agreed. Patients with psychogenic components to functional dysphonia were not included, only patients with functional dysphonia based in muscle tension. According to a power analysis, recruitment ceased after 20 patients.

Apart from the group of patients with functional dysphonia another three groups were recruited to studies 2 and 3. The group of patients with functional dysphonia formed a sample basis upon recruiting the other three groups, matching for age, general health and life situation. Due to this recruitment method, smokers were not excluded from the study, as they occurred in small number in the group with functional dysphonia. Recruitment for the two groups with high occupational work load (HLC and HLNC) was made from the participant base in the study by Lyberg-Åhlander, Rydell, and Löfqvist (2011). Participants who in this previous study had given written consent to take part in any further studies were asked. The question on voice problems was reiterated for the current study, as five years had passed. Participants for the voice healthy control group were mainly recruited through the patients with functional dysphonia, who were asked to bring a voice healthy female peer of the same age. Additional recruitment for the high vocal load (HLC, HLNC) and control (C) groups was made through inquiry on social media. In all n=50 participants were recruited in 4 groups. All participants gave written informed consent to take part in the study. Participants were not screened for hearing impairments, instead a thorough medical history covered hearing and overall health condition. Table 4 shows specific demography and grouping.

Table 4: Demographics of n=50 participants in papers 2 and 3.

Subgroup	Functional dysphonia	High load, complaints	High load, no complaints	Voice healthy controls
Short	FD	HLC	HLNC	C
n	20	10	10	10
Age	44 (22–70, SD: 15)	50 (33–62, SD: 10.6)	48 (32–65, SD: 9.8)	45 (25–67, SD 11.3)
Vocal health status	Patients with medically diagnosed functional dysphonia in different stages of voice therapy	Self-assessed functional/ occupational voice problems. Not saught help.	No self-assesed voice problems	No self-assesed voice problems
Medical status	asthma (5), allergies (4), ulcer (3), diabetes (1), high blood pressure (1), migraines (1), reflux (1),	high blood pressure (2), allergies (1), asthma (1), burn out (1)	depression (2), allergies (2), hypothyroi-dism (1), asthma (1), rheumatism (1)	depression (2), allergies (2), asthma (1)
Occupation	teacher (9), retired former teacher (4), student (3), administrator (2), counselor (1), priest (1)	teacher (9), singing teacher (1)	teacher (6), medical doctor (2), secretary (1), guide (1)	administrator (3), retired (1), student (1), medical doctor (1), chemist (1), translator (1), train driver (1), acoustical consultant (1)
Vocal load	Differing occupational vocal load	High occupa-tional vocal load	High occupa-tional vocal load	Low/no occupa-tional vocal load

Participants in Paper 4

Twenty (n=20) participants took part. The inclusion criteria were patients of 18 years or older, planned for surgical treatment of benign organic lesions in the vocal fold mucosa. Malignancies, lesions affecting muscle fibre, laryngeal papilloma and neurological voice pathologies were excluded. Patients were orally asked for participation by different surgeons in the phoniatric clinics in Lund, Malmö and Helsingborg in southern Sweden. Surgery was performed by three phoniatricians/phono surgeons through cold excision ad modum Bouchayer and Cornut, without use of fibrin glue in Reinke's oedema (Bouchayer and Cornut, 1992). Surgery and data collection were carried out in these clinics during a 2,5-year period from September 2013 to April 2016. Of all patients meeting inclusion criteria (86), 23% (n=20) participated, see Table 5. One participant from each randomized group did not participate in

the long-term check-up. According to clinical routine, participants were given sick leave for the duration of recovery (seven days) if their work was demanding for their vocal function.

Table 5: Demographics of n=20 participants in Paper 4.

	Absolute Voice Rest	Relative Voice Rest
Total n	10	10
Female	7	7
Male	3	3
Age, mean (range, SD)	53 (35–68, 12.2)	45 (25–73, 16.1)
Vocal pathology	Reinke's oedema (2 unilateral, 2 bilateral) cyst and granuloma (1, with excised granuloma) polyp (4) hyperplastic squamous epithelium (1)	Reinke's oedema (2 unilateral, 1 bilateral) cyst (3) polyp (2) sulcus (2)
Occupation	retired (3), teacher (2), chef (1), consultant (1), office worker (1) pharmacist (1), retail clerk (1)	retired (2), singer (2), teacher (1), singing teacher (1), chef (1), consultant (1), office worker (1), deli clerk (1)
Medical status	smoker (4), depressive (3), high blood pressure (2), chronic obstructive pulmonary disease (1), hyperthyroidism (1), reflux (2), ulcerative colitis (1), hearing impairment (0)	smoker (3), high blood pressure (2), depressive (1), multiple sclerosis (1), reflux (1), hearing impairment (0)
Voice therapy	preoperative (0) postoperative (0)	preoperative (0) postoperative therapy (1 singer)

The vocal loading task

This research project started out with developing a VLT for clinical conditions. The physical setup is depicted in Figure 5. The main aim for the VLT was setup with ecological validity taken into account, so that it would simulate everyday vocal loading during a period of time manageable for test administration and to attain vocal fatigue without causing any long-term damage to participants' vocal function. Processes involved in heavy vocal loading and recovery following said heavy vocal loading were tracked. As previously mentioned, vocal loading can not only be measured by comparing vocal function before and after prolonged voice use. Additional loading factors have to be added (Vilkman, 2004).
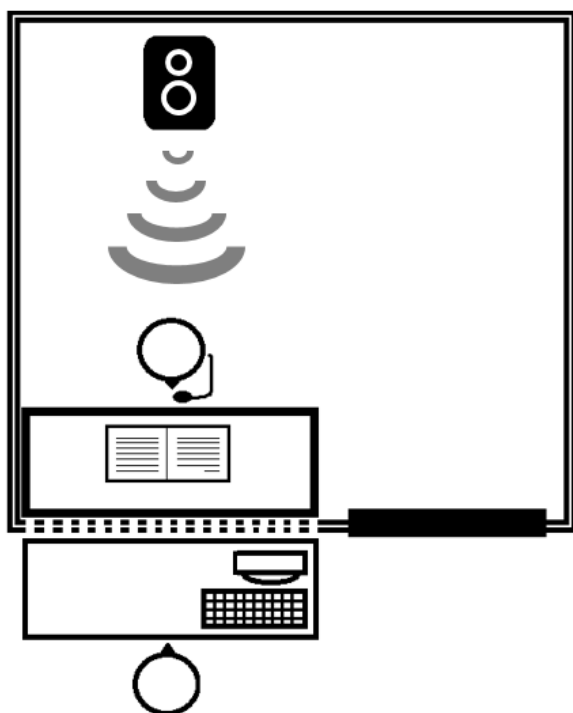
Figure 5: Physical setup of the vocal loading task

Participants phonated for a lengthy time through loud reading, seated at a desk in a soundproof double-walled booth, complying with the maximum permissible ambient SPL as specified in ISO 8253-1 (ISO 8253-1, 1989). The participants were asked to keep reading, only stopping for breathing and sips of water, which was provided. They were also asked to terminate the reading if and when they felt distinct discomfort from the throat, leading to differing task participation time. This uncommon approach prioritises individual vocal comfort, above comparable participation time. The approach takes Vilkman's multifaceted definition of vocal loading into account: participants will reach a state of vocal fatigue at different times depending on endurance and other individual factors (Vilkman, 2004).

In order to mimic everyday vocal loading and keep as close to spontaneous speech as possible without giving the participants time to pause, they were asked to choose an easy-to-read test in Swedish to read aloud during the VLT. Thus participants would avoid overloading their working memory. If they did not bring a text they were provided with an easy-to-read

text on kings and queens of Sweden. They were told it was not important they read correctly, only to keep going.

As they started reading the participants sat in a silent booth. After 30 seconds an International Speech Test Signal (ISTS) multi-talker speech babble with six male and six female North American voices (Holube, Fredelake, Vlaming and Kollmeier, 2010) was aired in free field from a loudspeaker (Fostex, SPA 12, Fostex Electric Company, Akishima, Japan), in order to further add to vocal loading. The ambient babble gradually increased from 55 dBA to 85 dBA during about 30 seconds. Gradual onset was chosen in order to let participants get accustomed with the loud noise. Babble SPL was controlled before testing, with a SPL meter (Svantek, model SV-102, Svantek Inc., Warsaw, Poland) placed at participant ear level. The babble stayed at 85 dBA for the remainder of the task. Given the high sound pressure level of the ambient noise, the time limit for the VLT was set to 30 minutes. The time limit was chosen because the Swedish Work Health Authority has set a time limit for exposure to noise in the work place at 85 dB A for a maximum of 8 hours, and there was no knowing what else participants would be subjected to the day of the VLT (AFS 2005:16, 2005). According to Echternach and colleagues, 10 minutes of vocal loading at 85 dB VL corresponds to the vocal dose of 45 minutes of teaching (Echternach, Nusseck, Dippold, Spahn and Richter, 2014), thus 10 minutes was chosen as a minimum time limit for Papers 2 and 3. If any participant wanted to stop before the lower time limit was up, they were prompted to continue if they could. This happened in 3/50 cases. Echternach time template and the fact that the VLT in the current investigation invoked loud phonation, which potentially could continue for 30 minutes compose the rationale of regarding vocal activity in the VLT *heavy* vocal loading. Paper 1 and 4 had no lower time limit, as the former was a pilot for the set up and the latter examined patients with organic voice disorders, which in some cases caused glottal obstruction, e.g. bilateral Reinke's oedema, making it difficult to draw deep breath during loud speech.

Signal-to-noise ratio was measured online by the test leader in Papers 2–4, through real-time phonetograms (Phog Interactive Phonetography System 2.5 for PC, Hitech Medical, Täby, Sweden) ensuring each participant reached the target sound pressure level, 85 dBA, which they were required to exceed. If they did not match 85 dBA they were reminded to speak louder by the test leader, who stood up, with spread arms, mouthing "louder" ("starkare" in Swedish). The voice signal was recorded using a head-mounted microphone (MKE 2, no 09_1, Sennheiser) calibrated at 94 dB SPL at a distance of 15 cm

from the mouth, converted to 30 cm. In Paper 1 the test leader sat in the booth with the participants to ensure they were making themselves heard.

For more details concerning the original setup of the VLT, see Paper 1, for application with voice patients, see Papers 3 and 4.

Voice accumulation

A voice dosimeter (VoxLog by SonVox AB, Umeå, Sweden) was used to track vocal behaviour in dissertation Papers 1, 2 and 4. Details are described in each Paper. The rationale behind choosing VoxLog is discussed under Methodological Considerations. The aim of voice accumulation was to track vocal behaviour in and out of a VLT. Each participant wore VoxLog during a number of days, measuring vocal behaviour (F0, SPL and phonation time). The VoxLog is a portable hardware device, made up of a neck collar, linked by a lead (90 cm) to a black plastic covered container/voice meter (11x8x2 cm). The device contains a battery which is charged overnight. The neck collar contains an accelerometer and a microphone.

Fast Fourier transformation extracts information from the accelerometer, recording high-speed skin vibrations, present on the neck during phonation during a pre-set time frame (1 minute during long-time accumulation). Candidates for fundamental frequency are selected from spectral peaks that fulfil the power criterion based on the global maximum. Final F0 selection is based on the lowest of these peaks, exceeding a proportion of the total signal power together with its lower harmonics (H1–H3). The total recording time is compared to voicing time, giving relative phonation time. In addition to F0 and relative phonation time the accelerometer gives estimates on cycle dose, i.e. how many hypothetic oscillatory cycles the vocal folds have completed, based on the estimate of fundamental frequency. This feature has not been used in the dissertation Papers, nor was the biofeedback function used.

The microphone does not record the spoken signal, preserving the integrity of the bearer. It is pre-calibrated and records phonatory SPL whenever phonation is registered by the accelerometer. Other noise, recorded when there is no active accelerometer signal, is regarded and recorded as ambient noise SPL. SPL of ambient noise have not been used in any of the current studies. The SPL value is measured in situ, i.e. there is no correction to standard distance of 15–30 cm from the mouth. Södersten and colleagues suggest subtracting 7 dB SPL to compensate for the lack of correction

(Södersten, Salomão, McAllister and Ternström, 2015). Data was analysed in VoxLog Connect (appertaining software by SonVox AB, Umeå, Sweden), which provides mean values for F0 and SPL values, but no standard deviation. Absolute values were extracted to Microsoft Excel (Microsoft, Redmond, WA, USA).

Tracking vocal recovery

The impact of and recovery from vocal loading was tracked through participants' self-assessments of vocal function in a voice activity questionnaire.

The voice activity questionnaire was constructed to shed light on activities and subjective voice symptoms connected to vocal behaviour. This was especially significant for Paper 2, in which the participants' vocal health was mapped during VLT, during work and during leisure. In order to match data from vocal dosimetry to the voice activity questionnaire participants were asked to fill out their questionnaire four times per day and concisely write down what they were doing and where (at home, at work etc.) at every entry. It was filled out before and after completion of the VLT and comprised three parts: (1) Assessment of unspecified voice problems measured on a 100 mm open-ended visual analogue scale (0=no voice problems, 100=maximal voice problems). (2) Ten voice health questions (10VQ) which were aimed at showing sudden fluctuations in specific voice symptoms, see Table 6. These were answered on a 5 point Likert scale. (3) Eight voice health questions (8VQ), which were aimed to reflect underlying voice problems, which would not fluctuate quickly and therefor would not have to be asked as frequently as the 10VQ. See Table 7. These were also answered on a 5 point Likert scale. Objective measurements of changes in vocal function were carried out before and after the VLT and at a follow-up assessment. These measurements are described below.

Table 6: Ten voice health questions (10VQ) evaluating sudden fluctuations in specific voice symptoms (based on VHI-T). These were filled out four times a day: morning, midday, afternoon and before bedtime.

#	10 voice health statements on specific voice symptoms
1	This is my current stress level.
2	My voice feels fatigued.
3	I need to clear my throat.
4	I need to cough.
5	My throat/neck (same word in Swedish: "hals") feels tense.
6	I am hoarse.
7	I am having a hard time making myself heard (like at a party).
8	My voice can suddenly change when I speak.
9	It is effortful to get my voice working.
10	I have a feeling of discomfort in my throat/neck.

Table 7: Eight voice health questions (8VQ) evaluating slow fluctuations in underlying voice problems (based on VHI-T). These were filled out once a day, at bedtime.

#	8 voice health statements on specific voice symptoms
11	I run out of breath when I speak.
12	My voice problems affect my private economy.
13	My voice problems restrict my private and social life.
14	My voice makes it hard for others to hear what I am saying.
15	Others ask me what is wrong with my voice.
16	I feel handicapped because of my voice.
17	I feel left out of conversations because of my voice.
18	I am worried by my voice problems.

Both sets of voice questions were selected mainly from VHI (Jacobson and colleagues, 1997) and VHI-T (Lyberg-Åhlander, Rydell, Eriksson, and Schalén, 2010) with an added question on general stress levels. The set of questions were used by Lyberg-Åhlander and colleagues (2014). Stability of the two sets of questions were checked with test-retest reliability in Paper 2, and not as is reported in the paper, with Chronbach's α. The two sets of questions showed high test-retest reliability (10VQ: $\rho=.86$, 8VQ: $\rho=.84$). In Paper 2 the 10 voice health questions are abbreviated *10VQ* and the 8 voice questions to *8VQ*. The latter set was not used in Paper 3, which holds the correct method description. In Paper 4, the abbreviations were changed from 10VQ to SVQ (shifting voice health questions) and from 8VQ to UVQ (underlying voice health questions). Time point prevalence for filling out the questionnaire are shown for each study in Table 8.

In paper 4 the voice accumulation process with voice dosimetry and voice health questionnaire was applied before phonomicrosurgery, for one

week following phonomicrosurgery and at a long-time follow up 3–6 months following phonomicrosurgery

Table 8: Time point prevalence for participants' filling out the voice activity questionniare in the different dissertation Papers. NA=not applicable.

Paper	VAS	10VQ/SVQ	8VQ/UVQ
1	4 times per day (4 days), spread out at will.	4 times per day, spread out at will	4 times per day, spread out at will
2	4 times per day (7 days). Morning, noon, afternoon, before bedtime.	4 times per day (7 days). Morning, noon, afternoon, before bedtime.	1 time per day (7 days) before bedtime.
3	4 times per day (7 days). Morning, noon, afternoon, before bedtime. Every 15 minutes for 1 h+ 1 time 2 hours following vocal loading task	4 times per day (7 days). Morning, noon, afternoon, before bedtime. Pre and post vocal loading task.	NA
4	4 times per day (7 days). Morning, noon, afternoon, before bedtime. Every 15 minutes for 1 h+ 1 time 2 hours following both vocal loading tasks Not filled out by absolute voice rest group following vocal surgery.	4 times per day (7 days). Morning, noon, afternoon, before bedtime. Only partly filled out by absolute voice rest group following vocal surgery. Questions direclty likned to voice use discarded.	1 time per day before bedtime during long-time measurement after surgery and at long-time check-up.

Objective measurements of vocal function

Objective measurements of changes in vocal function were carried out before the VLT and at a follow-up assessment of participants' vocal function. These measurements are accounted for below.

Acoustic measurements

All audio recordings were performed with participants sitting down, wearing a head-mounted microphone (MKE 2, no 09_1, Sennheiser Electronic GmbH and Co, Wedemark Germany). A calibration process was carried out before each test round, i.e. pre VLT, during VLT and post VLT recordings, standardizing at 94 dB SPL. Voice signals were digitized at 16 kHz with 16-bit resolution.

Changes in fundamental frequency caused by a VLT were examined in the spoken signal of each participant who were recorded reading the passage *Nordanviden och solen* [The North Wind and the Sun] (45 seconds) aloud in Swedish before and promptly after VLT. Fundamental frequency was examined using Soundswell Core 4.0 and Soundswell Voice 4.0 (Saven Hitech AB, Täby, Sweden).

Changes in intensity/phonatory SPL caused by a VLT were examined by use of speech range profiles (SRP) which compose real-time phonetograms using Phog Interactive Phonetography System 2.5 for PC (Saven Hitech AB, Täby, Sweden). The range/area is measured for size and the measure outcome is semitones x dB (STdB), i.e. the number of combined pixels each speaker has scored on a "sheet diagram" with fundamental frequency along the x axis and sound pressure level along the y axis. SRP's were required before and promptly after VLT from participants reading the passage *Nordanvinden och solen* [The North Wind and the Sun] (45 seconds) aloud.

Total voice range profiles were recorded ad modum Hallin and colleagues before the VLT for Paper 1, but were dropped before examining populations with voice pathology, as voice range profiles would cause to great a vocal loading in themselves (Hallin, Frost, Holmberg and Södersten, 2012).

Long-time average spectra (LTAS) gave information on changes in voice source spectral tilt caused by a VLT. LTAS were obtained ad modum Löfqvist and Mandersson (1987).

Perceptual assessments of audio recordings

To evaluate changes in voice quality caused by a VLT, a blind panel of two (Papers 3 and 4) or three (Paper 1) SLP's, experienced and specialized in voice care, assessed each audio recording of *Nordanvinden och solen* [The North Wind and the Sun] by consensus. The method was chosen to minimize problems of low inter-rater reliability (Shrivastav, Sapienza and Nandur, 2005). Each stimulus was presented in free field over a loudspeaker Fostex, SPA 12, Fostex Electric Company, Akishima, Japan). The panel was allowed free time limit and could re-listen to each stimulus as many times as they wished, assessing one participants voice at a time along the parameters in Table 9. Assessment were made with pen and paper ad modum SVEA: Stockholm Voice Evaluation Approach (Hammarberg, 1986) with 100 mm VAS (0=unaffected, 100=deviant in the greatest possible manner).

**Table 9: Seven voice quality parameters assessed by consensus
in perceptual analysis by a panel of trained speech and language pathologists.**

#	Parameters assessed in perceptual analysis
1	Hyperfunction/press
2	Breathiness
3	Instability
4	Roughness
5	Glottal fry
6	Sonority (reversed outcome scale)
7	Grade of overall voice pathology

Visual assessments of digital imaging of the vocal folds

Digital imaging was carried out before and promptly after a VLT. In order to evaluate the following parameters high resolution digital imaging (HRES Endocam, model 5562.9 colour; Wolf, Germany) was recorded by one and the same microsurgeon/phoniatrician throughout the studies, using a 70° rigid endoscope. Table 10 shows parameters assessed based on digital imaging filmed with high resolution. In order to assess the parameters of mucosal movement, shown in Table 11, digital imaging was also performed pre and post VLT.

**Table 10: Seven parameters of laryngeal morhpology assessed by consensus in
analyses of high resolution digital imaging by a panel of trained phonosurgeons/
phoniatricians.**

#	Parameters assessed in visual analysis of laryngeal morphology
1	Glottal shape
2	Glottal regularity
3	Morphological alteration
4	Adduction of both vocal folds (left and right analysed separately)
5	Abduction of both vocal folds (left and right analysed separately)
6	Corniculate tubercle symmetry during phonation
7	Corniculate tubercle symmetry during rest

**Table 11: Four parameters of mucosal movement assessed by consensus in
analyses of high-speed digital imaging by a panel of trained phonosurgeons/
phoniatricians.**

#	Parameters assessed in visual analysis of mucosal movement
1	Wave amplitude of both vocal folds (left and right analysed separately)
2	Wave propagation of both vocal folds (left and right analysed separately)
3	Phase difference
4	Activity in vestibular folds

The assessment protocol was carried out ad modum Bless and colleagues as adapted by Lyberg-Åhlander and colleagues (Bless, Hirano and Feder, 1987; Lyberg-Åhlander, Rydell and Löfqvist, 2012). The setup entails a blind panel of three experienced medical phonosurgeons/phoniatricians assessing each parameter in consensus to minimize any problems with poor inter rater reliability. Assessments are made on a 4 point scale (0=unaffected, 1= slightly affected, 2=moderately deviant, 3=highly deviant).[1]

Phonation threshold pressure

Subglottal air pressure alone does not control phonatory SPL. Features of vocal fold closure (duration and degree) also play an important part. Effective and durable closure of the vocal folds gives subglottal pressure time to build under the glottal plane while the vocal folds are closed (brought together by the Bernoulli effect). If the glottal resistance ($pressure \div flow$) is high, there will be considerable disturbance when the vocal folds are forced to open (Fant, 1982). On the other end of the spectrum, phonation threshold pressure (PTP) is the minimum lung pressure required to achieve phonation when air passes through the glottis. Subglottal pressure is important, because higher subglottal pressure is powerful enough to push through the closed vocal folds (Titze, 1992).

Apart from the pressure and velocity of air being pressed from the lungs, PTP is determined by the thickness and movement velocity of the superficial layers of the lamina propria (Hirano, 1975; Titze, 1992). Typical values during speech are 2–3 cm H_2O (Alton Everest, 2001). Values for PTP are equivalent to intraoral air pressure during an occluded bilabial, voiceless plosive [p] (Holmberg, 1980; Löfqvist, Carlborg and Kitzing, 1982; Smitheran and Hixon, 1981). Thus PTP could be examined non-invasively, using The Phonatory Aerodynamic System (PAS; model 6600; KayPENTAX, New Jersey), which includes hardware, a handheld mask covering mouth and nose, and by software for PC. The hypothesis was that

---

[1] This dissertation has been produced within the context of a collaboration with the technical University of Denmark. Results from high-speed digital imaging recorded before and after a vocal loading task in voice healthy and pathological participants in the current dissertation have been used as a basis for mathematical finite element models of collision forces. The results showed a mathematical model with differences between typical and atypical phonation. They were presented in the PhD thesis *Modelling and imaging of the vocal folds' vibration for voice health*, which was defended at DTU on 5 September, 2016.

PTP would increase as a sign of a change in the superficial mucosal layers, caused by the VLT.

As measurements were difficult for participants to perform and results showed no significant changes in PTP caused by the VLT, it was only examined in Paper 1 and subsequently discarded. Other measurements of vocal efficiency, such as vital capacity, maximum phonation time, phonation quotient of vital capacity to maximum phonation time were not examined, for logistic reasons and in order to narrow the scope of the current study.

Randomisation of voice rest advice

Randomising of patient groups for Paper 4 was conducted through block randomisation. This entailed preparing n=20 sealed envelopes containing voice rest advice for the recovery week (7 days) following phonomicrosurgery (Lachin, Matts, and Wei, 1988). Envelopes either contained advice for absolute voice rest (AVR) or advice for relative voice rest (RVR). Table 12 shows phrasing of voice rest advice, as translated from Swedish. Envelopes were manually mixed into random order by a colleague of the first author. The colleague was in no way involved in the research project. As patients were recruited to the study, the envelope at the top of the pile was handed out by the first author, who was also the test-leader.

Table 12: Voice rest advice applied for one week folliong phonomicrosurgery in two randomised groups: absolute voice rest (AVR) and relative voice rest (RVR). Advice has been translated from Swedish.

Group	AVR	RVR
Voice rest advice ENGLISH	Let your voice heal by being completely silent for the coming week, starting now. Do not use your voice AT ALL. Do not whisper, instead use gestures or writing for communicating. Please note that the voice is often used for more than speaking, e.g. when coughing, throat clearing and during physical activity. Do not use your voice at all.	Let your voice heal by speaking with a gentle, comfortable voice for the coming week, starting now. Do not whisper, do not use loud voice for speaking or shouting and do not sing.
Voice rest advice SWEDISH	Från och med nu och under veckan som kommer ska du låta rösten läka genom att vara HELT TYST och inte använda rösten alls. Du ska inte heller viska, utan istället skriva eller använda gester för att meddela dig. Tänk på att rösten ofta används till mer än prat: hosta, harkling och vid kroppslig ansträngning – försök att inte använda rösten alls.	Från och med nu och under veckan som kommer ska du låta rösten läka genom att tala med en försiktig, bekväm röst. Du ska inte viska, tala starkt, ropa eller sjunga.

Approach to statistical analyses

Statistical analyses in all studies were performed with IBM SPSS Statistics 21–23 (SPSS Inc., Chicago, IL, USA). Throughout the studies data distribution was thoroughly investigated. Statistical analyses (parametric or non-parametric) were chosen based on distribution and type of data. Distribution was explored using a Shapiro-Wilk's test (Razali and Wah, 2011; Shapiro and Wilk, 1965) of skewness and kurtosis (Cramer, 1998; Cramer and Howitt, 2004; Doane and Seward, 2011). Visual inspection of histograms, normal Q-Q plots, and box plots were performed for each outcome measure. Ordinal data was continuously explored using non-parametric analyses, based on the futility of comparing means across subjective ratings. Analyses have stayed close to the data, i.e. foremost direct comparisons have been performed, due to relatively low sample sizes.

Paper 1 is a pilot method study, without group comparisons, thus statistical analyses were merely used to explore effects of the VLT with outcome measures recorded before vocal loading being compared to measures recorded after vocal loading. Wilcoxon's signed rank test was performed, showing effect sizes with r = $(\frac{Z}{\sqrt{n}})$. For papers 2, 3 and 4 all data were examined according to Table 13. As most data were not normally distrubuted and all self-assessments were analysed on the ordinal scale mostly non-parametric statistics have been calculated. Repeated measures were corrected to avoid mass-significance, bu use of strict Bonferroni adjustment.

Table 13: Types of apriori statistical analyses chosen throughout the dissertation.

Data type	Distribution	Comparison	Statistical anslysis	Paper(s)
ordinal interval	not normal	related samples (2) repeated measures (2)	Wilcoxons' signed rank test	1, 2, 3, 4
		related samples (>2) repeated measures (>2)	Friedman's test	2, 3, 4
		independent samples (2)	Mann Whitney U test	2, 3, 4
		independent samples (>2)	Kruskal Wallis test	2, 3
		correlation	Spearman's ρ	2, 3
interval	normal	related samples (2) repeated measures (2)	Analysis of Variance	3
		related samples (>2) repeated measures (>2)	Repeated analysis of variance	2, 3
		independent samples (2)	Paired samples T test	2

Most analyses were carried out using non-parametric tests, meaning power curves were less affected by the assumtions needed for parametric analyses. However, it was important that only large group differences would result in statistically significant differences between groups. This is because if any changes in clinical routine are to be implemented it is important they be effective, especially time-consuming interventions, such as a VLT. Power effect size, mainly for self-assessment scores, was calculated using Cohen's d (Cohen, 1992). However effect size was diffcult to calculate, given few applicaple clinical studies to compare with. As previously mentioned, this problem was compensated by main use of non-parametric tests.

Repeated measures were treated according to Table 13. Due to relatively small sample sizes, it was interesting to do more than just hypothesis testing, thus measurements of repeated records of self-assessments were analysed in a different manner. To gain deeper insight into systematic variation between groups in score change from baseline during repeated measures, data from short-term recovery tracking in Paper 3 and 4 applied standardisation of scores ad modum Atkinson and colleagues. The method takes internal variance through within-group standard deviation at each time point into account (Atkinson, Nevill and Hopkins, 2000). More detail on this method is decribed in Paper 3.

Ethical considerations

Ethical considerations in all current studies comply with the World Medical Association's Declaration of Helsinki (World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects, 2013), ensuring integrity and autonomy of all participants. All studies (1–4) included in this dissertation hold ethical approval by the Regional Ethical Review Board in Lund, Sweden on April 25, 2013 (#2013/174). Some participants recruited for the HLC and HLNC groups in Papers 2 and 3 fall under the ethical approval in a previous study by Lyberg-Åhlander and colleagues (2014), which gained ethical approval by the Institutional Review Board at Lund University, Sweden (#248/2008). Participants in all studies gave informed written consent for participation.

Results in summary

Paper 1

The aim of this study was to setup and test run a clinical vocal loading task (VLT). Results showed a successful method, which attained self-perceived vocal loading and objective changes to vocal function in voice healthy participants, without causing damage to voice function. The method simulated authentic heavy vocal loading. Onset and recovery from self-perceived vocal loading were traceable through a voice activity questionnaire. The wide range of time spent in the VLT (3–30 minutes) was an unexpected finding, indicating the complexity of vocal loading.

Paper 2

This study compared everyday vocal loading/voice use to controlled vocal loading by use of long-time voice accumulation during work, leisure and a VLT in four vocal groups. Patients with functional dysphonia (FD) were compared to women with high everyday vocal loading with voice complaints (HLC), women with high everyday vocal loading with no voice complaints (HLNC) and voice healthy controls (C).

Results proposed reference values for maximum time for phonation during loud, prolonged speech based on loud reading in Swedish as measured with VoxLog around 60–70%, see Figure 6. Results also showed that vocal loading is not only dependent on prolonged phonation time at high intensity levels, but that it also seems to be reliant on prolonged phonation time at high fundamental frequencies. Lastly results showed that women with high everyday vocal load who experience voice problems (HLC) reported strain-induced voice problems during a VLT and during work. This set them apart from patients with functional dysphonia (FD) who exhibited voice problems in all conditions, even during leisure. It may explain why women with voice

problems associated with their work environment (group HLC) do not seek voice therapy.
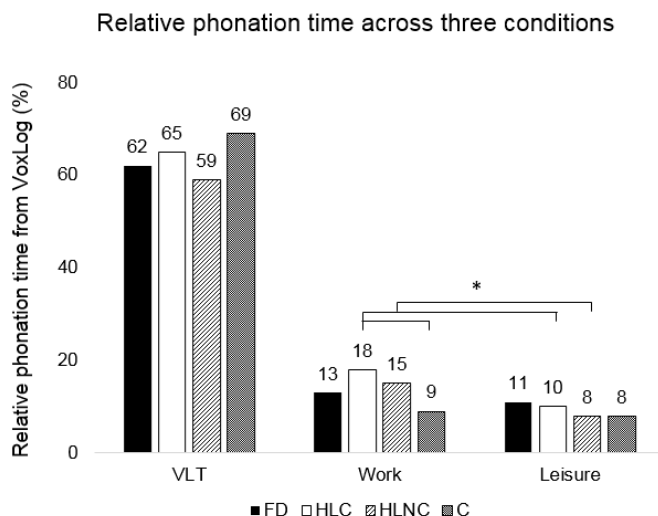
Figure 6: Relative phonation time (%) measured by VoxLog accelerometer under three conditions: vocal loading task (VLT) work and leisure and a vocal loading task across four (n=4) vocal subgroups (FD, HLC, HLNC and C). Relative phonation time in VLT was significantly higher than work and leisure for all groups, other significant group differences and differences within groups are marked with bars. Part of results from Paper 2.

Paper 3

In this Paper the same groups were studied as in Paper 2. Paper 3 explored the differences in recovery from a VLT in the groups. Short-term recovery was examined using participants' self-assessments of unspecified voice problems with a 100 mm visual analogue scale (VAS) every 15 minutes for the first 2 hours following the VLT. Long-term recovery was tracked by self-assessment of specific voice symptoms with voice health questions on a 5 point Likert scale four times per day during the days following the VLT.

Results show that short-term recovery is slower for patients with functional dysphonia than controls. However, their long-term recovery course, tracking more specific voice problems, is equivalent to others', implying short-term recovery is more important to track in the voice clinic. Patients were worse affected by the VLT than others. They presented significant perceptual changes toward hyperfunction/press and general voice

pathology due to vocal loading, as well as affected vibratory patterns seen in digital imaging of the vocal folds, compared to others. Typical mucosal movement was re-established at check-up, two days after the VLT. Women who are used to everyday vocal loading and who do not experience voice complaints (subgroup HLNC) are better than other groups at reacting adequately to, or identifying, heavy vocal loading, which may be due to coping factors such as shifting toward less hyperfunction during vocal loading, see Figure 7.
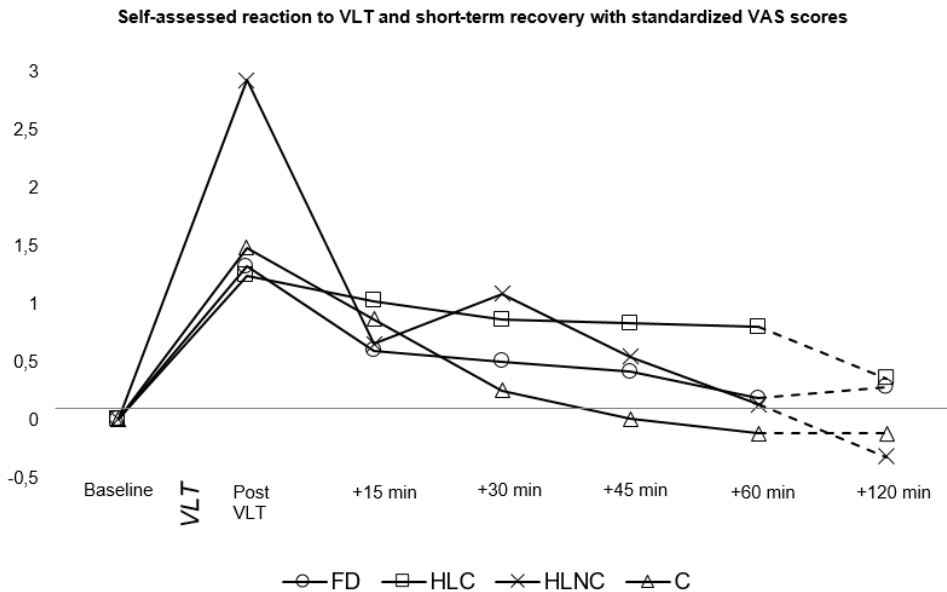
Figure 7: Standardized change in mean self-assessments of unspecified voice problems measured with 100 mm visual analogue scale (VAS) scores over time showing the immediate reaction and short-term recovery following a vocal loading task (VLT) in four vocal subgroups. Variance (within subject standard deviation) and differing levels at baseline and repeated measures taken into account. Standardized mean scores were calculated ad modum Vogel and Maruff, 2014. Part of results from Paper 3.

Paper 4

This paper explored differences in vocal function and vocal recovery in two groups undergoing phonomicrosurgery for benign lesions in the vocal fold mucosa. Both groups received voice care advice for the 7 days following surgery. One group was recommended absolute voice rest and the other relative voice rest.

Results show overall evidence to support relative voice rest over absolute voice rest following surgery of benign lesions in the vocal fold mucosa on the following rationale: (1) Patients showed difficulty in complying with absolute voice rest to a higher degree than relative voice rest, both regarding self-assessed compliance and objective measurements of relative phonation time. Patients phonated to low extent when recommended absolute voice rest, but they were not completely silent, see Table 14. (2) There were no significant group differences between absolute and relative voice rest groups in the short or long-term recovery of vocal function according to visual and perceptual assessments. However, there were significant short-term improvements in assessments of vocal fold digital imaging within the absolute voice rest group, which were not evident for the relative voice rest group. Both groups improved significantly regarding glottal shape, regularity and overall morphology and regarding press, breathiness and sonority. (3) Patients recommended for relative voice rest showed significantly better vocal stamina/vocal loading capacity and immediate recovery from vocal loading, as well as significantly improved within-group self-assessment of voice problems at long-term check-up.

Table 14: Participants' compliance with absolute or relative voice rest examined with VoxLog's relative phonation time (%) during 7 days' voice accumulation following phonomicrosurgery. Part of results from Paper 4.

	Participant	Phon time (%)		Participant	Phon time (%)
Absolute voice rest group	F1	1	**Relative voice rest group**	F2	4
	F3	0		F5	4
	F4	0		F6	3
	F7	14		F8	9
	F9	2		F10	28
	F11	0		F12	6
	F14	3		F13	1
	M2	2		M1	8
	M5	1		M3	0
	M6	2		M4	11

Discussion

This section revisits research questions and hypotheses of the dissertation. Within each paper's discussion clinical implications are presented. After this review follow discussions of important elements which have not been discussed in great detail in each paper. Subsequently follows a discussion of study limitations and finally, suggestions for future research are presented.

Revisiting research questions and hypotheses

Paper 1

In Paper 1 the question on how to investigate vocal loading and recovery in a clinical setting was posed. A vocal loading task (VLT) was developed, including loud reading in a noisy environment. This task was later changed to improve monitoring signal-to-noise ratio and tracking recovery. After changes the task was applied in Papers 2, 3 and 4. An important aspect of the setup was to let participants set the time limit for the VLT. This aspect was chosen because participants were meant to attain vocal fatigue, without permanent damage to the voice function, while vocal loading was in focus. The maximum time was 30 minutes and the loud phonation during that time was meant to reflect real life vocal loading during a whole working day. The same rationale was chosen by Echternach and colleagues, who concluded that 10 minutes' heavy vocal loading held an equivalent vocal dose to 45 minutes of real teaching (Echternach and colleagues, 2014). The hypothesis of this study was confirmed: *It is possible to setup a VLT that will cause vocal fatigue in voice healthy participants, without causing permanent damage to the vocal function*.

Paper 2

This study examined vocal behaviour and self-assessed vocal health across three conditions in patients with functional dysphonia as compared to groups of women with varying degrees of voice problems.

The hypotheses of this study were only partly confirmed. Hypothesis (1) *Female patients with functional dysphonia use their voice to a greater extent than women who also experience voice problems, but who do not seek medical voice treatment* was rejected. There were no significant group differences in relative phonation time in any condition (VLT, work and leisure). An unforeseen finding in this study was high fundamental frequency of phonation during heavy vocal loading (268–315 Hz across groups), with women with high everyday vocal load with voice problems they did not seek help for (group HLC), showing similar fundamental frequency also during work. This finding connects phonation at high pitch, not only at high intensity to increased vocal effort. This phenomenon was tracked well through voice dosimetry and is highly clinically relevant.

Hypothesis (2) *Patients with functional dysphonia report self-assessed voice problems to a greater extent than women in three groups on a spectrum of functional voice problems* was confirmed. During leisure the patients' self-assessed unspecified voice problems, measured with VAS, to a significantly higher extent than all other groups, but interestingly not in the heavy vocal loading situation brought about by the VLT. Specific voice symptoms scored significantly higher for patients than all other groups in self-assessments during leisure and work (and higher than two other groups during the VLT).

Paper 3

This study examined the impact of heavy vocal loading on voice function, and recovery from heavy vocal loading in patients with functional dysphonia as compared to groups of women with varying degrees of voice problems (same participants as Paper 2). There were three hypotheses in this study: (1) *Female patients with functional dysphonia will be worse affected by a VLT than women in three groups on a spectrum of functional voice problems.* The first hypothesis was confirmed, as patients with functional dysphonia increased perceived general voice impact and hyperfunction/press, as well as showing changes in vibration patterns of the mucosal tissue caused by vocal loading. (2) *Patients with functional dysphonia will terminate the VLT sooner*

than other groups. This hypothesis was rejected, as there were no significant differences in VLT participation time across groups, nor were patients prone to adequately identify heavy vocal loading, which women with high everyday vocal loading and no voice complaints (group HLNC) were. *(3) Recovery from vocal loading will take longer for patients with functional dysphonia compared to other groups.* This hypothesis was confirmed for short-term recovery over 2 hours following the VLT, as they did not return to baseline self-assessment of unspecified voice problems after VLT. Nor did they identify heavy vocal loading by a distinct reaction to it. The hypothesis was rejected for long-term recovery, when self-assessment of specific voice symptoms was tracked during the days following the VLT. There were no significant group differences, although voice healthy controls only took 7 hours to return to baseline, while the functional voice patients took 17 hours and the two groups with everyday vocal load took 20 hours. This implies that short-term recovery is more clinically relevant to track than long-term recovery.

Paper 4

This study investigated differences in compliancy with absolute or relative voice rest advice during 7 days following phonomicrosurgery in two randomised groups. It also explored vocal recovery in the two groups. There were two hypotheses in this study. (1) *It is difficult to comply with a recommendation of absolute voice rest following phonomicrosurgery*. This hypothesis was confirmed. Although patients in the absolute voice rest group phonated significantly less than the patients in the relative voice rest group, they were not completely silent, see Table 14. (2) *It will benefit healing processes in the vocal folds to let patients comply with relative voice rest advice instead of absolute voice rest for one week following phonomicrosurgery.* This hypothesis was partly confirmed. Short-term recovery (7 days after surgery) showed significantly improved assessments of vocal fold digital imaging in the absolute voice rest group, which was not evident for the relative voice rest group (see Paper 4). However, at long-term check-up (3–6 months following surgery) patients in the relative voice rest group had significantly improved their ability to identify heavy vocal loading, their vocal stamina (loading capacity) and all self-assessments of voice function. These positive changes were not evident at long-term check-up for

the absolute voice rest group, thus relative voice rest seems more beneficial for long-time vocal recovery.

The concept of vocal loading

The setup of the VLT in the current investigation prompts an important discussion of the concept of vocal loading, including three key elements: (1) *vocal loading,* (2) *vocal effort* and (3) *vocal fatigue*.

Many studies examine vocal loading, for example though VLT's (e.g. Doellinger and colleagues, 2009; Echternach and colleagues, 2014; Fujiki, Chapleau, Sundarrajan, McKenna, and Sivasankar, 2016; Hanschmann and colleagues, 2011; Hunter and Titze, 2009; Kelchner and colleagues, 2006; Lohscheller and colleagues, 2008; Remacle, Morsomme, Berrué and Finck, 2012; Sherman and Jensen, 1962; Södersten and colleagues, 2005). It is noteworthy that most researchers only imply what is meant by the concept, i.e. what is examined. The basic problem seems to be whether to regard vocal loading purely as voice use, as a function of the vocal instrument which is an evolutionary product of both tracheal protection and phonation, or to see it as something detrimental to the voice – which has sometimes been regarded as vocal overload (e.g. Vilkman, 2004; Ternström, Bohman, and Södersten, 2006). In terms of semantics, "load" can mean the *amount of work* assigned to or performed by a mechanical system, as well as it can signify something that *weighs down or oppresses like a burden*. The distinction is important to make when it comes to sustainability of a speaker's vocal function. The concept is often referred to as negative or cumbersome for the speaker and there is research put forth on safety limits for vocal loading (Švec, Popolo and Titze, 2003; Titze, 1999). Titze quite simply equates vocal load to vocal dose, with the following definition:

> *Vocal dose (also called **vocal load**) is the acoustic vocal power integrated over time. Thus, the daily occupational vocal dose is the acoustic vocal power integrated over the performance time.* Titze (2001, p. 4)

Vocal power relates to the sound pressure level of the phonatory signal (Alton Everest, 2001) and is determined by the subglottal pressure of air from the lungs passing through the glottis with power dependent on glottal resistance. This definition of vocal load basically entails a reaction in the vocal organ, leading to phonation at any power level. In order to achieve

increased vocal power, glottal resistance is loaded, which increases vocal effort. Vocal effort is a physiological response to loud phonation leading to an increase of SPL in the vocal signal. It can be measured objectively as a change in signal sound pressure level, and subjectively as changes in self-perceived vocal effort. Vocal effort increases when auditory feedback decreases (Bottalico, Graetzer and Hunter, 2016). The notion stands in contrast to vocal comfort, that to some extent increases with reverberation time of noise in a room (Pelegrin-García and Brunskog, 2012).

This fact builds phonation at high SPL levels into the vocal loading concept. An increase in vocal effort can be achieved by increased vocal loading, i.e. increased vocal power for a prolonged stretch of time.

The research community seems to have implicitly agreed: vocal loading is detrimental to the vocal function, as increased vocal effort leads to different kinds of vocal fatigue. For example, Titze has proposed effects such as laryngeal muscle fatigue, which was the main aim in the current VLT setup, and laryngeal tissue fatigue, meaning damage brought about in the lamina propria due to heavy vocal loading (Titze, 1999). Evidence of the latter was shown in patients with functional dysphonia in Paper 3. However, this becomes clinically confusing, as, according to Titze's definition above (Titze, 2001), no damage is implied by vocal loading in itself. Only mere voice use is implied. Echternach and colleagues subtly and perceptively make an addition to the concept witch they call *vocal loading capacity*, including the question "How much can you take?" (Echternach and colleagues, 2014). In Paper 4 vocal loading capacity is implied when the term vocal stamina is used.

Vilkman provides a more complex definition of vocal loading. He adds nuance by discussing the shift from vocal load, via vocal fatigue to vocal overload, which can be exacerbated by additional loading factors, such as bad air, poor acoustics and ergonomics. Vocal overload, according to Vilkman, manifests by increased subjective complaints, rather than objective changes of vocal behaviour (F0, SPL, time). Vilkman also adds grading to vocal loading and breaks the process into three stages: (1) vocal warm-up (also discussed by Hunter and Titze, 2009), (2) vocal fatigue, (3) voice rest (Vilkman, 2004). Possibly, vocal warm-up is what was shown in the HLNC group in Paper 3, who did not shift toward hyperfunctional phonation, but instead seemed to decrease the press of laryngeal muscles, to cope with heavy vocal loading, without dropping phonatory SPL. How they achieved it is unclear. It is possible participants in this group have an anatomical

constitution which gives advantages regarding effective resonance, however this has not been explored. Research on this topic is called for in Paper 3.

In this dissertation vocal loading has been determined to entail using the voice. The term *heavy vocal loading* was consciously chosen, to imply vocal effort caused by prolonged voice use at high intensity levels, causing vocal fatigue. The mode of testing vocal loading in this dissertation has assumed vocal loading to equal phonation, i.e. using the voice, but has taken into account what might significantly affect vocal loading, by making participants increase their vocal effort through loud, prolonged phonation. A novel and clinically important aspect of the VLT used in this dissertation is the fact that vocal loading capacity, or vocal stamina, was highlighted, by having participants decide when they experienced vocal fatigue and subsequently terminate the VLT. The setup sheds light mainly on self-assessment of vocal complaints. Figure 8 shows a schematic flow chart of how the process of vocal loading and recovery have been observed and used in the current dissertation.
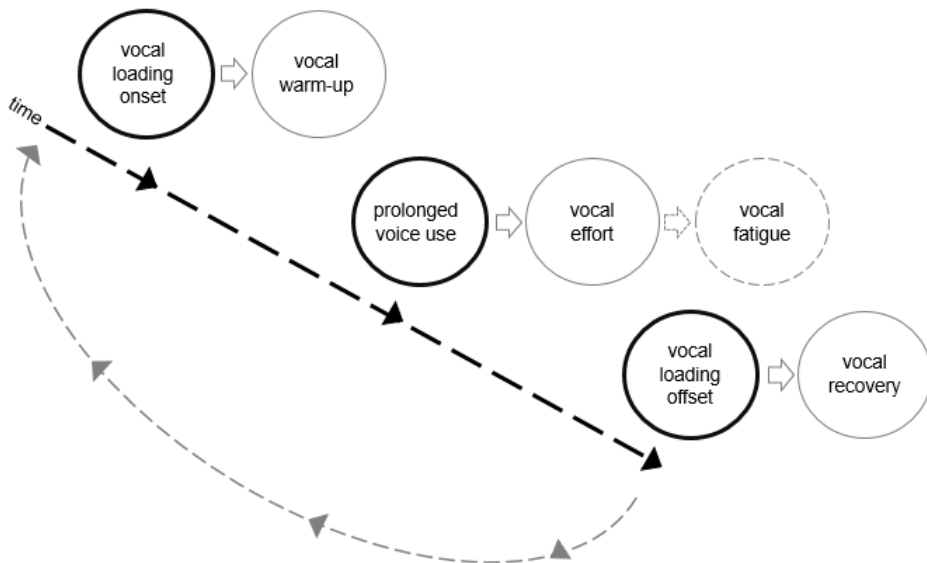
Figure 8: Flow chart of the processes involved in vocal loading and recovery used in the current dissertation.

Results of the current dissertation (all papers) suggests that when discussing vocal loading in a clinical setting, one cannot neglect to take the speaker's subjective experience into account. Papers 3 and 4 show that not all participants push their voices all the way into a state of vocal fatigue. This is why the circle around *vocal fatigue* is dashed in Figure 8). However, it seems there cannot be vocal loading in a noisy environment, without a reaction of some sort from the speaker. She is expected to subconsciously adapt to the situation, through the Lombard effect, by raising her phonatory sound pressure level to match the noise around her, causing the speaking pitch to also rise (Lane and Tranel, 1971). She may then, as is shown in Paper 3, react to the vocal loading context, by correctly identifying it and ceasing with the activity before vocal overload occurs. In Paper 3, this reaction was shown in a group of women who were used to a high vocal load, or dose, from their work and who did not experience any voice problems (the HLNC group). Their experience may have given them an ability for recognition schema, i.e. the ability to interpret propriosensory information from the voice, recognize when vocal ability is saturated, and set in a coping skill, e.g. by ceasing with the vocal behaviour or by reducing hyperfunction in phonation, as was seen in the HLNC group in Paper 3 (Schmidt, 1975). This topic is further discussed under *Vocal loading, recovery and intervention*. This is a solution-focused coping skill, opposed to the problem-focused coping strategies shown in teachers by Zambon and colleagues (Zambon, Moreti, and Behlau 2014). Such a reaction was not shown in patients with functional dysphonia. It could be they are too used to heavy vocal loading. Another explanation could be the connection presented by O'Hara and colleagues, between functional dysphonia and perfectionism, which would drive these patients not to give up, even though they are hurting themselves, so called problem-focused coping (O'Hara, Miller, Carding, Wilson, MacKenzie and Deary 2009, Zambon and colleagues, 2014). Unpublished data from papers 2 and 3 show that it was not necessarily the participants who experience severe vocal fatigue who terminated the VLT before the time was up. Some participants stated that they stopped *before* it got too bad, indicating good coping skills.

In summary, based on previous research and results in the current dissertation, it is suggested that vocal loading be clinically discussed in terms of heavy vocal loading, if subjective increase in patients' vocal effort and vocal fatigue are implied. As subjective assessments are in focus, it is also important to take coping skills of different kinds, into account. They will help reduce patients' struggle with speaking in noisy environments. An important

coping skill shown in papers 3 and 4 is an ability to identify heavy vocal loading and react to it by reducing hyperfunctional phonation (Paper 3).

Vocal recovery in the voice clinic

Vocal recovery and heavy vocal loading can be considered as opposite sides of the same coin. Vocal recovery includes wound healing processes in the vocal folds, but that is not all. Little is yet clinically established regarding main driving forces behind vocal recovery. One part of the problem may be the lack of uniform definitions of vocal loading and vocal fatigue, as previously mentioned. Another major difficulty is posed by the complexity of the research area focussing on vocal recovery, with participants constantly showing a high degree of interpersonal and contextual variation and with microscopic changes taking place in the lamina propria, that are clinically difficult to evaluate. Recovery is dependent on a multitude of objective and subjective, shown in Figure 9.

Factors behind vocal recovery

based mainly on Hunter and Titze 2009, Kahneman 1973, Roy and Bless 2000, Vilkman 2004, WHO 2001

vocal recovery

physiological factors

psychological factors

contextual factors

cell structure of lamina propria
external muscles of the larynx
internal muscles of the larynx
physical coping skills
breath support
resonance

mental coping skills
personality traits
stress

communication context
physical context
cognitive load
time
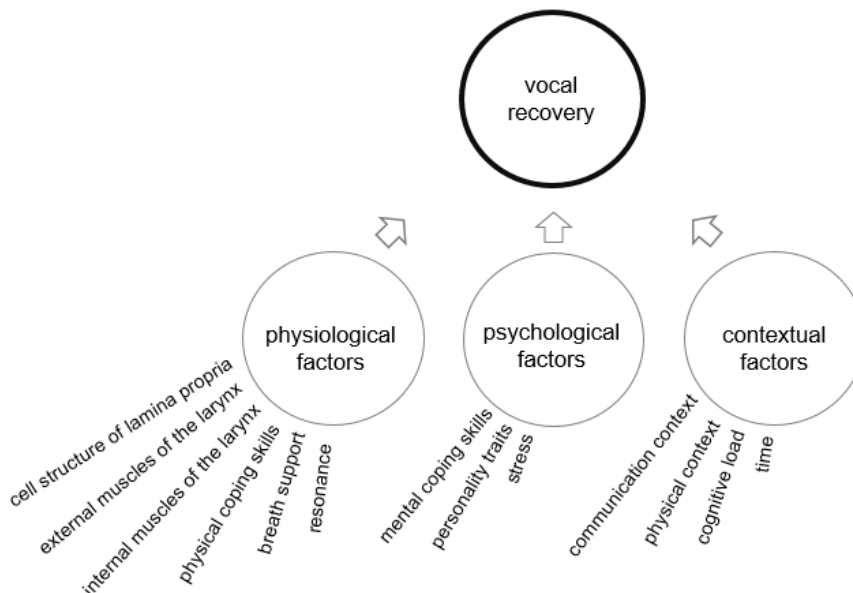
Figure 9: Factors behind vocal recovery

The current investigation has kept a patient perspective on recovery, with basis in the ICF model (WHO, 2001, see Figure 3), and has mainly focused on self- assessments of voice function. However, objective vocal recovery has not been disregarded at all. It is dependent on physiological factors, such as improved function of intra and extra laryngeal muscle (Hunter and Titze, 2009; Vilkman, 2004), healing metabolism and cell structure in the covering lamina propria, especially in the extracellular matrix (Verdolini Abbott, Li, Branski and colleagues, 2012; Li and colleagues, 2008) and functional breath support (Colton, Casper, and Leonard, 2011). In Paper 3 patients with functional dysphonia presented vibratory changes to vocal fold tissue due to heavy vocal loading, probably causing their slow short-term (2 hour) recovery. No other group in any other paper presented this deterioration, indicating a tangible need in patients with functional dysphonia for adequate vocal recovery time in their everyday life. In Paper 4 the absolute voice rest group showed significant short-term improvement of wound healing, however this group difference disappeared at long-term check-up 3– 6 months following phonomicrosurgery. This implies quicker short-term wound healing when laryngeal muscles and lamina propria are still, but long-term wound healing in favour of relative voice rest, which is much easier for patients to comply with.

Vocal recovery is dependent on coping skills, both physical, such as level of individual vocal training, and mental, such as being mentally armed for the speaking task at hand (Zambon and colleagues, 2014). Recovery is dependent on other psychological factors, such as personality traits (Baker, 2008; Roy and Bless, 2000; Roy, Bless and Heisey, 2000) and reducing stress levels (Kooijman, de Jong, Thomas and colleagues, 2006). Paper 3 showed evidence of physical coping skills in the group with high everyday vocal load and no voice complaints (HLNC), who identified heavy vocal loading and shifted toward hypofunction in order to save their voice. Unpublished data from Paper 3 also showed support for mental coping skills, as one patient with functional dysphonia who did not terminate the VLT, reported "I did not terminate the task, it was hard work, but it's like running; it's easier to stop, but that's not the goal". Paper 2 showed higher levels of stress, although not significantly higher, for patients with functional dysphonia across all explored conditions (VLT, work, leisure).

Recovery is also dependent on the context of communication and the physical context, in which phonation and research are taking place (WHO, 2001). Depending on who patients are talking to (children, adults, groups,

singles, even animals), and what they are talking about, different physiological and cognitive demands are out upon their voices (Vogel and Maruff, 2014). This was evident in Paper 4, when participant M2 reported great difficulty in complying with absolute voice rest, as he could not communicate with his dogs. Physical context includes room acoustics, air quality and humidity (Fujiki and colleagues, 2016; Pelegrin-Garcia, Smits, Brunskog and Jeong, 2011; Vilkman, 2004). These issues are discussed further under *Measurement instability.*

Gender is an interesting factor behind vocal recovery, as it may be discussed as a physiological, psychological and/or contextual factor. Biological differences between cisgender women and men on a group level contribute to differences in physiological vocal production (Titze, 1994b), and certain organic voice pathologies affect women more than men, e.g. Reinke's oedema (Paper 4). A physiological or indeed a psychological personality trait factor may make women more prone than men to report to having trouble making themselves heard in noisy environments (Södersten and colleagues, 2005). Gender may also be placed under contextual factors, as male, or low pitched, voices are still preferred e.g. for voters in political campaigns (Anderson and Klofstad, 2012), and because room reverberation is dependent on reflections of soundwaves produced at a specific pitch, i.e. voice produced at 100 Hz will be differently reflected by the same room than one at 200 Hz (Alton Everest, 2001).

This leaves a great deal of factors to control before attempting research in the area. The question remains: how is it possible to test all these factors? One highly important instrument, well highlighted in this dissertation, is patients' self-assessment of voice function. This may be the most powerful instrument the voice clinic holds for tracking vocal recovery, as it encompasses perceived physiological, psychological and contextual factors (Carding and colleagues, 2009). Depending on what part of the voice function is explored, it is certainly not unimportant to also systematically investigate targeted physiological, psychological and/or contextual factors contributing to vocal recovery, as they may not always correlate well to subjective ratings (Cheng and Woo, 2010). But combined with voice dosimetry, self-assessments give excellent support for the clinician to map and track patients' vocal behaviour.

Systematic tracking of vocal recovery using patients' and participants' self-assessments and voice dosimetry, combined with traditional clinical tools was applied in papers 1, 3 and 4. Results from Paper 3 were comparable to a study by Hunter and Titze. They hypothesised that vocal recovery would

follow one of two paths: it would either be brief, thus resembling acute wound healing or more prolonged in time, thus resembling chronic wound healing. It was hypothesised the latter would take place if participants were not given enough time to rest from vocal loading. This hypothesis may be confirmed by results in Paper 3 which showed short-term recovery according to self-assessments to be slower for patients with functional dysphonia. According to Hunter and Titze this could be due to fluid redistribution in the lamina propria, which also may explain the changes in vibratory patterns caused by heavy vocal loading in functional voice patients in Paper 3 (Hunter and Titze, 2009). Similar recovery patterns are discussed by Verdolini Abbot and colleagues, who mention that if injury to the lamina propria is repeated, reparative (scarless) wound healing or constructive (scarring) healing may occur (Verdolini Abbott and colleagues, 2012). The latter is more common in humans and this restorative process is the only one clinicians have any chance of clinically seeing through laryngoscopy (Gray, 2000).

A crucial aspect of clinical vocal recovery is whether or not vocal fold physiology benefits from total voice rest. As mentioned in the introduction, the research shows a running hypothesis that relative voice rest may be implemental, but clinical routines have no yet been changed, as the hypothesis has not been confirmed. The most common clinical routine following phonomicrosurgery is absolute voice rest for 7 days (Behrman and Sulica, 2003). In a later survey study, it was shown that relative voice rest is also administered post surgically, but to a lesser extent and without proper definition of the concept (Coombs and colleagues, 2013). The lack of data supporting length and type of voice rest was identified by Ishikawa and Thibeault, who conclude a need for increased knowledge on vocal fold healing and what effect mechanical stress has on the vocal fold mucosa (Ishikawa and Thibeault, 2010).

There is a traditional belief that voice use with onset directly following phonomicrosurgery would dampen healing processes in the mucosal tissue, and lead to further scarring (Roy, 2012). To date there have been no randomised studies published, which discuss the difference of total, or absolute, voice rest compared to relative voice rest with onset promptly following phonomicrosurgery of benign lesions in the vocal folds. Paper 4 may be the first study to make this comparison. Only different amount of days with absolute voice rest have been compared (Kaneko and colleagues, 2016; Kiagiadaki and colleagues, 2015). Paper 4 presents data supporting relative voice rest above absolute voice rest. The rationale is not based on significantly better wound healing processes in the relative voice rest group

compared to the absolute voice rest group. In fact, there were no differences regarding mucosal wound healing at long-time check-up 3–6 months following phonomicrosurgery. Instead the rationale is based on the fact that the relative voice rest group achieved significant improvement of vocal stamina, or vocal loading capacity at long-term check-up. They were better at identifying heavy vocal loading at check-up and they stayed significantly longer in the VLT than before phonomicrosurgery, without incurring any permanent damage. Finally, the relative voice rest group self-assessed their vocal function significantly better at check-up compared to before surgery. None of these beneficial changes were present in the group complying with absolute voice rest. In line with Rousseau and colleagues (2011), it was also shown that the absolute voice rest group struggled to comply with their voice rest advice, a problem that was much less evident in the relative voice rest group. Compliance with voice rest advice was systematically explored using voice dosimetry, as recommended by Misono, Banks, Gaillard, Goding and Yueh (2015).

Limitations of the current investigation

It is important to limit scientific studies to hypotheses that can be confirmed or rejected by the chosen method. Limitations within each study of this dissertation have been cumbersome. The main problem lies in the ambiguity in a concept essential to the dissertation: "vocal loading". The concept has not yet been properly established, nor have the limits of vocal load and vocal overload been thoroughly explored. A pronounced difficulty with this kind of exploration is the ethical aspect of demanding participants to push the limits of safe phonation as naturally their voice function should not be harmed!

For the purpose of clarity, the definition "vocal dose" is meant when vocal loading is mentioned in this dissertation, and *heavy* vocal loading is implied with the VLT (se rationale and discussion under *The concept of vocal loading*). This could have been made clearer in each paper. The aim of this dissertation was to try to unlock the vocal loading concept: to subject participants to a VLT, evaluate the method in Paper 1 and to investigate different parts of the population afflicted with functional and benign organic voice pathology in the subsequent papers. This has implied many different explorations on many levels, which may have made the scope for each paper

imprecise. One may argue that this dissertation has fallen for a temptation that has stricken others before, a problem phrased well by Ternström and colleagues:

> /.../ the act of phonation has so many control aspects and degrees of freedom that an exhaustive description of phonatory phenomena remains elusive. The clinician, the teacher, the training performer and the voice scientist all need ways of presenting multidimensional voice data as succinctly as possible. Ternström, Pabon and Södersten (2016, p. 268)

Risk of bias

Impact afflicted by each individual person participating in experimental setups for medical research cannot be overlooked. Selection bias is difficult to avoid and will wholly affect our findings, from methods, to results and conclusions (Kahan, Rehal, and Cro, 2015). It is also important to take into account how medical experiments may impact participants in the short and the long-term. In ethical consideration, medical voice researchers must respect the connection between mechanical and psychological voice function, lest participants are affected in unexpected ways. Clinical research based on human participants runs the risk of participants affecting the outcome of trials based on selection, or sampling bias. Selection bias entails the risk of failing to achieve true randomisation in different groups and not exploring a phenomenon in a true sample of the targeted population (Kahan and colleagues, 2015). One way of decreasing this risk in the current investigation, was not letting the test-leader recruit patients included Papers 2,3 and 4, but instead seeking the help of other voice professionals (SLP's and phoniatricians) than the test-leader. Participants in the current studies were required to come in to the voice clinic for two extra visits, that took approximately 2 hours each time. This probably means only very interested patients took part in the studies. This may involve people who have experienced problems with their voice or altruistic individuals who want to contribute to further understanding of the human voice, people who perhaps have been helped medically and who have relied on medical research for increased quality of life. The consequences of potentially recruiting patients and control participants with atypical vocal function was minimised by a vocal screening process (laryngeal exam and informal perceptual evaluation) carried out with all participants before entering. Patients in Paper 4 served as

their own voice healthy controls at long-term check-up and in papers 1–3 VLT as an intervention compared individuals to themselves within groups.

Research test-leaders run the risk of affecting participants and subsequently affecting results. This is not least easily done when data collection is stretched out in time, even though the same test leader conducts all data gathering. In the current study great care was taken on exact reiteration of oral instructions, giver over 2,5 years. Also all participants in Paper 4 were told not to tell test-leaders which voice care advice they had received until after the experiment was finished.

Block randomisation of envelopes containing voice rest advice limited the blindness of the study in Paper 4 towards the end of data collection. As the sample size was quite small there may have been a bias when authors examined the last patients. Test leaders could easily, by process of elimination, figure out what voice rest advice the patient had followed. At long-time follow up (3–6 months following phonomicrosurgery) it was impossible for test-leaders to remain blind. However, as perceptual and visual assessments of data were carried out by others than the test-leaders this bias is considered to be minimised.

Technical limitations

In the VLT, participants were seated in front of a pane of glass (see Figure 5). Reflections of sound caused by the glass were not controlled for. Reflections of direct sound may have caused unnaturally augmented feedback of the participant's own voice (Bottalico and colleagues, 2016). Care was taken not to cause any occlusion effect of the ambient sound, which is why it was aired in free field during the loud reading. The reflective pane of glass may have helped participants hear their own voice well, giving them increased acoustic support from the room (Bottalico and colleagues, 2016; Pelegrin-García, Brunskog, Lyberg-Åhlander, and Löfqvist, 2012; Lyberg-Åhlander and colleagues, 2011). This may give participants with functional voice problems an advantage they do not have outside a laboratory setting, as Lyberg-Åhlander and colleagues have shown that teachers with voice problems are more perceptive of the room acoustics (Lyberg-Åhlander, Rydell, Löfqvist, Pelegrin-García, and Brunskog, 2015).

The chosen voice dosimeter for tracking vocal behaviour was VoxLog (SonVox AB, Umeå, Sweden). There are some disadvantages when using VoxLog for research. There is standard calibration of the neck-worn microphone measuring sound pressure level of the bearer's voice signal and

surrounding sounds. It is not calibrated for each use, potentially leading to less specificity in sound pressure level measurements. Also, the placement of accelerometer compared to e.g. the Ambulatory Phonation Monitor (APM) (KayPentax New Jersey, USA), and microphone (no microphone in APM) is not as stable in the VoxLog. The APM requires the bearer to glue the accelerometer to the sternal notch, giving it measurement stability, whereas the accelerometer and microphone in the VoxLog are placed in a neck collar, which could compromise measurement stability. This was especially evident when developing the method of vocal loading and testing applicability of the VoxLog to measure vocal behaviour during the VLT in Paper 1. On the other hand, these disadvantages make VoxLog practical for clinical use. It can be used for several days without the need to calibrate, facilitating logistics of long-time voice accumulation. APM had been used in a previous study (Lyberg-Åhlander and colleagues, 2014) and allowed only shorter measurements of each participant, due to the need of calibration before each measurement period. During data collection for Paper 1, original VoxLog neck collars were used. They were ill-fitting (often too big/loose) for female participants, which was temporarily remedied with neck padding. As Paper 2 and 3 would examine female patients with functional dysphonia, it was important to adapt the neck collar to fit the purpose. After communication with SonVox AB, the collars were substituted to better fit the target group (which is also more relevant for clinical studies, as most patients who suffer from functional dysphonia are women (Fritzell, 1996), see Figure 10.
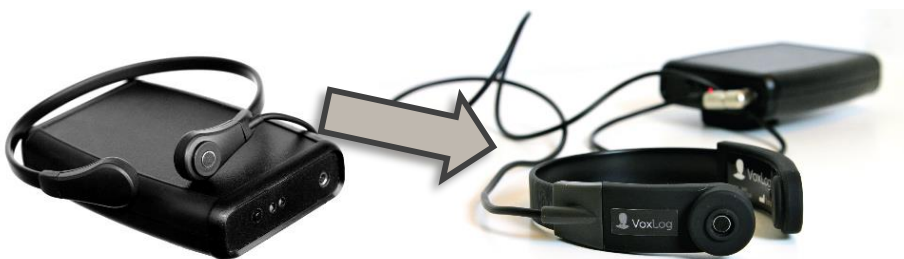
Figure 10: Voice dosimeter VoxLog with two different neck collars: large and unadjustabe to the left, exchanged for adjustable, more placement-stable to the right. Photo with permission from Sonvox AB.

Another suggested change, was to place the accelerometer on the same side of the neck collar as the microphone. This would further increase

measurement precision. This change was carried out by SonVox AB before data collection for papers 2, 3 and 4 were carried out.

Relative phonation time will not be equal to speaking time, due to voiceless segments in the speech signal – for North American English the ratio is about 1:2 (Löfqvist and Mandersson, 1987). It is important to map vocal behaviour, as, according to Misono and colleagues there is only moderate correlation (r=.62) between self-reported voice use and actual relative phonation time during long-time voice accumulation (Misono and colleagues, 2015). Three (n=3) of the participants with functional dysphonia in Paper 2 and 3 were interviewed in a master's thesis, co-supervised by the author. Data showed these participants guessed their speaking time during a working day was as high as 95%, which would correspond to phonation times found in the VLT (60–70%), i.e. much higher than actual phonation time during work, which was about 20% (Hammar, Whitling, Lyberg-Åhlander and Rydell, 2015; Paper 2). It would be interesting to further explore the clinical importance of the relationship between actual phonation time and self-perceived phonation time by instructing participants to consciously take vocal rests throughout their working days. Increased awareness of one's own phonatory behaviour may change the clinical subgroup, as self-assessment is a pivotal basis of diagnosing functional dysphonia (Behlau and colleagues, 2015). It is also important to test what effect high levels of background noise may have on measurements recorded with VoxLog. The biofeedback function of the dosimeter was not used in the current investigation – it would however be interesting to see additional development of this feature to include phonation time, so that each wearer could be reminded to take breaks from phonation throughout the day, e.g. when a certain level phonation time had accumulated.

Measurement instability

As repeated measures were used throughout the studies, the risk of variation and error was high. Vogel and Maruff name two areas that are sensitive to random error: biological and technological errors (Vogel and Maruff, 2014). Biological errors were partly addressed during data collection, as groups were matched for age, however diurnal changes were not taken into account, for logistical reasons. Technological measurement errors were also addressed, as learning effects were eliminated by using the third reading aloud of the standard passage *Nordanvinden och solen* [The North Wind and the Sun] for analyses and all equipment used for voice analyses was well tested for

stability. The previous limited use of VoxLog in research is addressed above. A technical disadvantage in the set-up of the VLT was lack of air quality control in the laboratory recording environment. However, all participants had access to a glass of water during testing to prevent dehydration, in line with Fujiki and colleagues, 2016).

Perceptual assessment of voice is known to be fraught with low inter-rater reliability. This problem was addressed by use of a panel of well experienced voice professional making all assessments by consensus (Shrivastav and colleagues, 2005). The same method was applied for assessments of digital imaging. There is the potential risk of only one individual oppinion taking over, though this was not a problem reported by any panel member.

The voice health questions used in the studies were subject to instability between studies. In Paper 1 all voice health questions were answered on a 4 point Likert scale (0=none, 1=somewhat, 2=moderately, 3=high). Administration of the voice activity questionnaire was changed before starting data accumulation for Papers 2–4. The Likert scale was changed to contain 5 points (0=none/not at all, 1=low/occasionally, 2=some/sometimes, 3=high/often, 4=very high/nonstop).

Future research

Voice production, vocal fatigue and motivation

Vocal fatigue plays a central part in this dissertation. This is a concept in need of further investigation. Vocal loading is the equivalent of voice use, and should not be confused with vocal overload. In order to examine the processes of vocal fatigue through increased vocal effort, heavy vocal loading was deliberately inflicted. The VLT was meant to reflect daily vocal dose during a short task. Self-assessed vocal fatigue was used as a marker, ensuring sufficient vocal loading had been inflicted, so that its impact could be evaluated. No other research compares vocal loading with vocal fatigue as a marker, letting participants be the judge of when heavy vocal loading has been achieved. This would greatly help improve the concept of vocal fatigue as a clinical instrument for diagnostics. A lot is known about the mechanics leading to vocal fatigue, although many researchers have pointed out the lack

of a specific definition of the concept (e.g. Kitch and Oates, 1994; Nanjundeswaran and colleagues, 2015; Solomon, 2008). In 2015 Nanjundeswaran and colleagues observed the important aspect of vocal fatigue as being symptom based and the need for it to be subjectively assessed. The idea was put forth by McCabe and Titze (2002) and by Solomon (2008). Nanjundeswaran and colleagues developed a self-assessment tool to help patients and voice professionals get to the bottom of their voice problem. The authors include two questions that link to motivation for speaking at all (1) I don't feel like talking after a period of voice use and (2) I avoid social situations when I know I have to talk more (Nanjundeswaran and colleagues, 2015, p. 440). Although the authors give muscular or tissue-related dysfunction greater emphasis, they may have started to close the gap between knowledge of mechanical voice use and cognitive aspects of voice production.

Cognitive motivation behind voice production also needs further exploration. For example, when interpreting research results it is important to understand what motivates a participant to give her all in a vocal loading context and what might drive another participant to be more restrained. Knowledge on how far voice patients are willing to push their vocal function in order to be heard in a certain context will give important clues to aiding patients who have problems with making themselves heard in noisy environments. This notion is found in a parallel field, when Pichora-Fuller, Kramer, Eckert and colleagues (2016) describe the same lack of knowledge in listening effort; how motivated are patients with hearing impairments to put in the effort needed to encode what is being said in a noisy environment? The authors draw upon the allegedly neglected neuropsychological concept of conation, which may also shed light on participants' behaviour in a vocal loading context. It reflects patients' ability to focus on and cope with task completion:

> *Conation is a term that has been used in psychology to refer to the ability to apply intellectual energy to a task, as needed over time, to achieve a solution or completion.* Reitan and Wolfson (2000, p. 444)

Little is known about the cognitive processes involved in the arousal, attention and effort that go in to listening (Pichora-Fuller and colleagues, 2016). The same is true for what motivates voice patients to load their vocal organs in different contexts. Levels of compliance or motivation in participants can serve as a source of error in voice research (Vogel and Maruff, 2014) and the drive to make oneself heard could depend on

personality traits, such as extroversion or introversion. As shown by Roy and Bless, vocal hyperfunction, or overuse of laryngeal mechanisms, is associated with extroversion (Roy and Bless, 2000). Hammar's master's thesis explored unpublished data from Papers 2 and 3. She investigated how participants answered the background question: "How would you describe your personality?" In line with Roy and Bless, results showed that out of the n=36 participants who had a clear answer to this question, 79% of participants with voice problems (groups FD and HLC) gave a version of the answer "extrovert". The number for participants without voice problems to give the same type of answer was 47% (Hammar and colleagues, 2015). Kahneman gives weight to the autonomic nervous system and connects effort to cognition and motivational arousal (Kahneman, 1973). Levels of motivation in a certain context may also depend on personal traits and/or pragmatic skills: e.g. letting others speak, rather than talking yourself.

Vocal loading, recovery and intervention

Patients presenting with organic voice pathology are mainly treated through medical intervention stretching beyond behavioural voice therapy. Patients with functional dysphonia are most common among adults seeking medical help for voice problems (Van Houtte, Van Lierde, D'Haeseleer and Claeys, 2010). These patients are commonly treated with behavioural voice therapy (Ruotsalainen, Sellman, Lehto, Jauhiainen and Verbeek, 2007). This is mainly because habitual muscular patterns are seen as the root of the problem if a patient's voice (i.e. laryngeal structures) is quickly fatigued (e.g. Sander and Ripich, 1983).

Iwarsson proposes for SLP's to intentionally break patients' habitual vocal behaviours by applying motor learning skills (Iwarsson, 2015). Motor learning theory links memory capacity to bodily movement, which require skill and practice in order to be automatized. If a person sets out perform a previously untried movement, they will use propriosensory memories of related movements in order to learn the new skill, making in clumsy and inaccurate to start with. The abstracted memories include four sets of memory trace information, including (1) proprioceptory memory from the body prior to movement, (2) planning of e.g. speed, (3) proprioception during implementation (4) evaluation of outcome. Motor learning theory encompasses schema theory, which implies that people can learn movement either through recall schema (by learning a movement which can be applied consciously) or recognition schema, when an already ongoing movement

prompts another correct movement, or both (Adams, 1971; Schmidt and Lee, 2011). This theory may shed light on the coping skill seen during heavy vocal loading, by women with high everyday vocal loading and no vocal complaints (the HLNC group) in Paper 3, which is also discussed under *The concept of vocal loading*. They correctly identified heavy vocal loading and recovered from it quickly. It may be that their healthy reaction to loading is partly due to well-developed proprioceptory feedback from the body. More studies implementing motor learning theory, as suggested by Iwarsson, are central for understanding physical coping skills.

The current study applies two set of interventions: a VLT in order to examine effects of vocal loading and recovery therefrom, and recommendations for voice rest following phonomicrosurgery of benign vocal fold lesions. In Paper 4 it was shown that relative voice rest was more beneficial for the participants than absolute voice rest. However, effects of any certain kind of behavioural voice therapy that would be beneficial for the participants was not explored within the context of heavy vocal loading. This area is important to target in future research, as voice therapy following phonomicrosurgery may alleviate patients' vocal discomfort (Ju and colleagues, 2013). It may also help dampen inflammation in the vocal fold mucosa (Verdolini Abbott and colleagues, 2012). It is crucial that voice therapy directed at a problem as complex and undefined as functional dysphonia be conducted in an effective manner. Kaneko and colleagues (2016), compared absolute voice rest in randomised groups, comparing 3 days to 7 days' voice rest. They found shorter voice rest to be more beneficial than longer. They also suggested controlled voice therapy after the inflammatory phase was over (2 days). It was not controlled, but tube phonation was suggested for early onset tissue activation, in order to enhance patients' motivation to comply with voice care advice and to promote beneficial fibroblast activity in the mucosa. It would be enlightening for the voice clinic for future research to target voice therapy intervention following phonomicrosurgery. In patients with functional dysphonia treatment need not only focus on decreasing hyperfunction, but increase patients' loudness by use of effective resonance rather than muscular force.

The term Evidence Based Practice (EBP) was coined by Dollaghan. It describes a three-way model for SLP's of steering intervention to voice patients based on the best available evidence (1) external (from systematic research) and (2) internal (from clinical practice) which is (3) the most beneficial for the fully informed patient (Dollaghan, 2007, p. 2). It has been established that high-evidence, systematic research is scarce in the voice clinic. Chan and colleagues evaluated the use of EPB in Australia and in line

with previous research they found that largely, SLP's did not apply evidence-based practice in voice therapy with patients with functional voice disorders. The main reason behind this is said to be shortage of time (Chan, McCabe and Madill, 2013). It is of utmost importance that clinical voice professionals be given time within their clinical curricula to find new external research and to gather their own data in order to systemize outcome of voice therapy.

Voice therapy aims to change vocal behaviour. Motivation is known to be important for reaching behavioural changes in voice therapy (Behrman, 2006). Pichora-Fuller and colleagues emphasise exploring how much effort a listener is motivated to dedicate in a certain context, e.g. a noisy environment, in which an interesting conversation topic will spark motivation (Pichora-Fuller and colleagues, 2016). This framework would be applicable for the required vocal effort when speaking and making oneself heard in a noisy environment.

Linking heavy vocal loading to other functions

The neurological link between voice an emotion is partly unexplored. Human vocalization holds a neuronal requirement for voluntary vocalization and speech compared to other species. One can decide whether or not to convey emotions semantically, but will the untrained voice withhold such information in the prosodic signal? It is suggested that what sets human vocalization apart from that of nonhuman primates is the location of laryngeal neuro motor control, found in the primary motor cortex, instead of the premotor cortex (Simonyan and Horwitz, 2011). Perception of emotion carried by the voice signal involves a complex network of cortico neural connections, with pathways running bilaterally between the superior temporal gyrus, close to the primary auditory cortex, and the inferior frontal gyrus (Ethofer, Bretscher, Gschwind, Kreifelts, Wildgruber and Vuilleumier, 2012). Temporal features of the laryngeal motor cortex are yet unexplored (Simonyan and Horwitz, 2011). This implies that underlying processes involved in vocal onset and offset, i.e. phonatory pauses, are unknown. This missing link could be important for patients with functional dysphonia, who seem dependent on short-term recovery (Paper 3). Perhaps more importantly, nothing is known about neuronal feedback between the laryngeal motor cortex and the primary somatosensory cortex (Simonyan and Horwitz, 2011). This relation may be very important in understanding what drives vocal behaviour within voice pathology.

The connection between personality, emotion and voice has recently begun being explored. Although it has been established for a while that there are certain personality traits behind different vocal pathologies, e.g. extroversion behind vocal nodules (Roy and colleagues, 2000), it is vital to knowledge of vocal recovery to find out more about what drives patients to push their voices in different situations. This "push" is something to be expected in trained actors and singers (Kitch and Oates, 1994), but what about patients with functional dysphonia? There is a new hypothesis, put forth by Baker and colleagues, that patients with muscle tension voice disorders to a higher extent than others show difficulty in identifying emotions driving their general behaviour (Baker, Oates, Leeson, Woodford, and Bond, 2014). Their hypothesis and the finding from O'Hara and colleagues, showing a connection between a sense of perfection and functional dysphonia, make a great platform to initiate further investigation in the area.

Take-home for the voice clinic

This dissertation has tested a vocal loading task (VLT) for the voice clinic, which entails exploring vocal loading and recovery behaviour in voice patients with functional dysphonia and benign organic lesions. This is something very few studies thus far have undertaken (e.g. Buekers, 1998). Based on current findings the following take-home messages are proposed to clinical voice professionals caring for these patients:

1. A VLT can be used diagnostically when looking at patients' reaction to vocal loading and recovery therefrom, by use of voice accumulation (dosimetry + voice health questionnaire).
2. Patients with functional dysphonia suffer from their voice problems all day, not only during work (Paper 2).
3. Phonation at high pitch implies increased vocal loading, which can be detrimental to the voice function (Paper 2).
4. Heavy vocal loading (speaking at 85 dB SPL for 30 minutes) is detrimental for vocal function of patients with functional dysphonia. It increases hyperfunction/press, overall grade of dysphonia and dampens vibrational movement in the vocal fold mucosa (Paper 3).
5. Vocal recovery according so self-assessment of unspecified voice problems (measured with VAS) is slower in patients with functional dysphonia than in other groups. They may not return to baseline values for 2 hours following heavy vocal loading (Paper 3). Short-term recovery is more important than long-term recovery to discuss with patients with functional dysphonia.
6. Adequate reaction to heavy vocal loading entails a vital coping skill: identifying heavy vocal loading (papers 3 and 4).
7. Absolute voice rest is difficult for patients to comply with.
8. More data is needed to confirm the findings in paper 4, but results from this study shows that relative voice rest following phonomicrosurgery is more beneficial for long-term recovery of vocal function than absolute voice rest.
9. Unpublished data show that patients with functional dysphonia rarely sense the discomfort from the throat which could lead them to terminate heavy vocal loading. When asked how their throats felt after finishing the vocal loading task, patients most often recognized the sensation of a tense, sore throat from their everyday life.

Conclusions

Study-specific conclusions

Paper 1

Vocal loading, which increasing vocal effort and causes a self-perceived sense of vocal fatigue, can be attained in a clinical environment by letting participants read aloud in ambient noise, without causing long-term damage. It is possible to let participants set the time limit for participation, only submitting them to heavy vocal loading until vocal fatigue is attained. This necessitates permitting participants' self-assessments of vocal effort and vocal fatigue to be in focus, rather than acoustical measurements of the voice signal.

Paper 2

Vocal loading may be associated, not only with high phonatory sound pressure level, but also high levels of fundamental frequency during longer periods of time. Patients with functional dysphonia (group FD) experience self-assessed voice problems during whole days, not only while they are at work, as compared to women who experience only dysphonia associated with their occupation (group HLC). This difference may explain why the latter group do not seek voice therapy.

Paper 3

Patients with functional dysphonia take longer than others to recover from unspecified voice problems in the short-term (2 hours). Their long-term

recovery from specific voice symptoms is equivalent to other groups'. Heavy vocal loading has more negative impact on patients with functional dysphonia than other groups. These patients present with alterations in the voice source caused by vocal loading, manifested through changed vocal fold tissue morphology and a shift toward perceptual hyperfunction/press and overall voice pathology.

Papers 3 and 4

The notion of vocal loading is not complete without a sense of a speaker's **reaction** to vocal loading or the required **recovery time** from vocal loading. These parameters give the concept holistic perspective, taking the speaker's assessment properly into account, and systematically quantifying repercussions of vocal loading.

Paper 4

It is more difficult for patients to comply with absolute voice rest compared to relative voice rest. Patients who were randomized to the relative voice rest group showed significantly better vocal stamina and immediate recovery from vocal loading, as well as significantly improved within-group self-assessment of voice problems at long-term check-up compared to patients randomized to the absolute voice rest group.

General conclusions

1. Patients with functional dysphonia are more vulnerable than others in a context of heavy vocal loading. Their voice function is worse affected and they recover more slowly than others. Care should be taken to arm these patients with coping skills during loud speech.

2. Absolute voice rest is not complied with by patients following phonomicrosurgery.

3. Relative voice rest may be beneficial for long-time vocal recovery following phonomicrosurgery of benign lesions in the vocal fold mucosa. Absolute voice rest, i.e. silence, is not recommended.

4. The concept of vocal loading is reinforced by measurements of reaction to heavy vocal loading and of recovery time following said loading.

Sammanfattning på svenska

Bakgrund

Det är angeläget att utveckla och förfina kliniska instrument som underlättar för diagnostik och behandling av patienter som lider av röststörningar. Ännu finns lite evidensbaserat underlag för diagnostik och omhändertagande av patienter som har svårt att göra sig hörda i bullriga miljöer, till exempel patienter med funktionell dysfoni (röststörning utan organisk sjukdom på stämvecksnivå). Det har hittills inte heller funnits några jämförande belägg bakom för- och nackdelar med total och/eller relativ röstvila direkt efter röstkirurgi. Man har trott att det är skadligt för röstfunktionen att använda den under läkning, men ny forskning tyder på att så inte är fallet. Avhandlingens delstudier fokuserar på röstbeteende under tung röstbelastning och på röståterhämtningsprocesser hos patienter med funktionell dysfoni och patienter med godartad organisk röststörning som genomgår röstkirurgi.

Metoder och resultat

Delstudie 1

Delstudie 1 är en metodutvecklingsstudie som presenterar ett kliniskt användbart röstbelastningstest. Metoden framkallade på ett framgångsrikt sätt rösttrötthet hos n=11 röstfriska forskningspersoner, utan att orsaka någon permanent skada på någon del av röstfunktionen. Metoden är beskriven under nästa rubrik.

Röstbelastningstestet

Röstbelastningstestet användes i samtliga delstudier. I korthet gick testet ut på att låta forskningspersoner läsa en text högt, sittande i en studio. Efter en stund fick forskningspersonen konkurrens av ett sorl bestående av inspelade röster, som skickades ut vid 85 dB SPL via högtalare i studion. Forskningspersonens huvuduppgift var att fortsätta läsa och försöka göra sig hörd och bara avsluta om/när hen upplevde tydligt obehag i halsen. Maxtiden, som var okänd för forskningspersonerna, var 30 minuter. Effekter på röstfunktionen orsakade av röstbelastningstestet undersöktes genom akustisk, perceptuell och spektral analys av röstinspelningar som spelades in precis före och strax efter belastningen. Fysiologiska förändringar på stämvecksnivå undersöktes genom högupplösta filmer och höghastighetsfilmer, inspelade genom rigidendoskopisk laryngoskopi före och efter belastningen. Självskattade subjektiva bedömningar av röstfunktionen noterades också före och efter belastningen. Bedömningar av inspelningar och filmer genomfördes blint. Röstbeteende och subjektiv skattning av röstfunktionen monitorerades under hela processen genom röstackumulering med en röstdosimeter och en röstdagbok. Till röstbeteendet räknades variationer i röstens tonhöjd och ljudtrycksnivå samt den relativa röstbildningstiden. I delstudierna pågick röstackumulering under flera dagar före och efter belastningstestet, för att kunna följa röstbelastningsbeteendet och återhämtning från röstbelastningen.

Delstudie 2

Delstudie 2 jämförde röstbeteende mellan n=20 kvinnliga patienter med diagnosticerad funktionell dysfoni och n=30 kvinnor i tre grupper med funktionella röstproblem på olika nivåer. Kvinnornas röstbeteende jämfördes under tre villkor: (1) röstbelastningstest, (2) arbetstid, (3) fritid. Resultaten från delstudie 2 visar att tung röstbelastning inte bara kan förknippas med röstbildning vid höga ljudtrycksnivåer, alltså stark röst, utan även med röstbildning vid höga frekvenser, alltså ljus röst med högt taltonläge. Detta fynd är viktigt att ta hänsyn till i röstkliniken. Det framkom även att patienter med funktionell dysfoni genom självskattade röstbedömningar, vittnar om röstbesvär som varar under hela dagen, medan andra som har arbetsrelaterade röstbesvär inte upplever besvär på fritiden, trots att röstbeteendet inte skiljer sig markant mellan grupperna under fritiden.

Delstudie 3

Delstudie 3 undersökte samma grupper som delstudie 2. Data från n=50 forskningspersoner samlades in samtidigt till de båda studierna. Delstudie 3 granskade belastningspåverkan, vilken visade sig vara störst för patienter med funktionell dysfoni. Patienterna uppvisade signifikanta förändringar i perceptuell bedömning (ökad press och generell röstpåverkan) samt vibrationsförändringar i stämvecksslemhinnan. Dessa förändringar märktes inte i de andra deltagargrupperna. Resultaten visade också att kvinnor som är vana vid hög röstbelastning i vardagen och som inte upplever röstbesvär (grupp HLNC) relativt sett reagerade starkast på röstbelastnings-testet, och uppvisade objektivt skifte till minskad press och ökat luftläckage under belastning. Dessa fynd tolkas som en copingstrategi, som yttrar sig genom god förmåga att identifiera och anpassa sig till tung röstbelastning. Kortsiktig återhämtning från röstbelastning, följdes genom täta självskattningar av ospecifika röstproblem under två timmar direkt efter röstbelastningstestet. Långsiktig återhämtning undersöktes genom självskattningar av specifika röstbesvärssymptom under ett par dagar efter belastningen. Resultaten visade långsammare kortsiktig återhämtning för patienterna med funktionell dysfoni än för de andra grupperna, medan långsiktig återhämtning var mer enhetlig på gruppnivå.

Delstudie 4

Delstudie 4 är den första i sitt slag. En blind, randomiserad studie som jämförde total röstvila med relativ röstvila (som tillåter försiktig röstbildning) under en vecka direkt efter röstkirurgi. Patienter (n=20) som skulle genomgå röstförbättrande kirurgi för godartade organiska förändringar i stämvecks-slemhinnan delades genom lottning in i två grupper med n=10 i varje. För att bedöma röstuthållighet genomgick samtliga patienter röstbelastningstestet före operation, samt på nytt 3–6 månader efter operation. Belastningar och återhämtning efter operation monitorerades med röstackumulering. Resultaten visade likartad röstfysiologisk långtidsåterhämtning för de båda grupperna, men gruppen som fick relativ röstvila skattade genomgående signifikant bättre röstfunktion vid långtidsuppföljningen än före operation, vilket gruppen som genomgick total röstvila inte gjorde. Resultaten visade också att det var svårt för patienter som rekommenderas röstvila att faktiskt vara helt tysta, trots att de bar en röstackumulator som bevakade dem.

Slutsatser

1. Patienter med funktionell dysfoni är mer sårbara under tung röstbelastning än andra: de påverkas mer negativt och återhämtar sig långsammare efteråt än andra kvinnor med funktionella röstproblem på olika nivåer. Dessa patienter bör ges copingstrategier för att både undvika, men även klara av, röstbelastning. Ökad medvetenhet och planering gällande återhämtningstid efter belastning rekommenderas.

2. Det är svårt för patienter att följa röstråd om total röstvila efter röstkirurgi.

3. Den sammanlagda långsiktiga återhämtning av röstfunktionen är bättre för patienter som rekommenderats relativ röstvila under en vecka efter röstkirurgi än för patienter som rekommenderats total röstvila.

4. Begreppet röstbelastning (engelska: *vocal loading*) förstärks genom tillägg av patienters reaktionsförmåga mot tung röstbelastning samt undersökning av återhämtningstid efter röstbelastning.

References

Adams, J. A. (1971). A closed-loop theory of motor learning. *J Mot Behav, 3*(2), 111-149.

AFS 2005:16 Buller [Noise], (2005). *Statutes of the Work Environment Authority.*

Alton Everest, F. (2001). *Master Handbook of Acoustics* (Fourth ed.). New York. U.S.A: McGraw-Hill.

Anderson, R. C., and Klofstad, C. A. (2012). Preference for Leaders with Masculine Voices Holds in the Case of Feminine Leadership Roles. *PLoS One, 7*(12), e51216. doi:10.1371/journal.pone.0051216

Antonovsky, A. (1987). Health, Stress and Coping. San Francisco: *Jossey-Bass Publishers.*

Atkinson, G., Nevill, A., and Hopkins, W. G. (2000). Measures of reliability in sports medicine and science. *Sports Medicine, 30*(5), 375-381 377p.

Baker, J. (2008). The role of psychogenic and psychosocial factors in the development of functional voice disorders. *International Journal of Speech-Language Pathology, 10*(4), 210-230.

Baker, J., Oates, J. M., Leeson, E., Woodford, H., and Bond, M. J. (2014). Patterns of emotional expression and responses to health and illness in women with functional voice disorders (MTVD) and a comparison group. *J Voice, 28*(6), 762-769. doi:10.1016/j.jvoice.2014.03.005

Behlau, M., Madazio, G., and Oliveira, G. (2015). Functional dysphonia: strategies to improve patient outcomes. *Patient Relat Outcome Meas, 6*, 243-253. doi:10.2147/prom.s68631

Behrman, A. (2006). Facilitating behavioral change in voice therapy: the relevance of motivational interviewing. *Am J Speech Lang Pathol, 15*(3), 215-225. doi:10.1044/1058-0360(2006/020)

Behrman, A., and Sulica, L. (2003). Voice rest after microlaryngoscopy: current opinion and practice. *Laryngoscope, 113*(12), 2182-2186. doi:10.1097/00005537-200312000-00026

Bless, D. M., Hirano, M., and Feder, R. J. (1987). Videostroboscopic evaluation of the larynx. *Ear Nose Throat J, 66*(7), 289-296.

Bottalico, P., Graetzer, S., and Hunter, E. J. (2016). Effects of speech style, room acoustics, and vocal fatigue on vocal effort. *J Acoust Soc Am, 139*(5), 2870. doi:10.1121/1.4950812

Bouchayer, M., and Cornut, G. (1992). Microsurgical treatment of benign vocal fold lesions: indications, technique, results. *Folia Phoniatr (Basel), 44*(3-4), 155-184.

Brewer, M. (2000). Research Design and Issues of Validity. In Reis, H. and Judd, C. (eds) Handbook of Research Methods in Social and Personality Psychology. Cambridge:*Cambridge University Press.*

Buckley, K. L., O'Halloran, P. D., and Oates, J. M. (2015). Occupational Vocal Health of Elite Sports Coaches: An Exploratory Pilot Study of Football Coaches. *Journal of Voice, 29*(4), 476-483. doi:http://dx.doi.org/10.1016/j.jvoice.2014.09.017

Buekers, R. (1998). Are voice endurance tests able to assess vocal fatigue? *Clin Otolaryngol Allied Sci, 23*(6), 533-538.

Buekers, R., Bierens, E., Kingma, H., and Marres, E. H. (1995). Vocal load as measured by the voice accumulator. *Folia Phoniatr Logop, 47*(5), 252-261.

Caraty, M.-J., & Montacié, C. (2014). Vocal fatigue induced by prolonged oral reading: Analysis and detection. *Computer Speech & Language, 28*, 453-466. doi:10.1016/j.csl.2012.12.003

Carding, P. N., Wilson, J. A., MacKenzie, K., and Deary, I. J. (2009). Measuring voice outcomes: state of the science review. *J Laryngol Otol, 123*(8), 823-829. doi:10.1017/s0022215109005398

Chan, A. K., McCabe, P., and Madill, C. J. (2013). The implementation of evidence-based practice in the management of adults with functional voice disorders: a national survey of speech-language pathologists. *Int J Speech Lang Pathol, 15*(3), 334-344. doi:10.3109/17549507.2013.783110

Chang, A., and Karnell, M. P. (2004). Perceived phonatory effort and phonation threshold pressure across a prolonged voice loading task: a study of vocal fatigue. *Journal of Voice, 18*(4), 454-466.

Cheng, J., and Woo, P. (2010). Correlation between the Voice Handicap Index and voice laboratory measurements after phonosurgery. *Ear Nose Throat J, 89*(4), 183-188.

Cohen, J. (1992). Statistical power analysis. *Current Directions in Psychological Science, 1*(3), 98-101.

Colton, R. H., Casper, J. K., and Leonard, R. (2011). *Understanding voice problems : a physiological perspective for diagnosis and treatment*: Philadelphia : Wolters Kluwer Health/Lippincott Williams and Wilkins, 20114. ed.

Coombs, A. C., Carswell, A. J., and Tierney, P. A. (2013). Voice rest after vocal fold surgery: current practice and evidence. *J Laryngol Otol, 127*(8), 773-779. doi:10.1017/s0022215113001485

Cramer, D. (1998). *Fundamental Statistics for Social Research: Step-by-Step Calculations and Computer Techniques Using SPSS for Windows*. London: Routledge.

Cramer, D., and Howitt, D. (2004). *The SAGE dictionary of statistics*. London: SAGE.

De Bodt, M. S., Wuyts, F. L., Van de Heyning, P. H., Lambrechts, L., and Vanden Abeele, D. (1998). Predicting vocal outcome by means of a vocal endurance test: a 5-year follow-up study in female teachers. *Laryngoscope, 108*(9), 1363-1367.

Dejonckere, P. H. (2000). Clinical implementation of a multidimensional basic protocol for assessing functional results of voice therapy. A preliminary study. *Rev Laryngol Otol Rhinol (Bord), 121*(5), 311-313.

Dejonckere, P. H., Bradley, P., Clemente, P., and colleagues (2001). A basic protocol for functional assessment of voice pathology, especially for investigating the efficacy of (phonosurgical) treatments and evaluating new assessment techniques: Guideline elaborated by the Committee on Phoniatrics of the European Laryngological Society (ELS). *European Archives of Oto-Rhino-Laryngology, 258*(2), 77-82. doi:10.1007/s004050000299

Doane, D. P., and Seward, L. E. (2011). Measuring Skewness: A Forgotten Statistic? *Journal of Statistics Education, 19*(2).

Doellinger, M., Lohscheller, J., McWhorter, A., and Kunduk, M. (2009). Variability of Normal Vocal Fold Dynamics for Different Vocal Loading in One Healthy Subject Investigated by Phonovibrograms. *Journal of Voice, 23*(2), 175-181. doi:10.1016/j.jvoice.2007.09.008

Dollaghan, C. (2007). *The handbook of evidence-based practice in communication disorders*. Baltimore, MD, USA: Paul H Brooks Publishing.

Echternach, M., Nusseck, M., Dippold, S., Spahn, C., and Richter, B. (2014). Fundamental frequency, sound pressure level and vocal dose of a vocal loading test in comparison to a real teaching

situation. *European Archives of Oto-Rhino-Laryngology*. doi:10.1007/s00405-014-3200-6

Ethofer, T., Bretscher, J., Gschwind, M., Kreifelts, B., Wildgruber, D., and Vuilleumier, P. (2012). Emotional voice areas: anatomic location, functional properties, and structural connections revealed by combined fMRI/DTI. *Cereb Cortex, 22*(1), 191-200. doi:10.1093/cercor/bhr113

Fant, G. (1982). Preliminaries to analysis of the human voice source. *STL-QPSR, 23*(4), 1-27.

Fritzell, B. (1996). Voice Disorders and Occupations. *Logopedics Phoniatrics Vocology, 21*, 7-12.

Fujiki, R. B., Chapleau, A., Sundarrajan, A., McKenna, V., and Sivasankar, M. P. (2016). The Interaction of Surface Hydration and Vocal Loading on Voice Measures. *J Voice*. doi:10.1016/j.jvoice.2016.07.005

Gelfer, M. P., Andrews, M. L., and Schmidt, C. P. (1991). Effects of prolonged loud reading on selected measures of vocal function in trained and untrained singers. *Journal of Voice, 5*(2), 158-167. doi:10.1016/S0892-1997(05)80179-1

Gelfer, M. P., Andrews, M. L., and Schmidt, C. P. (1996). Documenting laryngeal change following prolonged loud reading: A videostroboscopic study. *Journal of Voice, 10*(4), 368-377.

Gray, S. D. (2000). Cellular physiology of the vocal folds. In C. M. Rosen, T (Ed.), *The Otolaryngologic Clinics of North America*. Philadelphia, USA: WB Saunders Company.

Hallin, A. E., Frost, K., Holmberg, E. B., and Södersten, M. (2012). Voice and speech range profiles and Voice Handicap Index for males-- methodological issues and data. *Logoped Phoniatr Vocol, 37*(2), 47-61. doi:10.3109/14015439.2011.607469

Hammar, K., Whitling, S., Lyberg-Åhlander, V., Rydell, R. (2015). "*I coundn't even speak to the cat" – a qualitative study on womens' experiences of voice use and voice complaints*. Master's thesis [Swedish]. Dept. of Logopedics, Phoniatrics and Audiology, Faculty of Medicine. Lund University.

Hammarberg, B. (1986). *Perceptual and acoustic analysis of dysphonia. Doctoral dissertation.* (Doctoral Thesis), Karolinska Institute, Karolinska Institute, Stockholm.

Hanschmann, H., Gaipl, C., and Berger, R. (2011). Preliminary results of a computer-assisted vocal load test with 10-min test duration.

European Archives of Oto-Rhino-Laryngology, 268(2), 309-313. doi:10.1007/s00405-010-1428-3

Hirano, M. (1975). Phonosurgery: Basic and clinical investigation. Otologia, 21, 239-240.

Hirano, M. (1981). Clinical examination of voice. New York, USA: Springer.

Hogikyan, N. D., Wodchis, W. P., Terrell, J. E., Bradford, C. R., and Esclamado, R. M. (2000). Voice-related quality of life (V-RQOL) following type I thyroplasty for unilateral vocal fold paralysis. J Voice, 14(3), 378-386.

Holmberg, E. (1980). Laryngeal airway resistance as a function of phonation type. J. Acoust. Soc. Am. Suppl. 1, 68, 101.

Holube, I., Fredelake, S., Vlaming, M., and Kollmeier, B. (2010). Development and analysis of an International Speech Test Signal (ISTS). Int J Audiol, 49(12), 891-903. doi:10.3109/14992027.2010.506889

Hunter, E. J., and Titze, I. R. (2009). Quantifying vocal fatigue recovery: dynamic vocal recovery trajectories after a vocal loading exercise. Annals of Otology, Rhinology and Laryngology, 118(6), 449-460.

Hunter, E. J., and Titze, I. R. (2010). Variations in intensity, fundamental frequency, and voicing for teachers in occupational versus nonoccupational settings. Journal of Speech, Language, and Hearing Research, 53(4), 862-875. doi:10.1044/1092-4388(2009/09-0040)

Ishikawa, K., and Thibeault, S. (2010). Voice rest versus exercise: a review of the literature. J Voice, 24(4), 379-387. doi:10.1016/j.jvoice.2008.10.011

ISO 8253-1 (1989) Acoustics: Audiometric test methods part 1. Basic pure tone eir and bone conduction threshold audiometry. ISO 8253-1 International Organization for Standardization 8253-1.

Iwarsson, J. (2015). Facilitating behavioral learning and habit change in voice therapy--theoretic premises and practical strategies. Logoped Phoniatr Vocol, 40(4), 179-186. doi:10.3109/14015439.2014.936498

Jacobson, B. H., Johnson, A., Grywalski, C., Silbergleit, A., Jacobson, G., Benninger, M. S., and Newman, C. W. (1997). The Voice Handicap Index (VHI)Development and Validation. American Journal of Speech-Language Pathology, 6(3), 66-70. doi:10.1044/1058-0360.0603.66

Ju, Y. H., Jung, K. Y., Kwon, S. Y., and colleagues (2013). Effect of voice therapy after phonomicrosurgery for vocal polyps: a prospective, historically controlled, clinical study. *J Laryngol Otol, 127*(11), 1134-1138. doi:10.1017/s0022215113002454

Kahan, B. C., Rehal, S., and Cro, S. (2015). Risk of selection bias in randomised trials. *Trials, 16*, 405. doi:10.1186/s13063-015-0920-x

Kahneman, D. (1973). *Attention and effort*: Englewood Cliffs, N.J. : Prentice-Hall, 1973.

Kaneko, M., Shiromoto, O., Fujiu-Kurachi, M., Kishimoto, Y., Tateya, I., and Hirano, S. (2016). Optimal Duration for Voice Rest After Vocal Fold Surgery: Randomized Controlled Clinical Study. *J Voice*. doi:10.1016/j.jvoice.2016.02.009

Kelchner, L. N., Toner, M. M., and Lee, L. (2006). Effects of prolonged loud reading on normal adolescent male voices. *Language, Speech and Hearing Services in Schools, 37*(2), 96-103.

Kiagiadaki, D., Remacle, M., Lawson, G., Bachy, V., and Van der Vorst, S. (2015). The effect of voice rest on the outcome of phonosurgery for benign laryngeal lesions: preliminary results of a prospective randomized study. *Ann Otol Rhinol Laryngol, 124*(5), 407-412. doi:10.1177/0003489414560583

Kitch, J. A., and Oates, J. (1994). The perceptual features of vocal fatigue as self-reported by a group of actors and singers. *J Voice, 8*(3), 207-214.

Kooijman, P. G., de Jong, F. I., Thomas, G., and colleagues (2006). Risk factors for voice problems in teachers. *Folia Phoniatr Logop, 58*(3), 159-174. doi:10.1159/000091730

Lachin, J. M., Matts, J. P., and Wei, L. J. (1988). Randomization in clinical trials: conclusions and recommendations. *Control Clin Trials, 9*(4), 365-374.

Lane, H and Tranel, B. (1971). The Lombard Sign and the Role of Hearing in Speech. *Journal of Speech, Language, and Hearing Research, 14*(December), 677-709. doi:doi:10.1044/jshr.1404.677

Laukkanen, A., Järvinen, K., Artkoski, M., and colleagues (2004). Changes in voice and subjective sensations during a 45-min vocal loading test in female subjects with vocal training. *Folia Phoniatrica et Logopaedica, 56*(6), 335-346.

Lehto, L., Alku, P., Vilkman, E., and Laaksonen, L. (2006). Occupational voice complaints and objective acoustic measurements - Do they correlate? *Logopedics Phoniatrics Vocology, 31*(4), 147-152. doi:10.1080/14015430600654654

Li, N. Y., Verdolini, K., Clermont, G., Mi, Q., Rubinstein, E. N., Hebda, P. A., and Vodovotz, Y. (2008). A patient-specific in silico model of inflammation and healing tested in acute vocal fold injury. *PLoS One, 3*(7), e2789. doi:10.1371/journal.pone.0002789

Linville, S. E. (1995). Changes in glottal configuration in women after loud talking. *Journal Of Voice: Official Journal Of The Voice Foundation, 9*(1), 57-65.

Lohscheller, J., Doellinger, M., McWhorter, A. J., and Kunduk, M. (2008). Preliminary study on the quantitative analysis of vocal loading effects on vocal fold dynamics using phonovibrograms. *Annals of Otology, Rhinology and Laryngology, 117*(7 part 1), 484-493.

Lyberg-Åhlander, V. L., Rydell, R., and Löfqvist, A. (2011). Speaker's Comfort in Teaching Environments: Voice Problems in Swedish Teaching Staff. *Journal of Voice, 25*(4), 430-440. doi:http://dx.doi.org/10.1016/j.jvoice.2009.12.006

Lyberg-Åhlander, V. L., Rydell, R., and Löfqvist, A. (2012). How Do Teachers With Self-Reported Voice Problems Differ From Their Peers With Self-Reported Voice Health? *Journal of Voice, 26*(4), e149-e161. doi:http://dx.doi.org/10.1016/j.jvoice.2011.06.005

Lyberg-Åhlander, V., Pelegrin-García, D., Whitling, S., Rydell, R., and Löfqvist, A. (2014). Teachers' voice use in teaching environments: a field study using ambulatory phonation monitor. *J Voice, 28*(6), 841 e845-815. doi:10.1016/j.jvoice.2014.03.006

Lyberg-Åhlander, V., Rydell, R., Eriksson, J., and Schalén, L. (2010). Throat related symptoms and voice: development of an instrument for self assessment of throat-problems. *BMC Ear, Nose and Throat Disorders, 10*, 5-12. doi:10.1186/1472-6815-10-5

Lyberg-Åhlander, V., Rydell, R., Löfqvist, A., Pelegrin-García, D., and Brunskog, J. (2015). Teachers' Voice Use in Teaching Environment. Aspects on Speakers' Comfort. *Energy Procedia, 78*, 3090-3095. doi:10.1016/j.egypro.2015.11.762

Löfqvist, A., Carlborg, B., & Kitzing, P. (1982). Initial validation of an indirect measure of subglottal pressure during vowels. *J Acoust Soc Am, 72*(2), 633-635.

Löfqvist, A., and Mandersson, B. (1987). Long-time average spectrum of speech and voice analysis. *Folia Phoniatr (Basel), 39*(5), 221-229.

Ma, E. P., and Yiu, E. M. (2001). Voice activity and participation profile: assessing the impact of voice disorders on daily activities. *J Speech Lang Hear Res, 44*(3), 511-524.

McCabe, D. J., and Titze, I. R. (2002). Chant therapy for treating vocal fatigue among public school teachers: A preliminary study. *American Journal of Speech-Language Pathology, 11*(4), 356-369.

Misono, S., Banks, K., Gaillard, P., Goding, G. S., Jr., and Yueh, B. (2015). The clinical utility of vocal dosimetry for assessing voice rest. *Laryngoscope, 125*(1), 171-176. doi:10.1002/lary.24887

Nanjundeswaran, C., Jacobson, B. H., Gartner-Schmidt, J., and Verdolini Abbott, K. (2015). Vocal Fatigue Index (VFI): Development and Validation. *J Voice, 29*(4), 433-440. doi:10.1016/j.jvoice.2014.09.012

Neils, L. R., and Yairi, E. (1987). Effects of speaking in noise on vocal fatigue and vocal recovery. *Folia Phoniatr (Basel), 39*(2), 104-112.

Niebudek-Bogusz, E., Kotylo, P., and Sliwinska-Kowalska, M. (2007). Evaluation of voice acoustic parameters related to vocal-loading test in professionally active teachers with dysphonia. *INTERNATIONAL JOURNAL OF OCCUPATIONAL MEDICINE AND ENVIRONMENTAL HEALTH, 20*(1), 25-30.

O'Hara, J., Miller, T., Carding, P., Wilson, J., and Deary, V. (2011). Relationship between fatigue, perfectionism, and functional dysphonia. *Otolaryngol Head Neck Surg, 144*(6), 921-926. doi:10.1177/0194599811401236

Oates, J., and Winkworth, A. (2008). Current knowledge, controversies and future directions in hyperfunctional voice disorders. *Int J Speech Lang Pathol, 10*(4), 267-277. doi:10.1080/17549500802140153

Pelegrin-Garcia, D., and Brunskog, J. (2012). Speakers' comfort and voice level variation in classrooms: laboratory research. *J Acoust Soc Am, 132*(1), 249-260. doi:10.1121/1.4728212

Pelegrin-García, D., Brunskog, J., Lyberg-Åhlander, V., and Löfqvist, A. (2012). Measurement and prediction of voice support and room gain in school classrooms. *J Acoust Soc Am, 131*(1), 194-204. doi:10.1121/1.3665987

Pelegrin-García, D., Smits, B., Brunskog, J., and Jeong, C. H. (2011). Vocal effort with changing talker-to-listener distance in different acoustic environments. *J Acoust Soc Am, 129*(4), 1981-1990. doi:10.1121/1.3552881

Pichora-Fuller, M. K., Kramer, S. E., Eckert, M. A., Edwards, B., Hornsby, B. W., Humes, L. E., Wingfield, A. (2016). Hearing Impairment and Cognitive Energy: The Framework for Understanding Effortful Listening (FUEL). *Ear Hear, 37 Suppl 1*, 5s-27s. doi:10.1097/aud.0000000000000312

Ramig, L. O., & Verdolini, K. (1998). Treatment efficacy: voice disorders. J Speech Lang Hear Res, 41(1), S101-116.

Rantala, L., and Vilkman, E. (1999). Relationship between subjective voice complaints and acoustic parameters in female teachers' voices. *J Voice, 13*(4), 484-495.

Rantala, L. M., Hakala, S. J., Holmqvist, S., and Sala, E. (2012). Connections Between Voice Ergonomic Risk Factors and Voice Symptoms, Voice Handicap, and Respiratory Tract Diseases. *Journal of Voice, 26*(6), 819.e813-819.e820. doi:http://dx.doi.org/10.1016/j.jvoice.2012.06.001

Razali, N. M., and Wah, Y. B. (2011). Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. *Journal of Statistical Modeling and Analytics, 2*(1), 21-33.

Reitan, R. M., & Wolfson, D. (2000). Conation: a neglected aspect of neuropsychological functioning. *Arch Clin Neuropsychol, 15*(5), 443-453.

Remacle, A., Morsomme, D., Berrué, E., and Finck, C. (2012). Vocal impact of a prolonged reading task in dysphonic versus normophonic female teachers. *Journal of Voice, 26*(6), 820.e821-820.e813. doi:10.1016/j.jvoice.2012.06.002

Rosen, C. A., Lee, A. S., Osborne, J., Zullo, T., and Murry, T. (2004). Development and Validation of the Voice Handicap Index-10. *The Laryngoscope, 114*(9), 1549-1556. doi:10.1097/00005537-200409000-00009

Rousseau, B., Cohen, S. M., Zeller, A. S., Scearce, L., Tritter, A. G., and Garrett, C. G. (2011). Compliance and quality of life in patients on prescribed voice rest. *Otolaryngol Head Neck Surg, 144*(1), 104-107. doi:10.1177/0194599810390465

Roy, N. (2012). Optimal dose-response relationships in voice therapy. *International Journal of Speech-Language Pathology, 14*(5), 419-423. doi:10.3109/17549507.2012.686119

Roy, N., and Bless, D. M. (2000). Personality traits and psychological factors in voice pathology: a foundation for future research. *J Speech Lang Hear Res, 43*(3), 737-748.

Roy, N., Bless, D. M., and Heisey, D. (2000). Personality and voice disorders: a superfactor trait analysis. *Journal of Speech, Language and Hearing Research, 43*(3), 749-768.

Ruotsalainen, J. H., Sellman, J., Lehto, L., Jauhiainen, M., and Verbeek, J. H. (2007). Interventions for treating functional dysphonia in adults.

    Cochrane Database Syst Rev(3), Cd006373.
    doi:10.1002/14651858.CD006373.pub2

Sala E, H. U., Ketola R, Laine L, Olkinuora P, Rantala L, Sihvo M. (2009). *Voice Ergonomic Screening in Work Environment—Handbook and Check-list [in Finnish and Swedish]*. Helsinki, Finland: The Finnish Institute of Occupational Health.

Sander, E. K., and Ripich, D. E. (1983). Vocal fatigue. *Ann Otol Rhinol Laryngol, 92*(2 Pt 1), 141-145.

Schmidt, R. A., and Lee, T. D. (2011). *Motor control and learning: a behavioral emphasis* (5th ed.). Champaign, IL: Human Kinetics Publishers.

Shapiro, S. S., and Wilk, M. B. (1965). An Analysis of Variance Test for Normality (Complete Samples)*, 591.

Sherman, D., Jensen, P. J. (1962). Harshness and Oral-Reading Time. *Journal of Speech and Hearing Disorders, 27*(May), 172-177. doi:10.1044/jshd.2702.172

Shrivastav, R., Sapienza, C. M., and Nandur, V. (2005). Application of psychometric theory to the measurement of voice quality using rating scales. *Journal of Speech, Language and Hearing Research, 48*(2), 323-335.

Simonyan, K., and Horwitz, B. (2011). Laryngeal motor cortex and control of speech in humans. *Neuroscientist, 17*(2), 197-208. doi:10.1177/1073858410386727

Smitheran, J. R., & Hixon, T. J. (1981). A clinical method for estimating laryngeal airway resistance during vowel production. *J Speech Hear Disord, 46*(2), 138-146.

Solomon, N. P. (2008). Vocal fatigue and its relation to vocal hyperfunction. *International Journal of Speech-Language Pathology, 10*(4), 254-266.

Stemple, J. C., Stanley, J., and Lee, L. (1995). Objective measures of voice production in normal subjects following prolonged voice use. *Journal of Voice, 9*(2), 127-133. doi:10.1016/S0892-1997(05)80245-0

Stone, R. E., Jr., and Sharf, D. J. (1973). Vocal change associated with the use of atypical pitch and intensity levels. *Folia Phoniatr (Basel), 25*(1), 91-103.

Švec, J. G., Popolo, P. S., and Titze, I. R. (2003). Measurement of vocal doses in speech: experimental procedure and signal processing. *Logopedics Phoniatrics Vocology, 28*(4), 181.

Szabo Portela, A., Hammarberg, B., and Södersten, M. (2014). Speaking Fundamental Frequency and Phonation Time during Work and Leisure Time in Vocally Healthy Preschool Teachers Measured with a Voice Accumulator. *Folia Phoniatrica et Logopaedica, 65*(2), 84-90. doi:10.1159/000354673

Södersten, M., Salomão, G. L., McAllister, A., and Ternström, S. (2015). Natural Voice Use in Patients With Voice Disorders and Vocally Healthy Speakers Based on 2 Days Voice Accumulator Information From a Database. *J Voice*. doi:10.1016/j.jvoice.2014.09.006

Södersten, M., Ternström, S., and Bohman, M. (2005). Loud speech in realistic environmental noise: Phonetogram data, perceptual voice quality, subjective ratings, and gender differences in healthy speakers. *Journal of Voice, 19*(1), 29-46. doi:10.1016/j.jvoice.2004.05.002

Ternström, S., Bohman, M., and Södersten, M. (2006). Loud speech over noise: some spectral attributes, with gender differences. *J Acoust Soc Am, 119*(3), 1648-1665.

Ternström, S., Pabon, P., and Södersten, M. (2016). The Voice Range Profile: Its Function, Applications, Pitfalls and Potential. *Acta Acustica united with Acustica, 102*(2), 268-283. doi:10.3813/AAA.918943

Titze, I. R. (1992). Phonation threshold pressure: A missing link in glottal aerodynamics. *Journal of the Acoustical Society of America, 91*(5), 2926-2935. doi:10.1121/1.402928

Titze, I. R. (1994a). Mechanical stress in phonation. *J Voice, 8*(2), 99-105.

Titze, I. R. (1994b). *Principles of voice production*: Englewood Cliffs, N.J. : Prentice Hall, cop. 1994.

Titze, I. R. (1999). Toward occupational safety criteria for vocalization. *Logopedics Phoniatrics Vocology, 24*(2), 49-54. doi:10.1080/140154399435110

Titze, I. R. (2001). Criteria for occupational risk in vocalization. In P. H. Dejonckere (Ed.), *Occupational Voice: Care and Cure* (p. 1–10). The Hague, The Netherlands: Kugler.

Titze, I. R., Švec, J. G., and Popolo, P. S. (2003). Vocal Dose Measures: Quantifying Accumulated Vibration Exposure in Vocal Fold Tissues. *Journal of Speech, Language and Hearing Research, 46*(4), 919-932.

Van Houtte, E., Van Lierde, K., D'Haeseleer, E., and Claeys, S. (2010). The prevalence of laryngeal pathology in a treatment-seeking population

with dysphonia. *Laryngoscope, 120*(2), 306-312. doi:10.1002/lary.20696

Van Stan, J. H., Gustafsson, J., Schalling, E., and Hillman, R. E. (2014). Direct Comparison of Three Commercially Available Devices for Voice Ambulatory Monitoring and Biofeedback. *SIG 3 Perspectives on Voice and Voice Disorders, 24*(2), 80-86. doi:10.1044/vvd24.2.80

Verdolini Abbott, K. V., Li, N. Y. K., Branski, R. C., Rosen, C. A., Grillo, E., Steinhauer, K., and Hebda, P. A. (2012). Vocal exercise may attenuate acute vocal fold inflammation. *J Voice, 26*(6), 814.e811-814.e813. doi:10.1016/j.jvoice.2012.03.008

Vilkman, E. (2004). Occupational safety and health aspects of voice and speech professions... 28th World Congress of the International Association of Logopedics and Phoniatrics, Brisbane, Australia, 29th August to 2nd September 2004. *Folia Phoniatrica et Logopaedica, 56*(4), 220-253.

Vilkman, E., Lauri, E. R., Alku, P., Sala, E., and Sihvo, M. (1999). Effects of prolonged oral reading on F0, SPL, subglottal pressure and amplitude characteristics of glottal flow waveforms. *Journal of Voice, 13*(2), 303-312. doi:10.1016/S0892-1997(99)80036-8

Vintturi, J., Alku, P., Lauri, E.-R., Sala, E., Sihvo, M., and Vilkman, E. (2001). Objective analysis of vocal warm-up with special reference to ergonomic factors. *Journal of Voice, 15*(1), 36-53.

Vintturi, J., Alku, P., Sala, E., Sihvo, M., and Vilkman, E. (2003). Loading-related subjective symptoms during a vocal loading test with special reference to gender and some ergonomic factors. *Folia Phoniatrica et Logopaedica, 55*(2), 55-69.

Vogel, A. P., and Maruff, P. (2014). Monitoring change requires a rethink of assessment practices in voice and speech. *Logoped Phoniatr Vocol, 39*(2), 56-61. doi:10.3109/14015439.2013.775332

Welham, N. V., and Maclagan, M. A. (2003). Vocal fatigue: current knowledge and future directions. *J Voice, 17*(1), 21-30.

WHO: World Health Organization (2001) International Classification of Functioning, Disability and Health: ICF. *World Health Organization*, Geneva

World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. (2013). *Jama, 310*(20), 2191-2194. doi:10.1001/jama.2013.281053

Wrona, E. A., Peng, R., Amin, M. R., Branski, R. C., and Freytes, D. O. (2016). Extracellular Matrix for Vocal Fold Lamina Propria

Replacement: A Review. *Tissue Eng Part B Rev*. doi:10.1089/ten.TEB.2016.0015

Zambon, F., Moreti, F., and Behlau, M. (2014). Coping strategies in teachers with vocal complaint. *J Voice, 28*(3), 341-348. doi:10.1016/j.jvoice.2013.11.008

Paper I

Design of a Clinical Vocal Loading Test With Long-Time Measurement of Voice

*Susanna Whitling, *,†Roland Rydell, and *Viveka Lyberg Åhlander, *†*Lund, Sweden*

Summary: Objectives. The aim of this study was to design a clinical vocal loading task (VLT) and to track vocal loading and recovery in voice-healthy subjects.
Study Design. Pilot study.
Methods. Voice-healthy subjects (six female, five male) took part in a controlled VLT in the voice clinic. The VLT was designed to induce vocal fatigue. The subjects read aloud while making themselves heard through ambient speech-babble aired at 85 dB sound pressure level (SPL). Reading was terminated by the subjects when or if they felt any discomfort from the throat. The subjects wore a voice accumulator and filled out a voice activity questionnaire 1 day preceding and for 2 days following the VLT. Expert panels assessed vocal quality and laryngeal physiology from recordings.
Results. The subjects endured the VLT for 3–30 minutes. All subjects perceived vocal loading in the VLT. All subjects raised the fundamental frequency and SPL of their speech during the VLT. No match was shown between assessment of voice quality and laryngeal physiology. The subjects showed phonation quotients of 64–82% in the task. Measurements of phonation threshold pressure (PTP) were unstable and were not used. Self-perceived vocal loading receded after 24 hours.
Conclusions. An authentic vocal load was simulated through the chosen method. Onset and recovery from self-perceived vocal loading was traceable through the voice activity questionnaire. The range of endurance in the VLT was an unexpected finding, indicating the complexity of vocal loading.
Key Words: Vocal loading task–Vocal recovery–Long-time measurement of voice–Clinical voice assessment.

INTRODUCTION

In literature covering vocal loading, this concept is defined by a combination of prolonged voice use with added loading factors, such as high phonation at high sound pressure levels (SPLs). These factors may affect fundamental frequency, loudness, phonation modality, and/or laryngeal features, such as vibratory characteristics of the vocal folds or the external frame of the larynx.[1]

Prolonged voice use is commonly regarded as one of the most relevant factors in functional voice disorders.[2,3] Similarly we expect prolonged voice use to cause vocal fatigue also in voice-healthy subjects. Assuming typical loudness, the healthy, adult voice is expected to be fatigued after roughly 4–6 hours of use.[4,5] Numerous studies have tested vocal loading in voice-healthy subjects,[6–15] however a test of 4–6 hours is impractical to perform and administer in a clinical or laboratory setting.[6] Many studies have documented increased fundamental frequency and speech SPLs following vocal loading. This vocal behavior may follow the principle of the Lombard effect, which entails an involuntary rise of speaking pitch and loudness as a speaker strives to increase vocal audibility in ambient noise.[16] This has been regarded as adequate adaptation to loading[11,17,18] and has also been interpreted as being part of the voice function's circadian rhythm.[18]

Objective changes to the phonatory function are often sought in tests of vocal loading. It has been examined by many,[2,6,7,9–15,17–40] with little conformity,[41] making cross-study comparisons difficult.[21] In laboratory settings a time-limit has traditionally been set for intersubject comparability.[2,6,25,38]

It is essential to investigate symptom-based signs of vocal fatigue. By letting subjects phonate at high SPLs for a prolonged period of time in a controlled environment the vocal function can be studied for indirect evidence of vocal fatigue, before and after vocal loading. No previous studies have let subjects themselves set the time-limit for a controlled vocal loading task (VLT) to gain knowledge of what amount of increased phonation loudness will induce vocal fatigue. Asking subjects about the condition of their own voice is of great importance when assessing voice problems based on vocal fatigue. Solomon,[41] p. 254, argues that researchers and clinicians need to agree that vocal fatigue is a symptom-based voice problem, stemming from the self-report of an increased sense of effort with prolonged phonation, that is, symptoms voice professionals can neither hear, nor see.

The use of voice accumulation along with voice activity questionnaires could verify vocal loading and trace it from onset to regression. Although vocal behavior and spontaneous vocal load have been examined in the field through several different methods and for varying amounts of time,[8,19,42–47] the progression of controlled vocal loading has not yet been tracked through long-time voice accumulation.

Objectives

This pilot study was aimed to design a VLT which by extension should be possible to use in diagnostics of voice disorders.

Another aim is to track onset of vocal loading and vocal recovery.

The task was aimed to induce vocal fatigue in voice-healthy subjects by letting them read loudly until they perceive discomfort from the throat. Increased phonation loudness and vocal fatigue were tracked to observe objective changes in vocal acoustics and physiology following controlled vocal loading. Self-assessment was used to provide information on how the subjects experienced the program.

We wanted to learn how this method affects the subjects' time of participation, self-assessment of vocal function, fundamental frequency, speech SPL, perceived voice quality, and vocal fold physiology.

Method queries

- How do we design a VLT which is suitable for the voice clinic and which allows subjects to set the time limit for increased speaking loudness?
- How can recovery processes from potential vocal load/strain inflicted on voice-healthy individuals by a VLT be tracked?

Hypotheses

- The VLT will cause the subjects to experience vocal fatigue.
- Vocal recovery processes are traceable through voice accumulation and a structured voice activity questionnaire.

METHODS

In this study voice-healthy individuals were tracked during four working days with a VLT taking place midway. Processes involved in recovery from vocal loading have in previous research been tracked through self-evaluation of voice function.[14] The controlled vocal loading in this study was meant to represent vocal load accumulated over time, here compressed into a reading task of a maximum of 30 minutes in length. To learn more about what amount of increased phonation loudness will induce vocal fatigue in voice-healthy subjects, a timing aspect was also introduced, whereby the subjects themselves discontinued phonation if and when they perceived any discomfort from the throat while phonating.

Subjects

In total n = 11 subjects took part in this pilot study. Five male (mean age 37 years, range 28–55 years) and six female (mean age 36 years, range 28–47 years) were recruited among colleagues and friends through verbal inquiry and agreement. The subjects all had to meet the following requirements: adult (>18 years), voice-healthy, nonsmokers with or without vocal training. The exclusion criteria were: children (≤17 years), laryngeal pathologies, smokers, and professional singers. Vocal health was determined by oral interviews on medical history, carried out by the first author (S.W.) followed by a phoniatrical

TABLE 1.
Demographics of All Subjects (N = 11), Who Were All Voice-Healthy Nonsmokers

Subject	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11
Gender	F	F	F	F	F	F	M	M	M	M	M
Age	33	36	36	34	47	28	33	55	34	28	34
Better ear HL	25	20	20	20	20	20	20	20	20	20	20
Vocal work load	Low	Moderate	Low	Low	Moderate	Low	Some	Moderate	Moderate	Moderate	Moderate
Allergies	Pollen, nutritional, fur	None	None	Pollen, cat	Asthma, pollen, fur	Penicillin V	None	None	None	None	Undiagnosed: cat, horse
Medication	Antihistamines	None	None	None	Omeprazole	None	None	None	None	None	None
Stress level	Moderate, fluctuates	High, constant	High, fluctuates	Moderate, fluctuates	High, slow fluctuation	High, fluctuates	Moderate, fluctuates	High, constant	Moderate, constant	High, fluctuates	High, constant

laryngeal examination, carried out by the second author (R.R.). To reflect a variety of typical voices, the amount and nature of vocal training varied among the subjects. No professional singers were recruited. Some subjects had undergone singing voice training (n = 4) and some had trained speaking voices (n = 5). Three of the subjects had received no vocal training. The subjects underwent pure tone audiological screening at ≥20 dB SPL at 250 Hz, 500 Hz, 1 kHz, 4 kHz, and 8 kHz. All had normal hearing results. Demographics on the subjects are shown in Table 1.

Tracking vocal load and recovery
A VLT was preceded and followed by pre- (baseline) and post-referential audio voice recordings, measurements of phonation threshold pressure (PTP), and by laryngeal digital imaging. All voice recordings and the VLT took place in a double-walled soundproof booth (complying with the maximum permissible ambient SPL as specified in ISO 8252–1). They were recorded and lead by the first author (S.W.). Laryngeal digital imaging was carried out by the second author (R.R.), in a neighboring examination room. Follow-up recordings were made 2 days after the VLT. Objective tracking of voice production was carried out using a voice accumulator for four consecutive days, the VLT taking place on day 2. Subjective changes to the voice function of each subject were tracked using a structured voice activity questionnaire, designed for the study, throughout the voice accumulation process. Fundamental changes to speech tend not to occur when recordings made during the daytime are compared.[48] All recordings in this study were made before 5 PM For logistic reasons the recordings and digital imaging were made at different times of the day for the different subjects, however the times were maintained from before VLT recordings to follow-up recordings for each subject. Assessments of digital imaging and perceptual voice assessments were carried out by other voice professionals than the authors of this study. The digital imaging was assessed by three phoniatricians and the perceptual voice quality assessments were made by three speech-language pathologists specialized in voice care. None of the professionals knew the aims of the study. An overview of the VLT process is presented in Figure 1.

Vocal loading task
A controlled VLT was designed specifically for this study, to induce vocal fatigue in voice-healthy subjects. Each subject was required to read a randomly chosen passage from a handbook in audiology out loud. They were asked to make themselves heard though ambient noise, aired over a loudspeaker (HQ Power; model VDSABS8A; Vellerman NV, Gavere, Belgium) placed 1.5 m behind them. Each subject was instructed to keep reading and to terminate the reading when/if they felt any discomfort from the throat. If failing to make themselves heard they were prompted by the test-leader (S.W.) to speak louder, indicated by test-leader spreading her arms and mouthing "starkare" ("louder").

The Swedish Work Health Authority has set a time limit for exposure to noise in the work place at 85 dB SPL for a maximum of 8 hours. Not knowing what other levels of noise
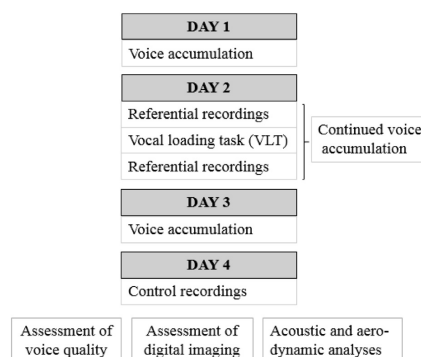
FIGURE 1. Process flowchart describing the VLT.

the subjects would be exposed to during the day of the VLT, a time limit of 30 minutes was set for the loading task to avoid impairing the subjects' hearing[49] and to make the task performable time-wise. The time limit was known to the subjects before they started reading.

Ambient noise
The noise consisted of ANL multi-talker speech-babble noise with 12 North American speakers. The noise was retrieved from the official AND CD (Arizona Travelodge; Cosmos Distributing Inc., Torrence, CA). It is identical to the noise used in the revised speech-in-noise test by Bilger et al, 1979.[50] Its long-time average spectrum does not differ significantly from Swedish babble. Giving the subjects a chance to adapt to the noise, they started reading in the otherwise silent booth. The ambient noise started airing at 55 dB SPL after 45 seconds, doubling in perceived loudness, by 10 dB SPL[51,52] every 10 seconds, until 85 dB SPL was reached after 30 seconds. The starting point at 55 dB SPL is the recommended background level for example a classroom, providing 99% speech intelligibility.[53,54] The SPLs were calibrated using an integrating-averaging sound level meter (type 2240, no. 00240502; Brüel & Kjær, Denmark).

Individual capacity of phonatory SPL
The VLT required the subjects to be able to phonate at ≥85 dB SPL (the SPL of the ambient noise in the VLT). To control for each subject's ability to phonate at high SPL, comparisons were made between SPL levels from maximum phonetograms (voice range profiles, VRP), SPL values of phonation in the VLT, extracted from VoxLog, and a general comparison of maximum SPL from VRP to 85 dB SPL. As fundamental frequency tends to rise with increased SPL of speech,[16] the value of mean fundamental frequency from the VLT, extracted from VoxLog, was used as a reference point, at which correlating SPL levels were examined for each subject.

The VRPs were performed with the subjects sitting down, wearing a head-mounted microphone (MKE 2, no 09_1,

Sennheiser, en-de.sennheiser.com/), calibrated at 94 dB SPL, at a distance of 15 cm from the mouth (converted to 30 cm from the mouth). In accordance with guidelines laid down by the European Union of Phoniatricians, the signal was converted to equal 30 cm from the mouth.[55] Recordings were made on a real-time phonetograph; *Phog Interactive Phonetography System 2.5 for PC* (Hitech Medical, Täby, Sweden). The subjects phonated with glissandos on [a:], trying to cover as large an area as possible in frequency and SPL. They started phonation at habitual levels and were free to take the time they needed. All instructions were given by the first author (S.W.). The VRPs were recorded close in time to the VLT, however, not within the 4 days of the process, not to inflict added vocal fatigue.

Audio voice recordings
A standard passage (The North wind and the Sun) of approximately 45 seconds was read in Swedish. For these recordings the same microphone was used in the same way as for the VRP (previously mentionted).

Analyses of audio recordings. The voice signal was digitized at 16 kHz with 16 bit resolution. The signal was retrieved and examined for fundamental frequency (mean, mode, and standard deviation) and speech SPLs (mean, mode, standard deviation) using *Soundswell Core 4.0* and *Soundswell Voice 4.0* (Hitech development). Real-time phonetograms (Speech Range Profiles, SRP) were retrieved and examined for fundamental frequency and speech SPL range using Phog. Interactive *Phonetography System 2.5 for PC* (Hitech development).

Information on voice source (spectral tilt) was obtained through long-time average spectra (LTAS),[49] *ad modum* Åhlander et al[56] (based on Löfqvist and Mandersson[57]).

Voice quality assessments (blind) from before and after VLT recordings were made by an expert panel of three experienced speech and language pathologists, specialized in voice care. The panel made consensus decisions when rating each stimulus,

they sat as long as they wanted (ca 2.5 hours) and they were able to listen to each stimulus multiple times. Once they had assessed a recording, they were not allowed to return to it. The assessments were made *ad modum* SVEA (Stockholm Voice Evaluation Approach), with 100 mm Visual Analog Scales, which were manually analyzed on paper. 0 = unaffected, 100 = deviant in the greatest possible manner. The degrees of the following parameters were assessed: hyperfunction; breathiness; instability; roughness; glottal fry; sonorance; affected voice/overall voice condition. The panel also assessed which of the before and after VLT voice recording they judged to be more deviant from typical voice quality. All speech-language pathologists were unaware of which audio recording was recorded before or after the VLT.

Assessment of laryngeal digital imaging
A digital documentation system was used for laryngeal digital imaging (HRES Endocam, model 5562.9 color; Wolf, Germany) with a 70° rigid endoscope. High resolution mode was used for imaging for evaluation of overall morphology, symmetry, adduction and abduction of the vocal folds. High-speed mode (4000 frames per second for female subjects, 2000 frames per second for male subjects) was used for imaging for evaluation of mode and symmetry of vibration on the glottal level. Of the 11 subjects, nine performed the examination without local anesthetic, one male and one female subject each required 1 ml of 40 mg/ml Lidocaine hydrochloride APP. All recordings were made by one of the authors (R.R.), who did none of the subsequent assessments.

The pre- and post-VLT digital imaging was assessed by an expert panel of three phoniatricians. The panel made consensus decisions when rating each stimulus. They sat as long as they wanted (ca 2.5 hours) and they were able to look at each stimulus multiple times. Once they had assessed a recording they were not allowed to return to it. Glottal open quotient was assessed from kymograms derived from the high speed digital imaging. Table 2 shows the qualitative parameters were

TABLE 2.
Described Qualitative Parameters From Digital Laryngeal Imaging Recorded Before and After the Vocal Loading Task, Assessed by a Panel of Three Phoniatricians

Glottal shape	Shape of vocal folds and opening between vocal folds
Glottal regularity	Regularity of vocal folds and opening between vocal folds
Morphological alteration	Overall deviation from typical morphology
Adduction right vocal fold	Right vocal fold's motion toward middle of glottis
Adduction left vocal fold	Left vocal fold's motion toward middle of glottis
Abduction right vocal fold	Right vocal fold's motion from middle of glottis
Abduction left vocal fold	Left vocal fold's motion from middle of glottis
Wave amplitude right vocal fold	Size of mucosal wave in right vocal fold
Wave amplitude left vocal fold	Size of mucosal wave in left vocal fold
Wave propagation right vocal fold	Dispersion of mucosal wave in right vocal fold
Wave propagation left vocal fold	Dispersion of mucosal wave in left vocal fold
Phase difference	Difference in wave amplitude and propagation of the vocal folds
Right vestibular fold	Activity in right vestibular fold
Left vestibular fold	Activity in left vestibular fold
Corniculate symmetry phonation	Symmetry of corniculate tubercles during phonation
Corniculate symmetry rest	Symmetry of corniculate tubercles during rest

assessed in line with Lyberg Åhlander et al (2012).[58] The panel also assessed which of the before and after VLT laryngeal digital imaging they judged to be more deviant from typical vocal physiology. All phoniatricians were unaware of which imaging was recorded prior to or following the VLT. The assessments were made, according to clinical routine, by grading each parameter with numbers 0–3 (0 = unaffected, 1 = slightly deviant, 2 = moderately deviant, 3 = highly deviant).

Phonation threshold pressure

The PTP is the minimum pressure drop required to achieve phonation when air passes through the glottis from the lungs and is normally 2–3 cm H_2O.[51] Values for intraoral air pressure and those for PTP have been assessed to be consistent. (Holmberg 1980, Löfqvist et al 1982, Smitheran and Hixon 1981). PTP was estimated from measurements of intraoral air pressure using the Phonatory Aerodynamic System (PAS, model 6600; KayPENTAX, New Jersey). The PAS is compiled by hardware, a handheld module with a mask covering the mouth and nose, and by software for PC.

Estimates for intraoral air pressure were calculated by using the software protocol for voicing efficiency. The intraoral air pressure was measured through a narrow plastic tube, placed between the lips during phonation on the aspirated, syllable [p^ha]. The syllable was repeated five times with voicing ranging from a whisper to soft phonation. After acquiring data, the transition from a whisper to soft phonation was determined through contours of fundamental frequency and SPL in the PAS software. To allow for a mean value to be calculated this procedure is repeated three times for each session. All measurements of PTP were made by the first author (S.W.).

Voice accumulation

To objectively track voice production, including vocal loading and recovery processes we used the portable voice accumulator VoxLog developed by SonVox AB (Umeå, Sweden). The voice activity questionnaire was used in order to make accumulated voice data comprehensive. Each subject wore VoxLog and logged their activities in the voice activity questionnaires for four consecutive days: 1 day before VLT, the day of the task and 2 days following the VLT.

VoxLog. The system consists of a voice meter (11,5 × 8 × 2 cm) covered with black plastic and a neck collar containing an accelerometer and a microphone, Figure 2. The accelerometer provides information on fundamental frequency, cycle dose and phonation time quotient. The acoustic speech signal of the bearer is never recorded. Fast Fourier transform approach is used to extract fundamental frequency from the accelerometer signal. Local maxima in the spectrum, spectral peaks, that fulfill a power criterion based on the global maximum, are selected as candidates for fundamental frequency. Among these, the peak of lowest frequency whose power, together with the power of its first few harmonics, S1, S2 and S3, exceed a certain quotient of the total power of the signal, is selected as the final estimate for fundamental frequency.

The SPL of the voice and the environment noise are calculated using a built-in microphone. The accelerometer signal is

FIGURE 2. VoxLog hardware. Accumulation device and neck collar with built-in microphone and accelerometer, SonVox AB.

used to determine whether phonation is present in the current speech frame, based on a moving average of the signal power. If phonation is registered, the SPL measured by the microphone is regarded as voice SPL (otherwise as noise). The reported value of the voice SPL is as measured at the location where the sensor is worn, ie, no conversion to SPL at 15 or 30 cm is made. Analyses of VoxLog data were made in appertaining software, *VoxLog Connect* (SonVox AB, Umeå, Sweden). Absolute numbers were extracted to *Microsoft Excel* (Microsoft, Redmond, WA). *VoxLog Connect* does not provide standard deviation of fundamental frequency or SPL values.

Voice activity questionnaire. To track subjective experience of vocal loading, vocal strain, and recovery processes a study specific structured voice activity questionnaire (based on Lyberg Åhlander et al, 2014) was used throughout the voice accumulation. The questions are based on the Voice Handicap Index, VHI.[59] To enable data from the voice accumulation to be matched with the voice activity questionnaire, it was crucial for the subjects to note the time of their entries, to be able to tell when activities were performed. The subjects were instructed to fill out the questionnaire as soon as they carried out any activity. The questionnaire was meant to give the examiner information of what kind of activities the subjects took part in, if the activities entailed work or leisure and how the subjects assessed their voice function and quality throughout the days. For the latter, a 100-mm visual analog scale was used, marking "how the voice feels right now", ranging from "no voice problems", to "maximum voice problems." The subjects were instructed to fill out the questionnaire as often as they wanted, at a minimum of four times per day. For further information, the following voice questions (from VHI-T Lund[60]) were filled out, answers were given on a 4-point scale (0 = none, 1 = somewhat, 2 = moderately, and 3 = high): Does your voice feel tired?; Does your voice fail you during speech?; Are you having difficulty making yourself heard?; Do you need to clear your throat?; Do you need to cough?; Is your throat sore?; Does your throat feel tense?; Are you hoarse?; How is your current stress level?; Do you have enough air to speak? The voice activity questionnaire ended on three questions evaluating the voice accumulation experience as a whole: Did anything unusual occur with your voice during the voice accumulation? Did
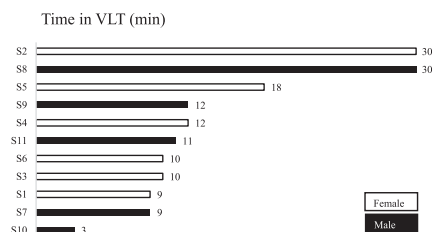
Time in VLT (min)

FIGURE 3. Time of performance in the VLT in minutes for all subjects (n = 11), in order of endurance.

your voice change in any way during the voice accumulation? How have you slept between the measured days?

What was your experience of wearing the voice accumulator?

Statistics

Nonparametric tests have been used because of the small samples tested. This small statistical examination is meant merely to complement the qualitative analysis of this VLT. Changes in mean fundamental frequency, mean SPL and mean PTP were explored statistically, using the Wilcoxon signed rank test to compare means across groups. Alpha (α) levels were set at $P < 0.05$. Effect size follows Cohen (1992), showing small effect-size at 0.1, medium effect-size at 0.3, and large effect-size at 0.5.[61] Fundamental frequency was compared among female subjects (n = 6) and male subjects (n = 5), with the two groups separated (as groups F and M). Mean SPL and mean PTP were explored in all subjects as one group (n = 11). All statistical data were computed using *SPSS 21* (SPSS, Inc., Chicago, IL). Increase and decrease in different values are accounted for in percentage of values attained before vocal loading.

Research ethical approval

The study was approved by the Regional Ethical Review Board in Lund, Sweden on April 25, 2013 (#2013/174).

RESULTS

Results from female and male subjects will mostly be presented separately in order of endurance in the VLT.

Time in VLT

Each subject was told to keep reading aloud, trying to make themselves heard. Figure 3 shows the times for loud reading in the VLT. Two of the subjects: S2 (female) and S8 (male) endured the maximum time of 30 minutes in the VLT. The mean time for endurance for all subjects was 14 minutes. The median time for all subjects was 11 minutes. The range of times was 3–30 minutes. For female subjects (n = 6), the mean time for endurance was 15 minutes, median: 11, range: 9–30 minutes. For male subjects (n = 5), the mean time for endurance was 13 minutes, median: 11 minutes, range: 3–30 minutes.

Individual capacity of phonatory SPL

SPL ability in VRP was compared with mean SPL in the VLT at the corresponding value mean frequency in the VLT for each subject (n = 11). Their SPL output in the VLT ranged from 88–100% of their SPL capacity in the VRP at a given fundamental frequency. Maximum SPL from the VRP was compared with 85 dB SPL at the mean fundamental frequency of the VLT for each subject. This showed capacity to exceed 85 dB SPL with 7–25% (Table 3).

Changes in self-assessed vocal function after VLT

None of the subjects reported high scores in their voice activity questionnaires from their field measurements preceding the VLT. The subjects stated different reasons for withdrawing from the VLT; however, they all (n = 11) felt some discomfort from the throat. Some reported feeling tired from the noise (n = 2), some withdrew, stating that they could have continued a little longer (n = 2). None of the voice-healthy subjects reported any voice problems before the VLT. The voice activity questionnaires were filled out freely, ie, not all at the same occurrence, which is why the results cannot be tracked with conformity, eg, hourly. The vocal discomfort experienced by

TABLE 3.
Individual Capacity of Phonatory SPL in All Subjects (N = 11) in Order of Performance in VLT

Subject	Gender	Time in VLT	F_0 Mean in VLT (Hz)	SPL Mean in VLT (dB SPL)	Max SPL in VRP (dB SPL)	SPL in VLT Comp. to VRP (%)	Max SPL in VRP Comp. to 85 dB SPL (%)
S10	M	3	130	86	94	91	111
S7	M	9	205	93	98	95	115
S1	F	9	320	94	94	100	111
S3	F	10	259	92	94	98	111
S6	F	10	340	90	91	99	107
S11	M	11	253	99	106	93	125
S4	F	12	260	88	93	94	109
S9	M	12	163	92	102	90	120
S5	F	18	337	93	95	98	112
S8	M	30	286	91	104	88	122
S2	F	30	340	92	99	93	116

Notes: The table shows increase or decrease in SPL when ability of high SPL phonation from total voice range profiles are compared to the SPL of the VLT and to 85 dB SPL. All comparisons are made of the SPL produced at the F_0 mean level of the VLT. Differences are expressed in %.

TABLE 4.
Self-Assessments of Vocal Comfort in all Subjects (N = 11) in Order of Performance in VLT

Subject	Gender	Time in VLT	VAS Before VLT	VAS After VLT	VAS After 1 h	VAS After 2 h	VAS After 3–5 h	VAS After 24 h
S10	M	3	0	20	20	0	0	10
S7	M	9	0	40	20	10	30	0
S1	F	9	10	30	20	10	0	0
S3	F	10	0	20	0	10	20	10
S6	F	10	40	20	20	30	30	10
S11	M	11	0	0	0	0	0	0
S4	F	12	0	20	0	0	0	0
S9	M	12	0	20	0	0	0	0
S5	F	18	10	100	80	80	70	0
S8	M	30	0	40	20	20	20	0
S2	F	30	0	10	0	0	0	0

Notes: Visual analog scale (100 mm VAS) entries 1 hour before the VLT; promptly after the VLT; after 1 hour, 2 hours, 3–5 hours, and 24 hours after the VLT.

each subject was tracked in their voice activity questionnaires, showing an increase in vocal discomfort after the VLT, and a decline to baseline after about 24 hours. Table 4 shows how the subjects all assessed their voice comfort in the hour before VLT (pre-VLT), within 1 minute after the VLT (post-VLT) and in the hours following the VLT.

Vocal changes due to VLT
Fundamental frequency was measured in two ways: by audio recordings pre- and post-VLT and with VoxLog, before, during, and after VLT. There was a significant rise with strong correlation to mean fundamental frequency extracted before and during vocal loading from VoxLog in female subjects (n = 6), $Z = -2.201$, $P = 0.028$, r = 0.90, and male subjects (n = 5), $Z = -2.023$, $P = 0.043$, r = 0.90. The values dropped significantly when fundamental frequency during vocal loading was compared with values after vocal loading, female subjects: $Z = -2.201$, $P = 0.028$, r = 0.90, male subjects: $Z = -2.023$, $P = 0.043$, r = 0.90. Changes in fundamental frequency are shown in Table 5 (female) and Table 6 (male). No significant difference was shown when mean fundamental frequency before vocal loading was compared with values after vocal loading. There were no significant changes to the post-

VLT audio recordings compared with the pre-VLT recordings, indicating a rapid regression of the F_0 raise.

The mean SPL of the speech signal was measured in two ways: by audio recordings before and after VLT with Phog, and before, during, and after VLT with VoxLog. Speech SPL was explored in all subjects as one group (n = 11). A significant rise with strong correlation was shown in values of speech SPL extracted from VoxLog prevocal loading compared with speech SPL during vocal loading, $Z = -2.934$, $P = 0.003$, r = 0.88. Values of speech SPL from VoxLog during vocal loading also decreased to a significant degree when compared with values after vocal loading, $Z = -2.934$, $P = 0.003$, r = 0.88.

A carryover effect is shown in a significant rise with strong correlation of mean speech SPL, extracted from Phog, when comparing values before vocal loading with values after vocal loading, $Z = -2.934$, $P = 0.003$, r = 0.88. Values from VoxLog confirm this carryover effect, showing a statistically significant rise of speech SPL postvocal loading, compared with prevocal loading recordings, $Z = -2.667$, $P = 0.008$, r = 0.88.

Changes in speech SPL are shown in Table 7. According to VoxLog data, all subjects were able to match the required SPL of the ambient babble (85 dB SPL) during the VLT. There were no differences to the mean frequency or mean SPL of

TABLE 5.
Rise in Mean Fundamental Frequency (Hz) During VLT and Decrease in Mean Fundamental Frequency after VLT in Female (N = 6) Subjects in Order of Performance in VLT

Subject	Time in VLT (min)	1 h Before VLT	During VLT	1 h After VLT
S1	9	221	320	235
S3	10	217	259	229
S6	10	252	340	257
S4	12	257	260	252
S5	18	245	337	244
S2	30	220	343	223

Notes: Data extracted from VoxLog.

TABLE 6.
Raise in Mean Fundamental Frequency During VLT and Decrease in Mean Fundamental Frequency After VLT in Male (N $=$ 5) Subjects in Order of Performance in VLT

Subject	Time in VLT (min)	1 h Before VLT	During VLT	1 h After VLT
S10	3	120	130	113
S7	9	160	260	131
S11	11	136	253	144
S9	12	133	163	129
S8	30	165	286	214

Notes: Data is extracted from VoxLog.

speech when values before VLT were compared with values from follow-up recordings 2 days later.

Long-time average spectrum
Changes to more hyperfunction, or increased muscle activity, in spectral energy were shown in female subjects (n $=$ 5), after the VLT. The voice signal of n $=$ 3 male subjects showed more hyperfunction, or increased muscle activity, in the distribution of spectral energy after the VLT. Figure 4 shows an overview of changes in phonatory energy.

Speech range profile
There were no significant changes to speech range profiles. Fundamental frequency range increased in all male subjects apart from subject S10, who endured the VLT for the shortest amount of time (3 minutes). The results regarding range of SPL were more diverse for each subject.

Phonation threshold pressure
There were no significant changes to mean PTP and these results were very diverse. No clear connection between decreased PTP and degree of previous voice training was revealed (Table 8).

Perceptual assessment of perceived voice quality
Roughness was never observed and breathiness had low ratings in all subjects. Instability increased only in subject S1 (female) after vocal loading. A majority of the subjects (n $=$ 7) increased

hyperfunction after VLT. However, subjects S9 and S10 were assessed to have less hyperfunction after the VLT, in both cases corresponding to LTAS results, as is shown in Figure 5.

Vocal sonorance decreased in n $=$ 6 subjects, whereas it increased in n $=$ 4, subject S11 did not change his vocal output according to assessment. Tables 9 (F) and 10 (M) give a detailed description of the assessments.

Voice quality of the passage recorded after VLT was deemed by the speech and language pathologists to be more generally affected in n $=$ 5 female subjects. For S6, who read for 10 minutes, the condition was reversed as her pre-VLT recording was assessed to be more affected than her post-VLT recording, indicating less vocal impact induced by the VLT and/or some vocal warm-up effect. For n $=$ 4 male subjects, the voice quality of the passage recorded before VLT was deemed by the SLPs to be more generally affected than the recordings after the VLT. Subject S8 was the only male subject whose post-VLT recording was assessed to be more affected, compared with his pre-VLT recording. He was also the only male to read for 30 minutes.

There was no match between SLPs' and phoniatricians' general assessments before and after VLT.

Digital imaging
For the most part, all subjects were assessed to have normal vocal fold physiology. Some of the parameters could not be assessed because of missing imaging (caused by subjects gagging, obstructing epiglottis, or difficult angles). Changes

TABLE 7.
Changes in Mean Sound Pressure Level of Speech During VLT, in All Subjects (N $=$ 11) in Order of Performance in VLT.

Subject	Gender	Time in VLT (min)	1 h Before VLT	During VLT	1 h After VLT
S10	M	3	71	86	73
S7	M	9	83	88	81
S1	F	9	79	94	80
S3	F	10	80	92	85
S6	F	10	76	90	81
S11	M	11	79	99	80
S4	F	12	83	88	81
S9	M	12	80	92	77
S5	F	18	79	93	82
S8	M	30	79	91	81
S2	F	30	84	92	84

Notes: Data is extracted from VoxLog.
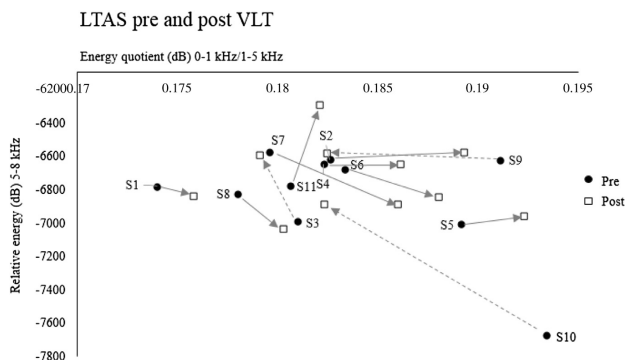
LTAS pre and post VLT

FIGURE 4. Change in long-time average spectrum in all subjects (n = 11) before and after VLT, showing increased hypofunction (decrease of muscle activity) after VLT in subjects S3, S9, and S10. Remaining subjects show increased hyperfunction (increase in muscle activity) after VLT.

in open quotient varied to a high degree among the subjects. For n = 4 female subjects, the phoniatricians judged the post-VLT digital imaging to be more deviant from typical vocal physiology. The imaging pre-VLT was assessed to be more deviant from typical vocal physiology in S1, who read for the shortest amount of time among female subjects (ca 9 minutes) and S2, who read for the longest amount of time among all subjects (30 minutes). For n = 3 male subjects, the phoniatricians judged the before VLT digital imaging to be more deviant from typical vocal physiology. The imaging post-VLT was assessed to be more deviant from typical vocal physiology in S7, who read for ca 9 minutes and S8, who read for the longest amount of time among all subjects (30 minutes). Tables 11 (F) and 12 (M) give a detailed description of the assessments.

There was no agreement between phoniatricians' and SLPs' general assessments before and after VLT.

Phonation quotient

The long-time measurements with VoxLog showed the subjects using their voices to different degrees during different activ-

ities. Compared with, eg, Titze et al,[62] most subjects exhibited low percentages of phonation time (or quotient), when all 4 days of voice accumulation were accounted for. However, in their VLT's, they attained very high percentages of phonation, compared with other periods of vocal activity in their data. Figure 6 shows the phonation quotients of the 4 day measurements (the VLT excluded in total data). The phonation quotients for the subjects' first day and for a period of at least an hour in which the subjects deemed they used their voices to a great extent are compared with controlled recordings with VoxLog: the first being the baseline recording, in which the subjects read the passage The *North Wind and the Sun* three times (*ca* 2 minutes) and the second, being the VLT.

DISCUSSION

Voice clinicians need more knowledge on what patients are vocally capable of outside the voice clinic. Earlier attempts have shown it difficult to mimic authentic vocal loading in a laboratory setting. This study was an attempt to find a suitable

TABLE 8.
Outcome of Phonation Threshold Pressure (cm H_2O) Measured Before and After VLT and at Follow-up Measurement 2 Days Later in All Subjects (N = 11) in Order of Performance in VLT

Subject	Gender	Time in VLT (min)	Before VLT (100% cmH$_2$O)	After VLT (%)	Follow-up (%)
S10	M	3	3.8	104	169
S7	M	9	6.2	80	99
S1	F	9	3.6	106	94
S3	F	10	6.5	84	96
S6	F	10	2.6	96	115
S11	M	11	5.8	88	74
S4	F	12	5.6	81	68
S9	M	12	5.0	118	145
S5	F	18	4.3	133	95
S8	M	30	5.1	138	101
S2	F	30	5.3	112	115

Notes: The increase or decrease of values compared to before VLT measurements are expressed in %.
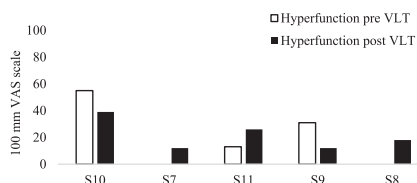
FIGURE 5. Assessment of male subjects' hyperfunction before and after vocal loading.

method of testing the course of vocal loading, ie, the onset and regression of increased phonation loudness and its consequences in the controlled environment of the voice clinic. Increased phonation loudness was inflicted through loud reading, which is not equivalent to spontaneous speech, but closer to it than, eg, sustained vowel phonation or repeated sentences. The ambient multitalker speech-babble provided the element of making oneself heard in a lifelike situation. The method is extensive and requires plenty of time for carrying out recordings and analyses. However, the majority of the time is spent on long-time measurement of voice, which is mostly carried out away from the clinic and thus not entirely controlled. The long-time voice accumulation provided information on everyday phonation. The mode of testing gives clinicians an extensive picture of a subject's voice, as it discloses vocal performance, endurance, and self-perception of vocal function.

Measurement errors

The measurements of PTP, compared over time indicated PTP to be too unreliable for conclusions to be drawn. Eg, PTP ranged from 3.9 cm H_2O at pretest to 6.7 cm H_2O at follow-up measurement in one subject.

Phonation quotient estimated by VoxLog was the vocal parameter which changed most drastically due to the VLT. The subjects' phonation quotient increased drastically when trying to make themselves heard through ambient noise, with female subjects reaching higher percentages than male sub-

jects, in line with Södersten et al.[9] According to Hunter and Titze[63] accumulated phonation time makes out as much as half the amount of speaking time. This proposition is not applicable in the present study, as all subjects reached phonation quotients of more than 64%.

Long-time voice accumulation

VoxLog does not require calibration before use, making it user friendly and facilitating long-time voice accumulation (ie, the subject can easily put it on and take it off over the course of many days). However, this convenience is problematic, as is makes SPL estimations unreliable. The SPL of the voice and of the surrounding noise are calculated using a built-in microphone. Fast Fourier transformation is used to extract fundamental frequency from the accelerometer signal and if phonation is registered by the accelerometer, the SPL measured by the microphone is regarded as voice SPL (otherwise as noise). The lack of calibration requires methodological considerations to be taken into account. Eg, the neck collar was too large for n = 5 of the female subjects, suggesting that placement of the built-in microphone might have moved during SPL estimations. For the same reason, the extraction of fundamental frequency is not entirely reliable. Because this study was carried out, the neck collar has been changed by SonVox AB into a sturdier, adjustable version, and research on validation of the VoxLog microphone is currently in progress at Umeå University, Sweden.

The long-time voice accumulation disclosed apparent measurements errors, eg, VoxLog mistook an hour of bassoon playing in subject S7 as phonation at 308 Hz and 95 dB SPL (when in fact the subject's overall fundamental frequency mean was 150 Hz, voice level 87 dB SPL). Subject S6 wore VoxLog during the first half of a high-intensity gym class. VoxLog made F0 and SPL estimations, thus the assumption that the subject phonated during her workout can be made (phonation quotient at 14% through 44 minutes). However, the subject attests she did not speak at all during the class, although she did occasionally groan and draw breath heavily. Heavy breathing being mistaken for phonation, could be one of the reasons for phonation quotients growing large during the VLT.

TABLE 9.
VAS-Assessment of Voice Quality for All Female Subjects (n = 6) in Order of Endurance in the Vocal Loading Task

Voice Quality Parameter	Subject											
	S1		S3		S6		S4		S5		S2	
	Before VLT	After VLT	Before VLT	After VLT	Before VLT	After VLT	Before VLT	After VLT	Before VLT	After VLT	Before VLT	After VLT
Affected	25	37	0	7	7	2	17	11	0	0	4	15
Hyperfunction	29	19	7	25	14	7	16	21	0	4	3	28
Breathy	0	0	7	17	0	0	14	0	0	2	0	0
Instability	9	43	0	0	2	5	0	0	0	0	0	6
Roughness	0	0	0	0	0	0	0	0	0	0	0	0
Glottal fry	64	14	2	1	2	4	14	2	0	0	13	4
Sonorance	54	46	88	82	76	83	68	77	100	89	93	74

Notes: Each parameter was graded on a 100 mm visual analog scale; 0 = unaffected, 100 = deviant in the greatest possible manner.

TABLE 10.
VAS-Assessment of Voice Quality for All Male Subjects (N = 5) in Order of Endurance in the Vocal Loading Task

Voice Quality Parameter	Subject									
	S10		S7		S11		S9		S8	
	Before VLT	After VLT	Before VLT	After VLT	Before VLT	After VLT	Before VLT	After VLT	Before VLT	After VLT
Affected	54	44	0	10	8	7	16	12	1	7
Hyperfunction	55	39	0	12	13	26	31	12	0	18
Breathy	5	3	0	0	0	0	0	0	0	4
Instability	9	0	1	2	0	0	7	9	0	0
Roughness	0	0	0	0	0	0	0	0	0	0
Glottal fry	98	72	5	11	54	3	59	4	7	28
Sonorance	48	52	85	72	63	62	60	64	82	70

Notes: Each parameter was graded on a 100 mm visual analog scale, 0 = unaffected, 100 = deviant in the greatest possible manner.

Designing a voice activity questionnaire

The voice activity questionnaire used in this study would have been more suited to track subjects' recovery from vocal loading if it had included more frequent entries immediately after the VLT and during the hours following. This would have made it possible to track self-perception of vocal status with a higher resolution. To make the questionnaire more user friendly for subjects, the time of entries could have been roughly preset, such as entries for morning, lunch time, afternoon, and evening.

Vocal fatigue and endurance

Very little seems to be known regarding what will evoke fatigue in voice-healthy individuals. The participating subjects showed different sensitivity for vocal loading. Vocal endurance based in vocal experience play a part in the results. As this pilot study

TABLE 11.
Assessment of Digital Imaging for All Female Subjects (N = 6) in Order of Endurance in the Vocal Loading Task

Voice Quality Parameter	Subject											
	S1		S3		S6		S4		S5		S2	
	Before VLT	After VLT	Before VLT	After VLT	Before VLT	After VLT	Before VLT	After VLT	Before VLT	After VLT	Before VLT	After VLT
Open quotient (%)	11	46	63	57	12	53	14	48	67	74	7	48
Glottal shape	0	0	1	0	1	0	1	1	0	0	0	0
Glottal regularity	0	0	0	0	0	0	0	0	0	0	0	0
Morphological alteration	0	0	1	1	0	0	0	0	0	0	0	0
Adduction right vf	0	0	0	0	0	0	0	0	0	0	0	0
Adduction left vf	0	0	0	0	0	0	0	0	0	0	0	0
Abduction right vf	0	—	0	0	0	0	0	—	0	0	0	—
Abduction left vf	0	—	0	0	0	0	0	—	0	0	0	—
Wave amplitude right vf	1	0	0	1	1	0	0	0	0	1	1	0
Wave amplitude left vf	1	0	0	1	0	0	0	1	0	1	0	0
Wave propagation right vf	1	0	0	0	1	0	0	0	0	0	1	0
Wave propagation left vf	1	0	0	0	1	1	0	0	0	0	0	0
Phase difference	0	0	0	0	0	0	0	0	0	0	1	0
False right vf	0	0	0	0	0	0	0	0	0	0	0	0
False left vf	0	0	0	0	0	0	0	0	0	0	0	0
Symmetry of corniculate tubercles during phonation	1	1	1	1	0	0	0	0	1	1	1	1
Symmetry of corniculate tubercles during rest	0	—	1	1	0	0	0	—	0	0	1	—

Abbreviation: vf, vocal fold.
Notes: Each parameter was graded with numbers 0–3 (0, unaffected; 1, slightly deviant; 2, moderately deviant; 3, highly deviant) from 0 (unaffected). Glottal open quotient was assessed from kymograms derived from the high speed digital imaging and is presented in %.

TABLE 12.
Assessment From Digital Imaging of All Male Subjects (N = 5) in Order of Endurance in the VLT

Voice Quality Parameter	Subject									
	S10		S7		S11		S9		S8	
	Before VLT	After VLT	Before VLT	After VLT	Before VLT	After VLT	Before VLT	After VLT	Before VLT	After VLT
Open quotient (%)	6	52	5	47	62	65	9	51	46	48
Glottal shape	2	1	1	0	2	1	1	1	0	1
Glottal regularity	0	0	0	0	0	0	0	0	0	0
Morphological alteration	1	1	0	0	0	0	0	0	1	1
Adduction right vf	0	0	0	0	0	0	0	0	0	0
Adduction left vf	0	0	0	0	0	0	1	1	0	0
Abduction right vf	0	—	—	—	0	—	0	—	0	—
Abduction left vf	0	—	—	—	0	—	1	—	0	—
Wave amplitude right vf	0	0	0	0	1	0	1	1	0	0
Wave amplitude left vf	0	0	0	0	0	0	2	1	0	0
Wave propagation right vf	0	0	0	0	1	0	2	2	0	0
Wave propagation left vf	1	0	0	1	1	0	3	1	0	1
Phase difference	0	0	0	0	1	0	2	1	0	1
False right vf	0	0	0	0	0	0	0	0	0	0
False left vf	0	0	0	0	0	0	0	0	0	0
Symmetry of corniculate tubercles during phonation	1	0	1	—	1	0	0	1	0	0
Symmetry of corniculate tubercles during rest	1	—	1	—	1	—	0	—	0	—

Abbreviation: vf, vocal fold.
Notes: Each parameter was graded with numbers 0–3 (0, unaffected; 1, slightly deviant; 2, moderately deviant; 3, highly deviant) from 0 (unaffected). Glottal open quotient was assessed from kymograms derived from the high speed digital imaging and is presented in %.

was meant to develop a new method of testing vocal endurance, subjects who varied in vocal experience needed to be tested, thus including subjects who had no vocal training at all and people who had some vocal training, but who were not professional singers or actors.

Another important factor when discussing vocal function and endurance is inherent personal traits. Patients in the voice clinic who suffer from functional dysphonia often possess certain personality traits (anxiety, depression, somatic complaints, and introversion) which could cause and/or maintain voice problems through behavioral consequences despite lack of structural or neurologic laryngeal pathology.[64–67] Eg, a trait could lead to inhibitory behavior of the larynx and elevated tension states, such as a feeling of uncertainty.[68] Individuals with vocal experience (trained voices) and patients who have undergone voice therapy are probably more wary than others of their vocal limitations. The subjects taking part in this study did not suffer from functional voice disorders per se, but one cannot disregard the possibility of some of them being at the end of the normophone spectrum closer to vocal dysfunction than others (eg, subject S5, who has allergies and medicates with omeprazole). The two subjects who endured the VLT for the full 30 minutes are both choir singers, who in addition give lectures often. They had not, however, had any speaking voice training (none of the other subjects had this particular voice profile). This back-

ground information could indicate that these individuals were accustomed to the feeling of a fairly strained voice, a sensation they might have assumed would pass after a rest.

Vocal recovery
In this study, possible effects of the VLT had to be monitored over time, however the rationale for follow-up laryngeal digital imaging and audio recordings taking place 2 days after the VLT is difficult to support or reject. The follow-up was performed to confirm that no damage to the physiology and/or function of the voice had resulted. Four days (the time span for the entire process) were deemed reasonable for the subjects to wear the voice accumulator and write frequent entries in their voice activity questionnaires. Very little is known regarding vocal recovery due to voice rest, without intervention. In fact Roy, p. 422, stated that, eg, postsurgical voice rest is a controversial area wherein nothing is known regarding optimal dose-response effects and that there is no proven standard for duration of voice rest.[69]

Hazards and clinical implications
Is testing for vocal loading dangerous? Could it harm the vocal function of the subjects? No previous VLT has proven to inflict any permanent damage to subjects' voice function (eg, a study by Kelchner et al[7]), although anterior glottal chinks of the vocal
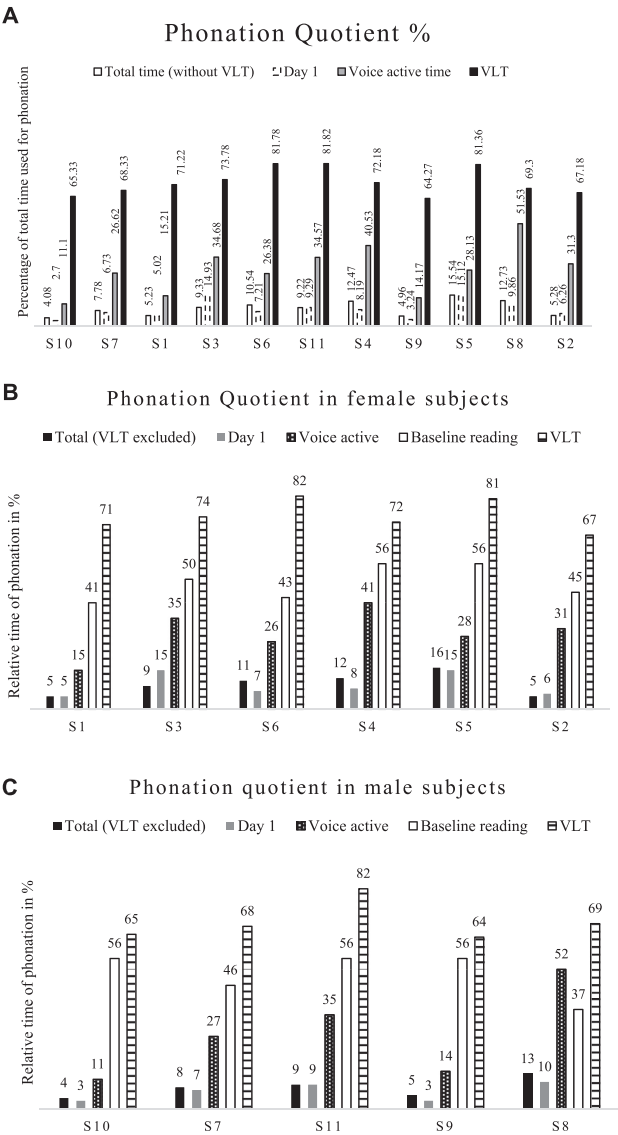
A

B

C

FIGURE 6. (**A**) Phonation quotient in all subjects (n = 11) in order of performance in VLT. Total phonation time (the data from VLT excluded); phonation time of the first day of accumulation; a vocally active period of at least 1 h; the phonation time of the baseline recording and phonation time for the VLT. (**B**) and (**C**) are alternatives to (**A**), separating female and male subjects' results, making them clearer. (**B**) Phonation quotient in female subjects (n = 6) from different activities: total phonation time (the data from VLT excluded); total phonation time of the first day of accumulation; a vocally active period of at least 1 h; the phonation time of the baseline recording and phonation time for the VLT. (**C**) Phonation quotient in male subjects (n = 5) from different activities: total phonation time (the data from VLT excluded); total phonation time of the first day of accumulation; a vocally active period of at least 1 h; the phonation time of the baseline recording and phonation time for the VLT.

folds have been induced from VLTs.[32] Increase of fundamental frequency and SPL of speech following the Lombard effect was seen in the present study, as eg, a study by Jonsdóttir et al[18] have found. Raised fundamental frequency and SPL will prove to be vocally demanding over time.[1] The design of the current VLT may in fact be less hazardous to the voice function than ones in previous research, as the subjects set their own time limit. The amount of time spent in a VLT like the current one could provide good clinical indications as to how far different individuals are prepared to test the limits of their voice. However, given the different time frames the subjects actually spent in the VLT (3–30 minutes) maybe we cannot trust vocally untrained individuals to know when they have reached their maximum time dose? N.B: it is impossible to be positively sure that vocal strain or fatigue was what prompted the subjects to discontinue their task (eg, n = 2 subjects complained of being tired by the ambient noise). It is essential to examine whether any subjects would endure the full 30-minute period if they are unaware of such a time limit.

Researchers run the risk of overinterpreting minimal changes in outcome of objective voice measurements preceding and following a VLT. Subjective observations of voice function have been proven somewhat difficult to measure objectively during prolonged speaking tasks. Noted changes may not be comparable with the severity of the subjects' experience.[32] The elevation of fundamental frequency recorded by VoxLog did not carry over to the post-VLT audio recording (compared with pre-VLT recordings). The elevation of speech loudness, however, did carry over to the post-VLT audio recording (compared with prerecordings). This can be an important example to use in the voice clinic, indicating the necessity to look at the bigger picture during voice assessments, than simply listening to audio recordings alone.

CONCLUSIONS

- The VLT appears to be successful in voice healthy subjects, as self-perceived vocal loading and objective changes to vocal function were attained through the chosen method without causing damage to voice function, thus simulating authentic vocal loading.
- Onset and recovery from self-perceived vocal loading were traceable through the voice activity questionnaire.
- The range of endurance in the VLT was an unexpected finding, indicating the complexity of vocal loading.

Acknowledgments

REFERENCES

1. Vilkman E. Occupational safety and health aspects of voice and speech professions… 28th World Congress of the International Association of Logopedics and Phoniatrics, Brisbane, Australia, 29th August to 2nd September 2004. *Folia Phoniatr Logop*. 2004;56:220–253.
2. Vintturi J, Alku P, Lauri E-J, et al. Objective analysis of vocal warm-up with special reference to ergonomic factors. *J Voice*. 2001;15:36–53.
3. Åhlander VL, Whitling S, Rydell R, Löfqvist A. Teacher's voice use in teaching environments: a field study using the Ambulatory Phonation Monitor. *J Voice*. 2014;28:841.e5–841.e15.
4. Böhme G. KlinikIn: *Sprach-, Sprech-, Stimm und Schluckstörungen*, 2003; Vol 1. Munich: Jena, Germany: Elsevier, Urban und Fischer; 2003.
5. Wendler J, Seidner W, Eysoldt U. *Lehrbuch der Phoniatrie und Pädaudiologie*. New York, NY: Stuttgart; 2005.
6. Hanschmann H, Gaipl C, Berger R. Preliminary results of a computer-assisted vocal load test with 10-min test duration. *Eur Arch Otorhinolaryngol*. 2011;268:309–313.
7. Kelchner LN, Toner MM, Lee L. Effects of prolonged loud reading on normal adolescent male voices. *Lang Speech Hear Serv Sch*. 2006;37:96–103.
8. Szabo Portela A, Hammarberg B, Södersten M. Speaking fundamental frequency and phonation time during work and leisure time in vocally healthy preschool teachers measured with a voice accumulator. *Folia Phoniatr Logop*. 2014;65:84–90.
9. Södersten M, Ternström S, Bohman M. Loud speech in realistic environmental noise: phonetogram data, perceptual voice quality, subjective ratings, and gender differences in healthy speakers. *J Voice*. 2005;19:29–46.
10. Doellinger M, Lohscheller J, McWhorter AJ, et al. Variability of normal vocal fold dynamics for different vocal loading in one healthy subject investigated by phonovibrograms. *J Voice*. 2009;23:175–181.
11. Laukkanen, A-M, Ilomäki I, Leppänen K, et al. Acoustic measures and self-reports of vocal fatigue by female teachers. *J Voice*. 2008;22:283–289.
12. Södersten M, Hammarberg B, Szabo A, et al. Vocal behavior and vocal loading factors for preschool teachers at work studied with binaural DAT recordings. *J Voice*. 2002;16:356–371.
13. Jilek C, Marienhagen J, Hacki T. Vocal stability in functional dysphonic versus healthy voices at different times of voice loading. *J Voice*. 2004; 18:443–453.
14. Hunter EJ, Titze IR. Quantifying vocal fatigue recovery: dynamic vocal recovery trajectories after a vocal loading exercise. *Ann Otol Rhinol Laryngol*. 2009;118:449–460.
15. Remacle A, Roche A, Morsomme D, et al. Vocal impact of a prolonged reading task at two intensity levels: objective measurements and subjective self-ratings. *J Voice*. 2012;26:e177–e186.
16. Lane HTB. The Lombard sign and the role of hearing in speech. *J Speech Hear Res*. 1971;14:677–709.
17. Rantala L, Vilkman E, Bloigu R. Voice changes during work: subjective complaints and objective measurements for female primary and secondary schoolteachers. *J Voice*. 2002;16:344–355.
18. Jónsdottir V, Laukkanen A, Vilkman E. Changes in teachers' speech during a working day with and without electric sound amplification. *Folia Phoniatr Logop*. 2002;54:282–287.
19. Lindstrom F, Ohlsson A, Sjöholm J, et al. Mean F0 values obtained through standard phrase pronunciation compared with values obtained from the normal work environment: a study on teacher and child voices performed in a preschool environment. *J Voice*. 2010;24:319–323.
20. Caraty MJ, Montacié C. Vocal fatigue induced by prolonged oral reading: analysis and detection. *Comput Speech Lang*. 2014;28:453–466.
21. Boucher VJ, Ayad T. Physiological Attributes of Vocal Fatigue and Their Acoustic Effects: a Synthesis of Findings for a Criterion-Based Prevention of Acquired Voice Disorders. *J Voice*. 2010;24:324–336.
22. Lohscheller J, Doellinger M, McWhorter AJ. Preliminary study on the quantitative analysis of vocal loading effects on vocal fold dynamics using phonovibrograms. *Ann Otol Rhinol Laryngol*. 2008;117(7 pt 1):484–493.
23. Niebudek-Bogusz E, Kotylo P, Sliwinska-Kowalska P. Evaluation of voice acoustic parameters related to the vocal-loading test in professionally active teachers with dysphonia. *Int J Occup Med Environ Health*. 2007;20:25–30.
24. Chang A, Karnell MP. Perceived phonatory effort and phonation threshold pressure across a prolonged voice loading task: a study of vocal fatigue. *J Voice*. 2004;18:454–466.

25. Laukkanen A, Järvinen K, Artkoski M, et al. Changes in voice and subjective sensations during a 45-min vocal loading test in female subjects with vocal training. *Folia Phoniatr Logop*. 2004;56:335–346.

26. Vintturi J, Alku P, Sala E, et al. Loading-related subjective symptoms during a vocal loading test with special reference to gender and some ergonomic factors. *Folia Phoniatr Logop*. 2003;55:55–69.

27. Vintturi J, Alku P, Lauri E, et al. The effects of post-loading rest on acoustic parameters with special reference to gender and ergonomic factors. *Folia Phoniatr Logop*. 2001;53:338–350.

28. Sihvo M, Laippala P, Sala E. A study of repeated measures of softest and loudest phonations. *J Voice*. 2000;14:161–169.

29. Vilkman E, Lauri E, Alku P, et al. Effects of prolonged oral reading on F0, SPL, subglottal pressure and amplitude characteristics of glottal flow waveforms. *J Voice*. 1999;13:303–312.

30. Sihvo M, Sala E. Sound level variation findings for pianissimo and fortissimo phonations in repeated measurements. *J Voice*. 1996;10: 262–268.

31. Gelfer MP, Andrews ML, Schmidt CP. Documenting laryngeal change following prolonged loud reading: a videostroboscopic study. *J Voice*. 1996;10:368–377.

32. Stemple JC, Stanley J, Lee L. Objective measures of voice production in normal subjects following prolonged voice use. *J Voice*. 1995;9:127–133.

33. Sherman D, Jensen PJ. Harshness and oral-reading time. *J Speech Hear Disord*. 1962;27:172–177.

34. Lindstrom F, Waye KP, Södersten M, et al. Observations of the relationship between noise exposure and preschool teacher voice usage in day-care center environments. *J Voice*. 2011;25:166–172.

35. Chen SH, Chiang S, Chung Y, et al. Risk factors and effects of voice problems for teachers. *J Voice*. 2009;24:183–190.

36. Lehto L, Alku P, Laaksonen L, et al. Changes in objective acoustic measurements and subjective voice complaints in call center customer-service advisors during one working day. *J Voice*. 2008;22:164–177.

37. Remacle A, Morsomme D, Finck C. Comparison of vocal loading parameters in kindergarten and elementary school teachers. *J Speech Lang Hear Res*. 2014;57:406–415.

38. Phyland DJ, Thibeault SL, Benninge MS, et al. Perspectives on the impact on vocal function of heavy vocal load among working professional music theater performers. *J Voice*. 2013;27:390.e31–390.e39.

39. Yiu EML, Wang G, Lo ACY, et al. Quantitative high-speed laryngoscopic analysis of vocal fold vibration in fatigued voice of young karaoke singers. *J Voice*. 2013;27:753–761.

40. Rosenthal AL, Lowell SY, Colton RH. Aerodynamic and acoustic features of vocal effort. *J Voice*. 2014;28:144–153.

41. Solomon NP. Vocal fatigue and its relation to vocal hyperfunction. *Int J Speech Lang Pathol*. 2008;10:254–266.

42. Ohlsson A-C, Brink O, Löfqvist A. A voice accumulation - validation and application. *J Speech Hear Res*. 1989;32:451–457.

43. Masuda T, Ikeda Y, Manako H, Komiyama S. Analysis of vocal abuse: fluctuations in phonation time and intensity in 4 groups of speakers. *Acta Otolaryngol*. 1993;113:547–552.

44. Švec JG, Popolo PS, Titze IR. Measurement of vocal doses in speech: experimental procedure and signal processing. *Logoped Phoniatr Vocol*. 2003;28:181.

45. Cheyne HA, Hanson H, Genereux RP, Stevens KN, Hillman RE. Development and testing of a portable vocal accumulator. *J Speech Lang Hear Res*. 2003;46:1457–1468.

46. Szabo A, Hammarberg B, Hakansson A, et al. A voice accumulator device: evaluation based on studio and field recordings. *Logoped Phoniatr Vocol*. 2001;26:102–117.

47. Szabo A, Hammarberg B, Granqvist S, et al. Methods to study pre-school teachers' voice at work: simultaneous recordings with a voice accumulator and a DAT recorder. *Logoped Phoniatr Vocol*. 2003;28:29–39.

48. Vogel AP, Fletcher J, Maruff P. Acoustic analysis of the effects of sustained wakefulness on speech. *J Acoust Soc Am*. 2010;128:3747–3756.

49. Arbetsmiljöverket TSWHA. Buller. In: Arbetsmiljöverket TSWHA, ed. *AFS 2005:16*. Solna, Sweden: Arbetsmiljöverket, The Swedish Work Health Authority; 2005.

50. Bilger RC, Rzcezkowski C, Nuetzel I, et al. Standardization of a test of speech perception in noise (SPIN). *J Acoust Soc Am*. 1979;65(suppl 1):S98.

51. Alton Everest F. *Master Handbook of Acoustics*. 4th ed. New York, NY: McGraw-Hill; 2001:599.

52. Fletcher H, Munson WA. Loudness, its definition, measurement and calculation. *J Acoust Soc Am*. 1933;5:82.

53. 3352, I.T., Acoustics - assessment of noise with respect to its effect on the intelligibility of speech. Geneva, Switzerland: International Organization for Standardization.

54. Pearsons K, Bennett R, Fidell S. *Speech Levels in Various Noise Environments*. Washington, DC: Office of Health and Ecological Effects, Office of Research and Development, U.S. Environmental Protection Agency; 1977.

55. Dejonckere PH, Bradley P, Clemente P, et al. A basic protocol for functional assessment of voice pathology, especially for investigating the efficacy of (phonosurgical) treatments and evaluating new assessment techniques. Guideline elaborated by the Committee on Phoniatrics of the European Laryngological Society (ELS). *Eur Arch Otorhinolaryngol*. 2001;258:77–82.

56. Åhlander VL, Rydell R, Löfqvist A. Speaker's Comfort in Teaching Environments: Voice Problems in Swedish Teaching Staff. *Journal of Voice*. 2011;25:430–440.

57. Lofqvist A, Mandersson B. Long-time average spectrum of speech and voice analysis. *Folia Phoniatr (Basel)*. 1987;39:221–229.

58. Åhlander VL, Rydell R, Löfqvist A. How do teachers with self-reported voice problems differ from their peers with self-reported voice health? *J Voice*. 2012;26:e149–e161.

59. Jacobson BH, Johnson A, Grywalski C, et al. The Voice Handicap Index (VHI):development and validation. *Am J Speech Lang Pathol*. 1997;6: 66–70.

60. Lyberg-Åhlander V, Rydell R, Eriksson J, et al. Throat related symptoms and voice: development of an instrument for self assessment of throat-problems. *BMC Ear Nose Throat Disord*. 2010;10:5–12.

61. Cohen J. Statistical power analysis. *Curr Dir Psychol Sci*. 1992;1:98–101.

62. Titze IR, Hunter EJ, Švec JG. Voicing and silence periods in daily and weekly vocalizations of teachers. *J Acoust Soc Am*. 2007;121:469–478.

63. Hunter EJ, Titze IR. Variations in intensity, fundamental frequency, and voicing for teachers in occupational versus nonoccupational settings. *J Speech Lang Hear Res*. 2010;53:862–875.

64. Aronson AE. *Clinical Voice Disorders: An Interdisciplinary Approach*. 3rd ed. New York, NY: Thieme; 1990.

65. House AO, Andrews HB. Life events and difficulties preceding the onset of functional dysphonia. *J Psychosom Res*. 1988;32:311–319.

66. Vanhoudt I, Wellens WAR, de Jong FICRS, et al. The background biopsychosocial status of teachers with voice problems. *J Psychosom Res*. 2008;65:371–380.

67. Meulenbroek LFP, Thomas G, Kooijman PGC, et al. Biopsychosocial impact of the voice in relation to the psychological features in female student teachers. *J Psychosom Res*. 2010;68:379–384.

68. Roy N, Bless DM, Heisey D. Personality and voice disorders: a superfactor trait analysis. *J Speech Lang Hear Res*. 2000;43:749–768.

69. Roy N. Optimal dose-response relationships in voice therapy. *Int J Speech Lang Pathol*. 2012;14:419–423.

Paper II

Long-Time Voice Accumulation During Work, Leisure, and a Vocal Loading Task in Groups With Different Levels of Functional Voice Problems

*Susanna Whitling, *Viveka Lyberg-Åhlander, and *,†Roland Rydell, *†Lund, Sweden

Summary: Objective. The study aimed to examine the vocal behavior and self-assessed vocal health in women with varying everyday vocal load and functional voice problems, including patients with functional dysphonia, in three conditions: work, leisure, and a vocal loading task (VLT).
Study Design. This is a longitudinal controlled, clinical trial.
Methods. Fifty (n = 50) female subjects were tracked during 7 days' voice accumulation accompanied by a voice health questionnaire, containing general assessments with visual analogue scale and specific voice health questions. Subjects were divided into four vocal subgroups according to everyday vocal load and functional vocal complaints. Accumulation time was divided into three conditions: a VLT, work, and leisure. The following behavioral parameters were measured: (1) relative phonation time (%), (2) phonatory sound pressure/voice level (dB sound pressure level), (3) ambient noise level (dB sound pressure level), and (4) phonatory fundamental frequency (Hz).
Results. Patients with functional dysphonia reported significantly higher specific voice problems across conditions and worse general voice problems during work and leisure than other groups. Women with high everyday vocal load and voice complaints showed higher phonation times and fundamental frequency during work than voice healthy controls. They also reported the highest incidence of general voice problems in the VLT.
Conclusions. Vocal loading relates to prolonged phonation time at high fundamental frequencies. Patients with functional dysphonia experience general and specific voice problems permanently, whereas women with everyday vocal load and voice complaints recover during leisure. This may explain why the latter group does not seek voice therapy.
Key Words: Voice accumulation–Long-time voice measurement–Vocal loading–Functional dysphonia–Occupational dysphonia.

INTRODUCTION

Vocal loading relates to vocal behavior within an individual as well as vocal context surrounding an individual. It is unclear what is meant by "vocal loading" in general and how it relates to vocal *overload*. Amplified and prolonged vocal intensity seems to be key elements leading to functional voice complaints.[1] Competing with ambient noise is considered a vocal loading activity, and female speakers are more prone to vocal overload in such conditions. The activity leads to significant increase of voice intensity, fundamental frequency, phonation ratio, and hyperfunctional voice behavior, as well as in vocal effort and strain.[2,3] People who rely on their voice for work and who do not have rigorous voice training, eg, teachers, often exhibit high everyday vocal load (high vocal dose competing with ambient noise). They are overrepresented in voice clinics, exhibiting functional voice problems[4] that occur in or are exacerbated at the workplace,[3,5–11] ie, presenting with occupational dysphonia. These conditions also hold true for other occupations. When classifying functional voice problems, it is important to investigate patients' vocal behaviors leading to their vocal habits.[12] Although it is already shown, for example, that teachers presenting vocal complaints phonate to greater extent than others, ie, they show a higher relative phonation time,[13,14] vocal *behavior* has

not been systematically, thoroughly compared in populations representing or manifesting different degrees of vocal loading and functional voice problems.

Inclusion criteria for clinical functional dysphonia are not yet established. There is an urgent need for high-quality outcome measures in clinical research to help voice clinicians and affected patients fully grasp underlying vocal difficulties associated with functional dysphonia.[15] As a diagnosis of exclusion, functional dysphonia is characterized by a lack of organic and neurologic laryngeal pathology by impaired regulation of laryngeal muscles.[12,16] Description of functional dysphonia requires multifaceted clinical knowledge and understanding. Elemental vocal behavior, for example, needs to be investigated in different contexts.[17] In fact, functional dysphonia can even be called behavioral dysphonia, and behavioral aspects of functional dysphonia need further investigation.[18] Such an investigation should be performed unobtrusively in contexts in which patients with functional dysphonia use their voice. Different contexts present different phonatory challenges for speakers, such as acoustics, ambient noise, dry air, group size and type of listeners, and the function of present speech.[13] The technology offered is voice accumulation (eg, Van Stan et al, [19]) combined with essential self-assessment tools.[20]

Objective

The study aimed to examine vocal behavior and self-assessed vocal health in a population with varying everyday vocal load and functional voice problems, including patients with functional dysphonia, under three different conditions: (1) work, (2) leisure, and (3) a vocal loading task (VLT).

Research questions

(1) How does vocal behavior in women with diagnosed functional dysphonia differ from others who also experience voice problems, but who do not seek therapy? (How do these two groups compare with control groups?)

(2) How does self-assessed vocal health in women with diagnosed functional dysphonia differ from others who also experience voice problems, but who do not seek therapy? (How do these two groups compare with control groups?)

METHODS

Subjects

Vocal behavior and vocal health were examined in four vocal subgroups in three contextual conditions: in a VLT, during working hours, and during leisure time. Fifty (n = 50) female subjects were recruited to four (n = 4) vocal subgroups representing different degrees of everyday vocal loading and functional voice problems: FD: n = 20 patients with functional dysphonia; HLC: n = 10 women with high everyday vocal load with voice complaints; HLNC: n = 10 women with high everyday vocal load with no voice complaints; and C: n = 10 women serving as voice healthy controls, with low everyday vocal load and no voice complaints. Subjects in the FD group stood on a voice therapy spectrum: some not having started, some having received a little, and some having finished one round of voice therapy. All expressed a wish for additional voice therapy. The subjects in the HLC group experienced self-perceived voice problems, for which they had not sought medical care. The distinction was determined *ad modum* Lyberg-Åhlander et al,[3] ie, each subject rating the statement "I have problems with my voice?" on a scale from 0 (=completely disagree) to 4 (=completely agree). Subjects with high occupational vocal load reporting ≥2 on this scale were placed in the HLC group. The fourth, remaining group comprised 10 (n = 10) voice healthy controls, with self-perceived low/no occupation vocal load. Patients with functional dysphonia were recruited specifically for the current study through medical voice care professionals (speech and language pathologists and phoniatricians) in and around Skåne University Hospital, in southern Sweden. All patients were cisgender females, thus identifying with the sex they were born as, having no desire for a fundamentally different function of their habitual voice. They were all of working age (or newly retired), with no organic voice disorders at laryngeal examination. No subject had any known hearing impairment. The group of patients with functional dysphonia formed a sample basis upon recruiting the other three groups, matching for age, general health, and life situation. Due to this recruitment method, smokers were not excluded from the study as they occurred in small number in the group with functional dysphonia. Recruitment for the two groups with high occupational workload (HLC and HLNC) was made from the subject base in Lyberg-Åhlander et al.[3] Subjects who in this previous study had written consent to take part in any further studies were asked (the question on voice problems earlier mentioned was reiterated for the current study, as 5 years had passed). Subjects for the voice healthy control group were mainly recruited through the patients with functional dysphonia, who were asked to bring a voice healthy female peer of the same age. Additional recruitment for the high vocal load (HLC, HLNC) and control (C) groups was made through inquiry on social media. No subjects were told any aims for the study, and all subjects gave written informed consent to take part in the current study. For age and group distribution, see Table 1.

Voice accumulation with *VoxLog*

Each subject underwent seven days' long-time measurement of voice using the portable voice accumulator *VoxLog* (Sonvox AB, Umeå, Sweden). *VoxLog* provided information on relative phonation time, ie, the amount of time during measurement during which each subject phonated, measured through skin vibration with an accelerometer, fundamental frequency, measured by said accelerometer and sound pressure level of phonation, and of ambient noise, measured by a microphone. The accelerometer and microphone were placed in a neck collar. The microphone does not correct to a standard distance, thus sound pressure level is reported *ad modum* Södersten et al, ie, ca 7 dB higher compared with measurements taken at a 30-cm distance.[21] All data from *VoxLog* were labeled with one of three conditions: (1) work, (2) leisure, and (3) a VLT. These conditions were categorized by each subject and noted in a voice health questionnaire throughout voice accumulation.

Self-assessment of vocal health with voice health questionnaire

A voice activity questionnaire based on the voice handicap index[22] accompanied the 7-day voice accumulation. It contained 18 voice health questions on a 5-point scale (0 = none/not at all, 1 = low/occasionally, 2 = some/sometimes, 3 = high/often, 4 = very high/nonstop). Ten voice questions (10VQ) were expected to fluctuate during a day, thus filled out four times a day (28 times in total). These were the following: 1: This is my current stress level, 2: My voice feels fatigued, 3: I need to clear my throat, 4: I need to cough, 5: My throat/neck (same word in Swedish: "hals") feels tense, 6: I am hoarse, 7: I am having a hard time making myself heard (like at a party), 8: My voice can suddenly change when I speak. 9: It is effortful to get my voice working, 10: I have a feeling of discomfort in my throat/neck. The remaining eight questions (8VQ) were filled out once a day (seven times in total). These were the following: 11: I run out of breath when I speak, 12: My voice problems affect my private economy, 13: My voice problems restrict my private and social life, 14: My voice makes it hard for others to hear what I am saying, 15: Others ask me what is wrong with my voice, 16: I feel handicapped because of my voice, 17: I feel left out of conversations because of my voice, 18: I am worried by my voice problems. A 100-mm visual analogue scale (VAS) with the question "how does your voice feel right now" (0 = no voice problems, 100 = maximal voice problems) was filled in four times a day (28 in total). The voice health questionnaire was structured to ensure point prevalence; therefore, each entry was followed by a notation of the time of day and current condition (VLT, work, or leisure). The 8VQ was only filled out during leisure (once a day, in the evening). Working and leisure hours were noted in the voice health questionnaire. The 10VQ and 8VQ were checked for internal consistency separately using Cronbach's α. A test form (T1) was filled out by each

TABLE 1.
Distribution of Fifty (n = 50) Female Subjects Divided Into Three Vocal Subgroups and One Voice Healthy Control Group

Subgroup	Acronym	n	Age	Vocal Health	Vocal Load	Occupation	Medication Status
Functional dysphonia	FD	20	44 (22–70, SD: 14.7)	Patients with medically diagnosed functional dysphonia in different stages of voice therapy	Differing occupational vocal load	Teacher (9), retired former teacher (4), student (3), administrator (2), counselor (1), priest (1)	Asthma (5), allergies (4), ulcer (3), diabetes (1), high blood pressure (1), migraine (1), reflux (1)
High load, complaints	HLC	10	50 (33–62, SD: 10.6)	No medical care for declared functional/ occupational voice problems	High occupational vocal load	Teacher (9), singing teacher (1)	High blood pressure (2), allergies (1), asthma (1), burnout (1)
High load, no complaints	HLNC	10	48 (32–65, SD: 9.8)	No declared voice problems	High occupational vocal load	Teacher (6), medical doctor (2), secretary (1), guide (1)	Depression (2), allergies (2), hypothyroidism (1), asthma (1), rheumatism (1)
Voice healthy controls	C	10	45 (25–67, SD: 11.3)	No declared voice problems	Low/no occupational vocal load	Administrator (3), retired (1), student (1), medical doctor (1), chemist (1), translator (1), train driver (1), acoustical consultant (1)	Depression (2), allergies (2), asthma (1)

Abbreviation: SD, standard deviation.

subject 20 minutes prior to voice accumulation start, and a retest form (T2) was filled out 5 minutes prior to accumulation start. The internal consistency for 10VQ was α = .92 (T1 item mean = 1.33, T2 item mean = 1.17). The internal consistency for 8VQ was α = .94 (T1 item mean = .61, T2 item mean = .60). Both exceed acceptable minimum consistency.[23]

Vocal loading task
Each subject underwent a VLT midweek/mid-voice accumulation, which served as a control condition where heavy vocal loading (prolonged phonation at high sound pressure levels) would definitely take place. The VLT was set up *ad modum* Whitling et al.[24] There were slight changes to the setup as follows: (1) Phonatory signal-to-noise ratio was monitored during the VLT, ensuring minimum phonation sound pressure levels of 85 dB A were met. It was monitored with real-time phonetograms (Phog Interactive Phonetography System 2.5 for PC, Hi-Tech Medical, Täby, Sweden) using a head-mounted microphone (MKE 2, no 09_1, Sennheiser, Wedemark, Germany) calibrated at 94 dB sound pressure level at a distance of 15 cm from the mouth (converted to 30 cm). (2) The subjects were free to choose a text for loud reading in the VLT. Most of them chose the alternative text chosen by the test leader: an easy to read excerpt from a book on the history of kings and queens of Sweden. (3) Oral instructions on entering the VLT were modified: the subjects were not informed of the 30-minute maximum participation time in order to prevent subjects terminating or not terminating the VLT for any other reason than a distinct sensation of discomfort from the throat.

Statistical analyses
Statistical analyses were performed with *IBM SPSS Statistics 23* (SPSS Inc., Chicago, IL, USA). All samples were explored for normal distribution and statistical analyses were chosen accordingly. A Shapiro-Wilk's test[25,26] of skewness and kurtosis[27–29] and a visual inspection of their histograms, normal Q-Q plots, and box plots were performed for each parameter examined as basis for choice of statistical analyses (see each result). A statistical power analysis of sample sizes was based on vocal self-assessment results from Whitling et al.[24] It showed that in order for an effect size of .5 to be detected at a 97% chance as significant at the 1% level,[30] a sample of 10 participants in each compared group was needed.

Ethical research approval
The study was approved by the Regional Ethical Review Board in Lund, Sweden on April 25, 2013 (#2013/174).

RESULTS
Time in vocal loading task
Sample characteristics and statistical analyses
The analysis of sample characteristics revealed that distribution was not normal for all vocal subgroups for minutes spent in the VLT, disqualifying parametric analyses. The data were explored using the Kruskal-Wallis H test, which showed no significant group differences for time in VLT.

All vocal subgroups spent a similar amount of time (minutes) in the VLT: FD (M = 27, Med = 30), HLC (M = 25, Med = 30), HLNC (M = 26, Med = 30), and C (M = 28, Med = 30).

Vocal behavior observed through voice accumulation

Sample characteristics and statistical analyses
The inspection of the distribution of all voice accumulation parameters (relative phonation time, phonatory sound pressure level, and fundamental frequency) showed normality, qualifying parametric analyses. A two-way repeated analysis of variance tested the relative phonation time, phonatory sound pressure level, and fundamental frequency separately, in the three conditions—VLT, work, and leisure—across four vocal subgroups (FD, HLC, HLNC, C). The analyses applied Bonferroni adjustment for multiple vocal subgroups (4) and conditions (3) and with Greenhouse-Geisser correction as a consequence of violation of assumption of sphericity for two variables: relative phonation time: $\chi^2(2) = 29.980$, $P < 0.01$, and fundamental frequency: $\chi^2(2) = 29.408$, $P < 0.01$. Differences within groups were explored using paired samples t tests.

Relative phonation time
Relative phonation time differed significantly across all three conditions work, leisure, and VLT ($F[1.3, 44.1] = 466.09$, $P < 0.01$, $\eta^2 = .931$). A pairwise comparison of conditions, *work* to *leisure*, *work* to *VLT*, and *leisure* to *VLT*, all showed $P < 0.01$. There was no significant overall interaction between vocal subgroup affiliation and condition. A pairwise comparison of each vocal subgroup within each condition showed a significant difference between HLC and C in the work condition ($P = 0.034$). A comparison of work to leisure within each subgroup showed significantly higher phonation time during work than leisure for HLC ($t = 4.721$, $P = 0.002$) and for HLNC ($t = 4.634$, $P = 0.004$). The mean scores of relative phonation time under three conditions across four vocal subgroups are shown in Figure 1. A summary of results for vocal subgroups and conditions is found in Table 2.

Phonatory sound pressure level
Phonatory sound pressure level (voice level) differed significantly across the three conditions—VLT, work, and leisure ($F[2, 70] = 75.568$, $P < 0.01$, $\eta^2 = .686$). Pairwise comparisons *work* (M = 86 dB sound pressure level) to *VLT* (M = 94.5), and *leisure* (M = 86.5) to *VLT*, both showed significance at $P < 0.01$. A pairwise comparison of conditions *work* to *leisure* was not statistically significant. The overall interaction of affiliation to vocal subgroup and condition was not statistically significant. The summary of the results for vocal subgroups and conditions is found in Table 2.

Ambient noise under three conditions: work, leisure, and a vocal loading task
The sound pressure level of ambient noise (noise level) was compared with the phonatory sound pressure level (voice level) as measured by *VoxLog* across the four vocal subgroups within one condition at a time using a two-way repeated measures analysis of variance. The *work* condition gave general voice levels
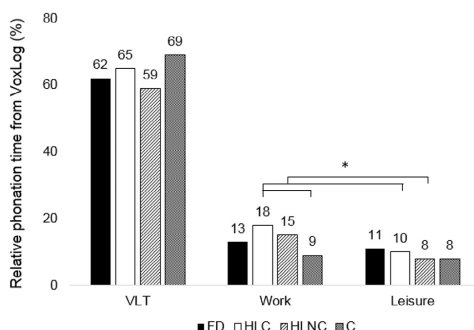

Relative phonation time across three conditions

FIGURE 1. Relative phonation time (%) measured by *VoxLog* accelerometer under three conditions: vocal loading task (VLT), work, and leisure, and a VLT across four (n = 4) vocal subgroups (FD, HLC, HLNC, and C). Relative phonation time in VLT was significantly higher than work and leisure for all groups; other significant group differences and differences within groups are marked with bars.

(M = 86 dB sound pressure level), which were significantly higher than the ambient noise levels (M = 71 dB sound pressure level): $F(1, 36) = 328.015$, $P < 0.01$, $\eta^2 = .901$. There were no significant differences between the vocal subgroups: all groups significantly overpowered the sound pressure level of their working environments,[31] which were comparable across groups. The same pattern was repeated for the *leisure* condition and the VLT, giving general voice levels that were significantly higher than the ambient noise levels (*leisure*: M = 72 dB sound pressure level, $F[1, 41] = 389.092$, $P < 0.01$, $\eta^2 = .905$; and *VLT*: M = 84 dB sound pressure level, $F[1, 42] = 283,456$, $P < 0.01$, $\eta^2 = .871$). There were no significant differences between vocal subgroups: all groups managed to phonate at higher intensity levels than their ambient noise levels.

Fundamental frequency
Fundamental frequency differed significantly overall across conditions, $F(1.3, 44.1) = 16.247$, $P < 0.01$, $\eta^2 = .317$. A pairwise comparison of conditions *work* to *leisure* was not statistically significant. Pairwise comparisons of *work* (M = 246 Hz) to *VLT* (M = 283 Hz), and *leisure* (M = 249 Hz) to *VLT*, both showed the *VLT* leading to significantly higher fundamental frequency in all groups for *work* to *VLT* (overall: $P < 0.01$, FD: $t = -2.597$, $P = 0.019$, HLC: $t = -2.903$, $P = 0.023$, HLNC: $t = -2.08$, $P = 0.083$, C: $t = -3.892$, $P = 0.008$). When *leisure* was compared with *VLT*, post hoc paired samples t tests showed significant differences for HLC ($t = -5.576$, $P < 0.01$) and C ($t = -3.337$, $P = 0.012$). A pairwise comparison of vocal subgroups within each condition showed one significant difference between HLC and C in the *work* condition ($P = 0.024$). The overall interaction of affiliation to vocal subgroup and condition was not statistically significant. The mean scores of fundamental frequency under three conditions across four vocal subgroups are

TABLE 2.
Overview of All Results, Noting Significant Differences in Comparisons Between Groups and Conditions

Group	Measure	VLT	Work	Leisure	Comparison *Groups*	Comparison *Conditions*
FD	Phonation time (%)	62	13	11	NS	VLT higher than work and leisure
	Voice Level (dB SPL)	95	86	84	NS	VLT higher than work and leisure
	Noise level (dB SPL)	85	69	73	NS	VLT higher than work and leisure
	Fundamental frequency (Hz)	277	243	250	NS	VLT higher than work
	10VQ (maximum of 40 points)	21	15	13	Higher than HLNC and C in VLT Higher than all groups during work and leisure	VLT higher than work and leisure
	VAS (100 mm)	66	52	49	Higher than HLNC and C during VLT and work Higher than all groups during leisure	VLT higher than work and leisure
	Stress (0–4 points)	1.2	1.9	1.4	Higher than C in VLT Higher than all groups during work and leisure	
HLC	Phonation time (%)	65	18	10	Higher than C during work	VLT higher than work and leisure Work higher than leisure
	Voice level (dB SPL)	94	88	86	NS	VLT higher than work and leisure
	Noise level (dB SPL)	83	72	71	NS	VLT higher than work and leisure
	Fundamental frequency (Hz)	315	268	254	Higher than C during work	VLT higher than work
	10VQ (maximum of 40 points)	16	7	4	Higher than C in VLT Lower than FD during work and leisure Higher than HLNC and C during work	VLT higher than work and leisure
	VAS (100 mm)	67	32	17	Higher than C across all conditions Higher than HLNC and C during work Lower than FD during leisure	VLT higher than leisure
	Stress (0–4 points)	1.1	1.1	0.7	Higher than C during VLT and work Lower than FD during work and leisure	Work higher than VLT and leisure
HLNC	Phonation time (%)	59	15	8	NS	VLT higher than work and leisure Work higher than leisure
	Voice level (dB SPL)	94	88	86	NS	VLT higher than work and leisure
	Noise level (dB SPL)	83	72	71	NS	VLT higher than work and leisure
	Fundamental frequency (Hz)	268	243	250	NS	VLT higher than work
	10VQ (maximum of 40 points)	12	3	3	Lower than FD in VLT Lower than FD during work and leisure Lower than HLC during work	VLT higher than work
	VAS (100 mm)	31	9	6	Lower than FD during VLT and work Lower than HLC during work Lower than FD during leisure	VLT higher than leisure
	Stress (0–4 points)	0.6	1.3	0.5	Lower than FD during work and leisure	Work higher than leisure
C	Phonation time (%)	69	9	8	Lower than HLC during work	VLT higher than work and leisure
	Voice level (dB SPL)	94	88	86	NS	VLT higher than work and leisure
	Noise level (dB SPL)	83	72	71	NS	VLT higher than work and leisure
	Fundamental frequency (Hz)	274	228	242	Lower than HLC during work	VLT higher than work
	10VQ (maximum of 40 points)	5	2	1	Lower than FD and HLC during VLT Lower than FD during work and leisure	VLT higher than work and leisure
	VAS (100 mm)	12	3	2	Lower than FD and HLC across all conditions	NS
	Stress (0–4 points)	0.4	0.5	0.4	Lower than FD and HLC during VLT and work Lower than FD during leisure	NS

Abbreviations: 10VQ, ten voice questions; NS, not significant; SPL, sound pressure level; VAS, visual analogue scale; VLT, vocal loading task.
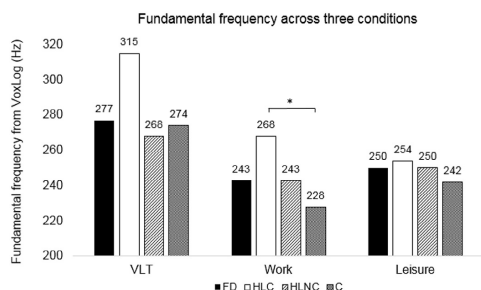
FIGURE 2. Fundamental frequency (Hz) measured by *VoxLog* accelerometer under three conditions: vocal loading task (VLT), work, and leisure, and a VLT across four (n = 4) vocal subgroups (FD, HLC, HLNC, and C). Fundamental frequency in the VLT was significantly higher than work and leisure for all groups; the only significant group difference between HLC and C during work is marked with a bar.

shown in Figure 2. The summary of results for vocal subgroups and conditions is found in Table 2.

Self-assessed vocal health recorded in voice health questionnaire

Sample characteristics and statistical analyses

The analysis of sample characteristics revealed that distribution was not normal for any vocal subgroup for any voice activity parameter measured in the voice health questionnaire, disqualifying parametric analyses. As the eight voice questions were not filled out across all three conditions, they were not used for comparison. The distribution of scores from 10 voice questions (10VQ) in four (n = 4) vocal subgroups across three conditions (VLT, work, and leisure) is shown in Figure 3, and the maximum score was 40. The distribution of VAS scores is shown in Figure 4, and the maximum score was 100. The distribution of stress scores is shown in Figure 5, and the maximum score was 4. The correlation of self-assessment scores (10VQ and VAS) was explored using Spearman's ρ. The summary of results for vocal subgroups and conditions is found in Table 2.
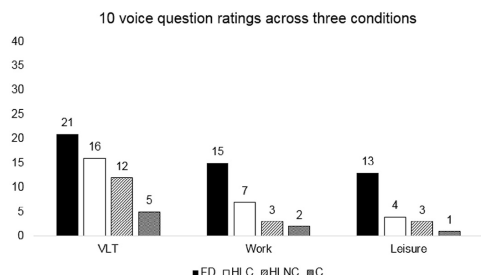

FIGURE 3. Mean scores from 10 voice questions (maximum score of 40) across three conditions: vocal loading task, work, and leisure in four vocal subgroups (FD, HLC, HLNC, and C).
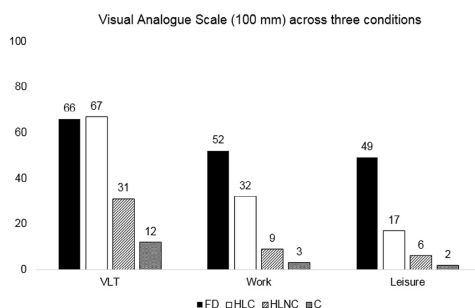

FIGURE 4. Mean scores from visual analogue scale (VAS) (100-mm VAS) across three conditions: vocal loading task, work, and leisure in four vocal subgroups (FD, HLC, HLNC, and C).

Vocal loading task condition

A Kruskal-Wallis H test uncovered statistically significant differences between group means for 10VQ, VAS, and stress in the VLT condition, filled out promptly after terminating the task: 10VQ: $\chi^2(3) = 18.326$, $P < 0.01$, VAS: $\chi^2(3) = 25.259$, $P < 0.01$, stress: $\chi^2(3) = 8.145$, $P = 0.043$. Post hoc comparisons of 10 voice questions after the VLT with the Mann-Whitney U test showed that FD (M = 21) differed with statistical significance from HLNC (M = 12) ($U = 23$, $P < 0.002$) and C (M = 5) ($U = 19$, $P = 0.001$). HLC differed significantly from C ($U = 15.5$, $P = 0.013$). Post hoc comparisons of VAS scores after the VLT with the Mann-Whitney U test showed FD (M = 66) differed with statistical significance from HLNC (M = 31) ($U = 14$, $P < 0.01$) and from C (M = 12) ($U = 2$, $P < 0.01$). HLC differed significantly from C ($U = 9.5$, $P = 0.002$). Post hoc comparisons of stress scores in the VLT with the Mann-Whitney U test showed that C (M = .4) differed significantly from FD (M = 1.2) ($U = 47$, $P = 0.02$) and HLC (M = 1.1) ($U = 24$, $P = 0.03$). There were no other significant group differences.
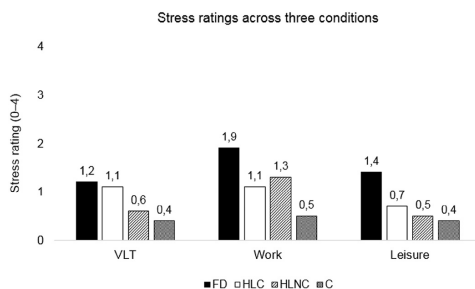

FIGURE 5. Mean scores from rating of current stress level (0 = none, 1 = low, 2 = some, 3 = high, 4 = very high) across three conditions: vocal loading task, work, and leisure in four vocal subgroups (FD, HLC, HLNC, and C).

TABLE 3.
Results From Post Hoc Testing With a Bonferroni-Adjusted Wilcoxon Signed-Rank Test, Comparing Median Scores From 10 Voice Questions Across Three Conditions (VLT, Work, and Leisure) in Four Vocal Subgroups (FD, HLC, HLNC, and C)

Group	VLT to Work	Sig.	VLT to Leisure	Sig.	Work to Leisure	Sig.
FD	21 to 15	$P = 0.001$*	21 to 13	$P < 0.01$*	15 to 13, NS	NS
HLC	16 to 7	$P = 0.013$*	16 to 4	$P = 0.005$*	7 to 4*	$P = 0.012$*
HLNC	12 to 3	$P = 0.012$*	12 to 1	$P = 0.017$*	3 to 1, NS	NS
C	5 to 1.5	$P = 0.012$*	5 to 1	$P = 0.011$*	1.5 to 1, NS	NS

* Significant differences at $P < 0.017$.
Abbreviations: NS, not significant; VLT, vocal loading task.

Work condition

A Kruskal-Wallis H test uncovered statistically significant differences between group means for 10VQ, VAS, and stress in the work condition: 10VQ: $\chi^2(3) = 28.238$, $P < 0.01$, VAS: $\chi^2(3) = 21.036$, $P < 0.01$, stress: $\chi^2(3) = 16.534$, $P < 0.01$. Post hoc comparisons of 10VQ during work with the Mann-Whitney U test showed FD (M = 15) differing with statistical significance from all other groups: HLC (M = 7) ($U = 41.5$, $P = 0.042$), HLNC (M = 3) ($U = 6$, $P < 0.01$), and C (M = 2) ($U = 0$, $P < 0.01$). HLC differed statistically significantly from HLNC ($U = 15.5$, $P = 0.027$) and from C ($U = 7$, $P = 0.003$). Post hoc comparisons of VAS during work with the Mann-Whitney U test showed significantly higher scores for FD (M = 52) than for HLNC (M = 9) ($U = 8.5$, $P = 0.001$) and C (M = 3) ($U = 3$, $P < 0.01$). HLC (M = 32) scores significantly higher than HLNC ($U = 7$, $P = 0.015$) and C ($U = 5$, $P = 0.008$). Post hoc comparisons of stress scores during work with the Mann-Whitney U test showed FD (M = 1.9) scoring significantly higher than all other groups: HLC (M = 1.1) ($U = 26$, $P = 0.005$), HLNC (M = 1.3) ($U = 48$, $P = 0.043$), and C (M = 0.5) ($U = 12$, $P = 0.001$). HLC differed significantly from C ($U = 13$, $P = 0.026$). There were no other significant group differences.

Leisure condition

A Kruskal-Wallis H test showed significant differences between group means for 10VQ, VAS, and stress in the leisure condition: 10VQ: $\chi^2(3) = 27,976$, $P < 0.01$, VAS: $\chi^2(3) = 22.793$, $P < 0.01$, stress: $\chi^2(3) = 15.124$, $P = 0.002$. Post hoc comparisons of 10VQ during leisure with the Mann-Whitney U test showed FD (M = 13) differing with statistical significance from all other groups: HLC (M = 4) ($U = 26.5$, $P = 0.002$), HLNC (M = 3) ($U = 6$, $P < 0.01$), and C (M = 1) ($U = 6$, $P < 0.01$). HLC differed significantly from C ($U = 23$, $P = 0.041$). Post hoc comparisons of VAS scores during leisure with the Mann-Whitney U test showed FD (M = 49) differing with statistical significance from all other groups: HLC (M = 17, $U = 49$, $P = 0.035$), HLNC (M = 6, $U = 8$, $P < 0.01$), and C (M = 2, $U = 8$, $P < 0.01$). HLC differed significantly from C ($U = 14$, $P = 0.021$). Post hoc comparisons of stress scores during leisure with the Mann-Whitney U test showed FD (M = 1.4) differed with statistical significance from all groups: HLC (M = .7, $U = 40$, $P = 0.01$), HLNC (M = .5, $U = 41.5$, $P = 0.014$), and C (M = .4, $U = 25.5$, $P = 0.001$). There were no other significant group differences.

Comparison across conditions and statistical analyses

When the three conditions VLT, work, and leisure were examined with Friedman's test within each vocal subgroup, there were statistically significant differences in ratings of 10VQ for all groups: FD: $\chi^2(2) = 14.000$, $P = 0.001$, HLC: $\chi^2(2) = 15.600$, $P < 0.01$, HLNC: $\chi^2(2) = 9.750$, $P = 0.008$, C: $\chi^2(2) = 13.067$, $P = 0.001$; in VAS scores for all groups: FD: $\chi^2(2) = 20.588$, $P < 0.01$, HLC: $\chi^2(2) = 7.000$, $P = 0.030$, HLNC: $\chi^2(2) = 10.571$, $P = 0.005$, C: $\chi^2(2) = 8.960$, $P = 0.011$; and in stress ratings for FD: $\chi^2(2) = 9.509$, $P = 0.009$ and HLC: $\chi^2(2) = 6.686$, $P = 0.035$. Post hoc comparisons were carried out with a Bonferroni-adjusted Wilcoxon signed-rank test, rendering a significance level at $P < 0.017$ (.05/3). For 10VQ, scores from the VLT were significantly higher than scores from the work condition in all vocal subgroups. For the leisure condition, this held true for FD, HLC, and C. The results of post hoc tests from 10VQ are presented in Table 3.

For VAS, scores from the VLT were significantly higher than scores from the work and leisure conditions only in FD (VLT to work: $Z = -3.506$, $P < 0.01$, VLT to leisure: $Z = -3.823$, $P < 0.01$). The VAS scores from VLT were significantly higher than leisure for HLC ($Z = -2.395$, $P = 0.017$) and HLNC ($Z = -2.380$, $P = 0.017$). None of the vocal subgroups displayed differences between work and leisure conditions for VAS scores. For FD, the median stress scores from the work condition (Med = 2) were significantly higher than the in the VLT (Med = 1), $Z = 2.244$, $P = 0.025$, and in the leisure condition (Med = 1.4), $Z = 3.108$. $P = 0.002$. For HLC, the median stress scores from the work condition (Med = 1.2) were significantly higher than the in the VLT (Med = 1), $Z = 2.244$, $P = 0.025$, and in the leisure condition (Med = .6), $Z = 2.684$, $P = 0.007$. For HLNC, the stress scores from the work condition (Med = 1) were significantly higher than in the leisure condition (Med = .5), $Z = 2.325$, $P = 0.02$.

Consistency of voice parameters in voice health questionnaire

A Spearman rank order correlation was performed to explore how well overall 10VQ ratings matched overall VAS scores for all vocal subgroups together. There was a strong positive correlation between the rank of 10 voice questions and VAS scores ($\rho = .84$, $P < 0.01$), as shown in Figure 6. When broken down
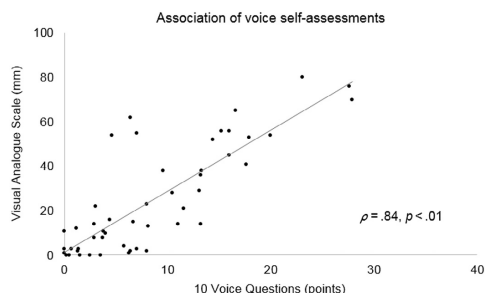
FIGURE 6. A Spearman rank correlation for all vocal subgroups showing a large positive, significant correlation between scores from 10 voice questions (10VQ, maximum of 40 points) and visual analogue scale (VAS) (100-mm VAS).

into subgroups, the correlations were moderate (FD: $\rho = .62$, HLC: $\rho = .62$, HLNC: $\rho = -.43$, C: $\rho = .68$), all nonsignificant.

Research questions revisited

(1) How does vocal behavior in women with diagnosed functional dysphonia (FD) differ from others who also experience voice problems, but who do not seek therapy? (How do these two groups compare with control groups?)
There were no significant group differences between FD and the other three groups regarding vocal behavior across conditions. The other group with voice problems (HLC) produced notably higher fundamental frequencies during VLT and work than all other groups. They also showed a tendency to phonate more than others during work, but not during leisure or VLT. This may explain why HLC experience voice problems but, contrary to FD, do not seek help for them.

(2) How does self-assessed vocal health in women with diagnosed functional dysphonia differ from others who also experience voice problems, but who do not seek therapy? (How do these two groups compare with control groups?)
FD self-assessed voice problems across all conditions, whereas HLC did not during leisure. HLNC and C did not record any voice problems other than in the extreme VLT condition. FD self-assess higher levels of stress than other groups across conditions.

DISCUSSION

The purpose of this study was to observe vocal behavior and self-assessed vocal health across three conditions in a population with varying functional voice problems. We wanted to know what sets women with functional voice problems who do seek voice therapy (FD) apart from those who do not seek therapy (HLC), and how these relate to voice healthy individuals with (HLNC) and without (C) vocal load. It seems the answer to this question is twofold: (1) HLC speak more and at a higher fundamental frequency in vocally demanding contexts compared with others; and (2) HLC

tend to mimic self-assessment of FD regarding incidence of general voice problems (recorded with VAS) in vocally demanding contexts, but not when asked specific voice questions. HLC react to high vocal loading, prompted prolonged loud phonation, through increased fundamental frequency. This coping mechanism sets HLC apart from, eg, the clinical population with Bogart-Bacall syndrome, who are located at the opposite end of the spectrum, using speaking fundamental frequencies at the very bottom of their frequency range. For comparison, it would be of interest to use our method to examine professional voice users who have very high requirements on their speaking voices and who more often than others are afflicted by Bogart-Bacall syndrome.[32] HLC also react to high vocal loading through an increased general sensation of discomfort, which does not necessarily cause specific, mechanically measurable parts of the vocal function to strain, but something vaguer, something that is more difficult to track with our clinical instruments. Consistency between the two self-assessment tools 10VQ and VAS was generally high, and the pattern could be seen throughout conditions and groups. The only exception was HLC's high VAS assessments in the VLT, which even exceeded FD's assessment.

An important difference emerged between the two groups with voice problems: HLC was the only vocal subgroup to self-assess the work condition with significantly higher VAS scores than the leisure condition, indicating that there might be such a thing as a "true" clinical population with occupational voice problems—probably due to high vocal strain during working hours. FD, on the other hand, was the only group to report high prevalence of voice problems (10VQ and VAS) during leisure. In fact, their report did not differ a lot across conditions. In accordance with Lyberg-Åhlander et al,[13] there is not enough time during leisure for vocal recovery to occur for the FD group, which is vital for a healthy voice function.[33] HLC scored high on the general VAS assessment, but not accordingly high on more specific 10VQ assessment. A possible explanation is an inability in this group to properly assign their voice problems to separate voice symptoms. There could also be an effect of bias: HLC were not recruited as *patients* with functional dysphonia, and might have been careful to score too high when it comes to voice symptoms, but the general VAS was perhaps easier or more adequate for them to use. The functional dysphonia patients (FD) in the study, on the other hand, showed high self-assessed voice health scores across all conditions, which probably explains why they seek help in the voice clinic. FD also report higher prevalence of general stress across conditions. High stress levels may be associated with the fact that they are constantly reminded (by the voice health questionnaire) of their voice problems, which do not diminish over time.

In a previous study, we developed a method of loading the vocal function of voice healthy participants.[24] We considered the difficulties of simulating true vocal loading by loud speech during a length of time set by the participants themselves and suggested changes to the method which were applied in the current study. Onset and regression of vocal loading effects were successfully tracked with the former study's voice health questionnaire. In order to ensure comparable compliance to the method during the long-time voice measurement, we are currently

investigating recovery from vocal loading in a parallel study. In the current study, all subjects filled out the voice health questionnaire according to a schedule: once every morning, midday, afternoon, and evening. An unexpected finding in the former study was the pronounced differences in endurance time in the VLT. In order to minimize this effect, the oral instructions were altered. Unlike the subjects in the former study, the current subjects were not informed that the time limit in the VLT was 30 minutes. They were also told not to terminate the VLT until they perceived a *distinct* discomfort (Swedish: "tydligt obehag") from the throat. The VLT in these studies generate an augmentation of any vocal load in reality, subjecting the participants to a very heavy vocal load consisting of speech at high sound pressure levels for as long as they could muster. Relative phonation time averaged over 1 day is expected to vary.[21] However, the results from the VLT show high time ratios of phonation compared with the other conditions, as this was the only condition showing prolonged high phonatory intensity. The mean relative phonation time for all subgroups during the VLT was approximately 65%. We can argue for 60%–70% as reference value for maximum time for phonation during loud, prolonged speech based on loud reading in Swedish as measured with *VoxLog*. This implies that 100% speech during 30 minutes equals 60%–70% phonation time. This gives a smaller factor (1.5) than proposed by Titze et al (factor ca 2),[34] which is not surprising owing to language-specific phonological typology.

The current study includes people with confirmed functional voice problems (diagnosed and undiagnosed), which has been called for in previous research.[15] Setting the risk of selection bias aside, this setup runs the risk of subjects being pressured to certain expectations, depending on what criteria they were recruited on. For example, one FD subject who had undergone a full round of voice therapy told the test leader explicitly that she could no longer report severe voice problems, because she and her treating speech language pathologist had agreed that her voice now was greatly improved. However, the patient was not sure which aspect of the vocal function had improved, thus making a case for placebo effect in voice health care. The FDs who had already undergone voice therapy had other words to describe their voice problems, which may have had an impact on their self-assessments. This points out the importance of exact and proper phrasing of oral instructions and discussions of vocal function with patients. The subject base may also have affected the results due to menopausal voice change. HLC held the highest percentage (60%) of women above the age of 50 (FD: 20%, HLNC: 40%, C: 20%). This fact might be one of the causes of their vocal function being particularly sensitive to high vocal use during working hours, as voice disorders are sensitive to age.[35] HLC's constant high fundamental frequency could be a coping mechanism, trying to compensate for a lack of sex hormone, which theoretically would lead their fundamental frequency to decrease.[36] Such coping could be productive to address in voice therapy for women in vocally demanding occupations.

Another important difference in the current study is the improved sturdiness of the neck collars now supplied with the *VoxLog* by Sonvox, rendering more stable measurements by both accelerometer and microphone. However, compared with other vocal dosimeters, *VoxLog* has high mean errors in both fundamental frequency and sound pressure level measurements.[37]

The relationship between personality traits and psychological health in relation to vocal function needs to be closely examined, as it may affect vocal behavior.[38] The clinical diagnoses of the patients included in the current study did not discriminate between muscle tension voice disorders and psychogenic voice disorders *ad modum* Baker et al.[12] On recruitment, all subjects were asked about current psychological health. None of the FD group declared any underlying psychological health markers that would indicate PVD, other than generally high stress levels (as was subsequently confirmed in the current study). In all other groups, however, depression and/or fatigue caused by burnout was noted. One explanation of HLC's nonspecific recognition of vocal loading though VAS, but not 10VQ, could be in line with Baker et al.[39] Baker et al found lower levels of emotional awareness and higher levels of interest for external processes, compared with internal processes, in women with pure muscle tension voice disorders who were compared with patients with more severe voice problems and voice healthy controls.

The 10VQ was chosen from the voice handicap index[22] with the addition of a question rating general stress levels, as high stress levels could cause fundamental frequency to increase.[40] It was important to keep the recurring questions brief and sensitive to changes in vocal condition. The 8VQ was not as sensitive to change and seems to be appropriate to use in diagnostics rather than when tracking change of the vocal function over time within an individual. There was no clear correlation between stress levels and eg constant high fundamental frequency in HLC, which could mean that the question of stress either needs to be investigated in a different manner or the connection between a certain vocal behavior and stress is not as clear-cut. From the perspective of the International Classification of Functioning, Disability and Health,[41] it is important not only to look at the individual when assessing impact of any functional impairment, but also to see the bigger picture surrounding the individual, which sheds light on how much needs to be incorporated into an assessment of vocal function (health condition: disorder or disease, body functions and structure, activity, participation, environmental, and personal factors). Self-assessments of vocal quality of life through voice activity questionnaires may be a good way of individualizing this process.[42] Customizing self-assessment to reflect vocal function in different contexts could be a helpful way of further charting functional voice problems.

As has been shown in previous studies,[13,14] vocal loading seems to be related, not only to high levels of sound pressure during prolonged speech, but also to high fundamental frequency and high levels of relative phonation time. This will cause a strain to the laryngeal mechanism, but also perhaps lead to higher levels of general stress. Findings from the current study and previous field studies of female voices call for a general revisit of the standard fundamental frequency of 180–220 Hz.[43]

CONCLUSIONS

- Vocal loading is not only dependent on prolonged phonation time at high intensity levels, it also seems to be

reliant on prolonged phonation time at high fundamental frequencies.

- Women with high everyday vocal load who experience voice problems (HLC) reported strain-induced voice problems during a VLT and during work. This sets them apart from patients with functional dysphonia (FD) who exhibited voice problems in all conditions, even during leisure. This may explain why people with voice problems associated with their work environment do not seek voice therapy.

Acknowledgments

REFERENCES

1. Vilkman E. Occupational safety and health aspects of voice and speech professions. *Folia Phoniatr Logop*. 2004;56:220–253.
2. Södersten M, Ternström S, Bohman M. Loud speech in realistic environmental noise: phonetogram data, perceptual voice quality, subjective ratings, and gender differences in healthy speakers. *J Voice*. 2005;19:29–46.
3. Lyberg-Åhlander V, Rydell R, Löfqvist A. Speaker's comfort in teaching environments: voice problems in Swedish teaching staff. *J Voice*. 2011;25:430–440.
4. Fritzell B. Voice disorders and occupations. *Logoped Phoniatr Vocol*. 1996;21:7–12.
5. Ohlsson AC, Andersson EM, Södersten M, et al. Voice disorders in teacher students–a prospective study and a randomized controlled trial. *J Voice*. 2015; 26:629–634.
6. Roy N, Merrill RM, Thibeault S, et al. Prevalence of voice disorders in teachers and the general population. *J Speech Lang Hear Res*. 2004;47:281–293.
7. Laukkanen AM, Ilomäki I, Leppänen K, et al. Acoustic measures and self-reports of vocal fatigue by female teachers. *J Voice*. 2008;22:283–289.
8. Rantala L, Vilkman E, Bloigu R. Voice changes during work: subjective complaints and objective measurements for female primary and secondary schoolteachers. *J Voice*. 2002;16:344–355.
9. Jónsdottir V, Laukkanen AM, Vilkman E. Changes in teachers' speech during a working day with and without electric sound amplification. *Folia Phoniatr Logop*. 2002;54:282–287.
10. Lindström F, Persson Waye K, Södersten M, et al. Observations of the relationship between noise exposure and preschool teacher voice usage in day-care center environments. *J Voice*. 2011;25:166–172.
11. Chen SH, Chiang S, Chung Y, et al. Risk factors and effects of voice problems for teachers. *J Voice*. 2009;24:183–190.
12. Baker J, Ben-Tovim DI, Butcher A, et al. Development of a modified diagnostic classification system for voice disorders with inter-rater reliability study. *Logoped Phoniatr Vocol*. 2007;32:99–112.
13. Lyberg-Åhlander V, Whitling S, Rydell R, et al. Teachers' voice use in teaching environments: a field study using ambulatory phonation monitor. *J Voice*. 2014;28:841, e5–15.
14. Szabo Portela A, Hammarberg B, Södersten M. Speaking fundamental frequency and phonation time during work and leisure time in vocally healthy preschool teachers measured with a voice accumulator. *Folia Phoniatr Logop*. 2014;65:84–90.
15. Behlau M, Madazio G, Oliveira G. Functional dysphonia: strategies to improve patient outcomes. *Patient Relat Outcome Meas*. 2015;6:243–253.
16. Roy N. Functional dysphonia. *Curr Opin Otolaryngol Head Neck Surg*. 2003;11:144–148.
17. Hunter EJ, Titze IR. Variations in intensity, fundamental frequency, and voicing for teachers in occupational versus nonoccupational settings. *J Speech Lang Hear Res*. 2010;53:862–875.
18. Pedrosa V, Pontes A, Pontes P, et al. The effectiveness of the comprehensive voice rehabilitation program compared with the vocal function exercises method in behavioral dysphonia: a randomized clinical trial. *J Voice*. 2015;30:377, e11, 9.
19. Van Stan JH, Gustafsson J, Schalling E, et al. Direct comparison of three commercially available devices for voice ambulatory monitoring and biofeedback. *Perspect Voice Voice Disord*. 2014;24:80–86.
20. Solomon NP. Vocal fatigue and its relation to vocal hyperfunction. *Int J Speech Lang Pathol*. 2008;10:254–266.
21. Södersten M, Salomão GL, McAllister A, et al. Natural voice use in patients with voice disorders and vocally healthy speakers based on 2 days voice accumulator information from a database. *J Voice*. 2015;29:646, e1–9.
22. Jacobson BH, Johnson A, Grywalski C, et al. The voice handicap index (VHI) development and validation. *Am J Speech Lang Pathol*. 1997;6:66–70.
23. Lance CE, Butts MM, Michels LC. The sources of four commonly reported cutoff criteria: what did they really say? *Organ Res Methods*. 2006;9:202–220.
24. Whitling S, Rydell R, Lyberg-Åhlander V. Design of a clinical vocal loading test with long-time measurement of voice. *J Voice*. 2015;29:261, e13–27.
25. Shapiro SS, Wilk MB. *An Analysis of Variance Test for Normality (Complete Samples)*. London: Biometrika Office, University College; 1965:591.
26. Razali NM, Wah YB. Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. *J Stat Model Anal*. 2011;2:21–33.
27. Cramer D. *Fundamental Statistics for Social Research: Step-by-Step Calculations and Computer Techniques Using SPSS for Windows*. London: Routledge; 1998.
28. Cramer D, Howitt D. *The Sage Dictionary of Statistics*. London: Sage; 2004.
29. Doane DP, Seward LE. Measuring skewness: a forgotten statistic? *J Stat Educ*. 2011;19:1–18.
30. Cohen J. Statistical power analysis. *Curr Dir Psychol Sci*. 1992;1:98–101.
31. Lane H, Tranel B. The Lombard sign and the role of hearing in speech. *J Speech Lang Hear Res*. 1971;14:677–709.
32. Koufman JA, Blalock PD. Vocal fatigue and dysphonia in the professional voice user: Bogart-Bacall syndrome. *Laryngoscope*. 1988;98:493–498.
33. Dejonckere PH, Bradley P, Clemente P, et al. A basic protocol for functional assessment of voice pathology, especially for investigating the efficacy of (phonosurgical) treatments and evaluating new assessment techniques: guideline elaborated by the Committee on Phoniatrics of the European Laryngological Society (ELS). *Eur Arch Otorhinolaryngol*. 2001;258:77–82.
34. Titze IR, Hunter EJ, Švec JG. Voicing and silence periods in daily and weekly vocalizations of teachers. *J Acoust Soc Am*. 2007;121:469–478.
35. Thibeault SL, Merrill RM, Roy N, et al. Occupational risk factors associated with voice disorders among teachers. *Ann Epidemiol*. 2004;14:786–792.
36. Lindholm P, Vilkman E, Raudaskoski T, et al. The effect of postmenopause and postmenopausal HRT on measured voice values and vocal symptoms. *Maturitas*. 1997;28:47–53.
37. Bottalico P, Ipsaro Passione I, Hunter EJ. Vocal dosimeter devices and their uncertainty. *J Acoust Soc Am*. 2016;139:2018.
38. Roy N, Bless DM, Heisey D. Personality and voice disorders: a superfactor trait analysis. *J Speech Lang Hear Res*. 2000;43:749–768.
39. Baker J, Oates J, Leeson E, et al. Patterns of emotional expression and responses to health and illness in women with functional voice disorders (MTVD) and a comparison group. *J Voice*. 2014;28:762–769.
40. Tse ACY, Wong AWK, Whitehill TL, et al. Analogy instruction and speech performance under psychological stress. *J Voice*. 2014;28:196–202.
41. WHO. *International Classification of Functioning Disability and Health: ICF*. Geneva, Switzerland: World Health Organization; 2001.
42. Hancock AB. An ICF perspective on voice-related quality of life of American transgender women. *J Voice*. 2016; In press.
43. Pegoraro Krook MI. Speaking fundamental frequency characteristics of normal Swedish subjects obtained by glottal frequency analysis. *Folia Phoniatr (Basel)*. 1988;40:82–90.