**Validation of a fast and easy screening method
for voice disorders**

Alma Zenger, 35660

Master's Thesis

Supervisors: Sofia Holmqvist-Jämsén and Daniel Fellman

Faculty of arts, psychology and theology

Logopedics

Åbo Akademi University

2019

Alma Zenger

ÅBO AKADEMI UNIVERSITY

FACULTY OF ARTS, PSYCHOLOGY AND THEOLOGY

Abstract for master's thesis

Subject: Logopedics
Author: Alma Zenger
Title of the work: Validation of a fast and easy screening method for voice disorders
Supervisors: Sofia Holmqvist-Jämsén and Daniel Fellman
Abstract: Vocal symptoms and voice disorders are a common health issue, especially among persons with vocally demanding occupations. Thus, vocal symptoms should be observed and detected in an early stage to avoid them from developing to voice disorders. A fast and easily-administered screening method developed for capturing early signs of a possible voice disorder could have the potential to prevent more severe voice-related issues. The aim of this thesis was to investigate the validity of an updated voice-related questionnaire, Screen11, which could be used as a potential screening tool for voice disorders. The investigation included analyses examining the internal consistency and convergent validity of Screen11. Additionally, the aim was to explore a cut-off score to distinguish individuals at risk of developing voice disorders from vocally healthy individuals. The study included 24 persons seeking help for their voice problems and 24 control persons. The help seekers were recruited from the *Phoniatric Outpatient Clinic of Turku University Hospital* and the control persons were matched with the help seekers according to age, gender and occupation. All participants received three different voice-related questionnaires: Screen11, *Voice Handicap Index (VHI)* and *Voice Activity and Participation Profile (VAPP)*. The results indicated that Screen11 possessed a good internal consistency. Moreover, it exhibited strong convergent validity by correlating well with VHI and VAPP. Furthermore, four or more frequent vocal symptoms occurring daily or weekly was shown to be the most optimal cut-off score for separating the help seekers from the control persons. The results indicate a successful validation of the screening method. Screen11 could be a useful addition to the toolbox of voice-related questionnaires that screens for possible voice disorders at an early stage.
Key words: Screen11, vocal symptom, Voice Activity and Participation Profile, voice disorder, Voice Handicap Index, voice screening
Date: 11.01.2019	Pages: 31 + 10

Alma Zenger

ÅBO AKADEMI

FAKULTETEN FÖR HUMANIORA, PSYKOLOGI OCH TEOLOGI

Sammanfattning av avhandling pro gradu

Ämne: Logopedi	
Författare: Alma Zenger	
Arbetets titel: Validering av ett snabbt och enkelt screeningformulär för röststörningar	
Handledare: Sofia Holmqvist-Jämsén och Daniel Fellman	

Sammanfattning:

Röstsymtom och röststörningar kan vara en stor belastning i vardagen och försämra livskvaliteten, speciellt hos individer med röstkrävande yrken. För att minska förekomsten av röststörningar borde röstsymtom upptäckas och identifieras i ett tidigt skede. Ett sätt att ingripa tidigt i de eventuella förestående röstsvårigheterna är genom screening av dem.

Syftet med den här pro gradu-avhandlingen var att undersöka ett nytt potentiellt screeningverktyg för röststörningar, Screen11. Metoden baserar sig på ett frågeformulär som består av elva symtomrelaterade påståenden som besvaras och poängsätts enligt de givna svaren. De specifika syftena för avhandlingen var att kontrollera formulärets interna konsistens och konvergenta validitet samt bestämma ett gränsvärde av upplevda röstsymtom som kunde tyda på en eventuell röststörning.

I studien deltog 24 personer som sökte hjälp för sina röstproblem och 24 röstfriska kontrollpersoner. De hjälpsökande personerna rekryterades via *Åbo Universitets Centralsjukhus foniatriska poliklinik* och de röstfriska personerna matchades enligt de hjälpsökandes ålder, kön och yrke. Alla 48 deltagare fick besvara tre olika röstrelaterade frågeformulär varav ett var Screen11 och de två andra var *Voice Handicap Index (VHI)* och *Voice Activity and Participation Profile (VAPP)*.

Resultaten visar att Screen11 har en stark intern konsistens och att det korrelerar starkt med de två andra röstfrågeformulären VHI och VAPP. Gränsvärdet i Screen11 som urskiljer de hjälpsökande personerna från röstfriska personer var fyra eller flera upplevda röstsymtom varje vecka eller dagligen. Studiens resultat påvisar en lyckad validering av screeningformuläret Screen11.

Nyckelord: Screen11, röstscreening, röststörning, röstsymtom, Voice Activity and Participation Profile, Voice Handicap Index	
Datum: 11.01.2019	Sidantal: 31 + 10

Alma Zenger

Contents

1 Introduction

Vocal symptoms and voice disorders have a large impact on people's daily activities and quality of life (Verdolini & Ramig, 2001), which often results in sick-leaves and thus increased financial costs for the community (Van Houtte, Claeys, Wuyts, & Van Lierde, 2010). Hence, voice disorders should be detected as early as possible, particularly for those populations that have an increased risk for developing voice disorders (Vilkman, 2004). One way of detecting voice disorders at an early stage is screening for voice disorders and vocal symptoms (National Screening Unit, 2018), as in a study by Simberg, Sala, Laine and Rönnemaa (2001) where a fast and easy voice-screening method administered among teacher students is described. At present, some validated screening questionnaires for voice disorders are available. Of these, perhaps the most frequently one used is the Voice Handicap Index (VHI; Jacobson et al., 1997), capturing the subjective perception of disability due to a voice disorder. Another commonly used questionnaire is the Voice Activity and Participation Profile (VAPP; Ma & Yiu, 2001), which gathers information about limitations of activity and participation of persons with voice disorders.

Although both of the aforementioned questionnaires have shown to be reliable and valid instruments for measuring the influence of vocal symptoms on individuals' quality of life (Jacobson et al., 1997; Ma & Yiu, 2001; Seifpanahi, Jalaie, Nikoo, & Sobhani-Rad, 2015; Sukanen et al., 2007), their assessment can be quite time-consuming due to their extensive battery of questions. Moreover, the aim of these questionnaires is not only to screen for the physical vocal symptoms and possible voice disorders, but to capture emotionally and socially related issues that the voice disorder might cause. Thus, a fast and easily-administered screening questionnaire focusing solely on the symptomatic issues related to voice disorders is lacking, deserving further scrutiny. The purpose of this thesis was therefore to examine whether a voice questionnaire encompassing only eleven statements about vocal symptoms could be an adequate screening method to identify possible voice disorders.

1.1 Vocal symptoms and voice disorders

In previous literature, most researchers have made a clear distinction between the terms *vocal symptom* and *voice disorder*. A vocal symptom is referring to a temporary symptom in the throat, voice or the vocal mechanism, whereas a voice disorder exists when the

individual is concerned about her or his voice being deviant or not matching with the demands of her or his daily voice use (American Speech-Language-Hearing Association [ASHA], 2018). This distinction and these definitions have been used in this thesis.

Studies investigating the prevalence of vocal symptoms and voice disorders have shown highly varying results, mainly stemming from different definitions concerning what a voice disorder constitutes and how they are operationalized. For example, according to Vilkman (2004), a voice disorder is present when an individual's voice fails to meet the occupational criteria and demands, while Roy et al. (2001) define a voice disorder any time the voice does not work, perform, or sound as it normally should. Verdolini and Ramig (2001), on the other hand, define a voice disorder as a condition of sufficient concern for the patient to seek treatment in an otolaryngologic clinic, while Aronson (1990) defines a voice disorder as a condition when a person's voice is deviant on pitch and range compared with persons with the same age, gender, cultural background and geographic location.

An operationalization of voice disorders has been put forth by Simberg, Laine, Sala and Rönnemaa (2000); here, a voice disorder is identified when individuals report two or more frequently occurring voice symptoms per week on a self-reporting questionnaire or if the voice sounds remarkably deviant. The cut-off score of two or more frequently reported vocal symptoms is based on the results from a subsequent study by Simberg et al. (2001) showing that those reporting two or more symptoms weekly, most often had organic changes in the larynx. Visible organic changes in the laryngeal area might be a sign of an organic voice disorder (ASHA, 2018). Using this method for defining a voice disorder, a large number of studies have investigated the prevalence of voice disorders in different occupational groups such as teachers (Ohlsson, Andersson, Södersten, Simberg, & Barregård, 2012), nurses (Pekkarinen, Himberg, & Pentti, 1992), and priests (Hagelberg & Simberg, 2015). This self-reporting method was later developed and coined as the Screen6 questionnaire (Ohlsson et al., 2012).

1.2 Factors affecting the prevalence of vocal symptoms and voice disorders

Previous studies have shown highly varying results regarding the prevalence of voice disorders, with a prevalence rate spanning from 3% to 30% in different populations. In a recent Swedish study of voice disorder in the general population ($n > 100\ 000$), the prevalence rate was estimated to 16.9% (Lyberg-Åhlander, Rydell, Fredlund, Magnusson,

& Wilén, in press). Studies from the US show a prevalence between 3% and 9% of the US population (Ramig & Verdolini, 1998; Roy, Merrill, Gray, & Smith, 2005), suggesting that the prevalence rates are slightly varying across countries. Besides the inconsistent definitions as to what constitute voice disorders, the varying prevalence rates of voice disorders might also be related to the highly varying characteristics of the samples recruited in prior studies. The prevalence of vocal symptoms and voice disorders has typically been studied in specific occupational groups, with teachers being the most commonly one studied. The prevalence of voice disorders among teachers is spanning from 11% to 57.7%, with the prevalence rates varying depending on whether the symptoms have been defined as currently occurring (Roy et al., 2004; Sapir, Keidar, & Mathers; Schmidt, 1993; Smith, Lemke, Taylor, Kirchner, & Hoffman, 1998) or as vocal symptoms occurring during lifetime (Roy et al., 2004). Vocal symptoms or voice disorders occur among 13–17% of students (Linnasalo, 1990; Simberg, Sala, & Rönnemaa, 2004), whereas vocal students (e.g. becoming singers) and acting students suffer from voice disorders slightly more with a prevalence of 27–61% (Sapir et al., 1993; Winkworth & McCabe, 2001; Timmermans et al., 2002). A survey of voice problems among Catholic priests ($n = 340$) showed a prevalence of 15.9% (Hocevar-Boltezar, 2009).

Several risk factors have shown to increase the exposure for developing voice disorders, including both individual and environmental factors. One circumstance that consistently has shown to increase the risk for having voice disorders is vocally demanding occupations (Vilkman, 2000; Williams, 2003). Such occupations are for instance teachers, day care personnel, priests, soccer coaches, singers and actors. Each of these occupations are characterized by an extensive and heavy vocal load (Vilkman, 2004). Vocal load is a term that describes the physical effort on the vocal organ in prolonged periods of speaking (Vilkman, 2004; Whitling, Rydell, & Lyberg-Åhlander, 2015). Early stages of vocal loading can be experienced as a warm-up feeling on the vocal folds. If the vocal organ becomes too exposed to constant heavy vocal loading, it might lead to vocal fatigue, which refers to a negative sensation (Vilkman, 2004). Work-related ergonomic factors, such as working posture (Morrison & Rammage, 1993) or usage of a voice amplifier in suitable situations (Buekers, 2001; Vilkman, 2004), have an impact on the individual's vocal load. A population of active and working adults has a higher prevalence of vocal symptoms and voice disorders than the general population (Van Houtte et al., 2010).

Besides heavy vocal load, poor working conditions, including poor inside air quality and poor acoustic environment, have also shown to contribute to voice disorders. High

background noise and long working hours without vocal rest have shown to have a negative impact on the voice (Morton & Watson, 1998; Rantala, Hakala, Holmqvist, & Sala, 2013). In combination with individual factors such as general health (Stemple, Stanley, & Lee, 1995), psychosocial situation (Kooijman et al., 2006), personality (Dietrich, Abbott, Gartner-Schmidt, & Rosen, 2008), stress (Holmqvist, Santtila, Lindström, Sala, & Simberg, 2013) and heredity (Simberg et al., 2009), the negative environmental working conditions lead to higher prevalence of disordered voice in vocally demanding occupations (Cutiva, Vogel, & Burdorf, 2013).

Another consideration for prospective voice disorders is related to gender-specific differences, with females typically reporting a higher prevalence of vocal symptoms (Martins et al. 2016; Roy et al., 2005) and certain voice disorders (Cohen, Kim, Roy, Asche, & Courey, 2012) than males do (Fritzell, 1996; Remacle, Petitfils, Finck, & Morsomme, 2017). The reason for this might partly be explained by differences in laryngeal structures between females and males. Females tend to have shorter vocal folds and a smaller vocal tract than males (Titze, 1994). The fundamental frequency of the vocal folds is therefore higher in women than in men, resulting in a greater load and affection for the vocal folds (Butler, Hammond, & Gray, 2001; Hunter, Tanner, & Smith, 2011). The potentially lower levels of hyaluronic acids on the superficial area of the vocal folds in females might also influence this parameter, potentially resulting to less protection from vocal overuse which could partly explain why females more often suffers from phonotrauma than males (Butler et al., 2001). There is, however, a lack of consensus regarding the association between gender differences and levels of hyaluronic acids (Butler et al., 2001; Hunter et al., 2011). The higher rates of voice disorders among women might also be related to occupational factors, as women tend to be overrepresented in occupations with heavy vocal loading, such as teaching (Vilkman, 2004).

Furthermore, one factor that has an impact on vocal symptoms is the age parameter. Various age groups are vulnerable for different vocal symptoms of diverse etiologies: vocal fold nodules and cysts are dominant among children, adults suffer mostly from functional voice disorders and reflux, while elderly people are diagnosed with presbyphonia and Reinke's edema. This distribution might depend on anatomical changes that take place due to ageing (Martins et al., 2016).

1.3 Screening methods for vocal symptoms

The term *screening* refers to a procedure where the aim is to detect potential diseases, physical and mental disorders or other malfunctions in an early stage among individuals who might have an increased risk of a particular condition (National Screening Unit, 2018). The procedure may be initiated by concerns from the individual, relatives or other relationships (ASHA, 2018). Ideally, the screening procedure should be fast, simple and inexpensive to perform. In the context of voice disorders, the screening procedure is typically conducted using self-reporting questionnaires (National Screening Unit, 2018). If the initial screening results indicate deviations from a normal voice use, further examinations are needed (ASHA, 2018).

1.3.1 Validated voice screening questionnaires

The VHI (Jacobson et al., 1997) is one of the most popular self-reporting voice questionnaires used in clinics (Seifpanahi et al., 2015). The VHI is considered as a valid and reliable diagnostic instrument by the Agency of Healthcare Research and Quality (2002), and it has been translated and adapted into several languages during the past decades (Amir et al., 2006; Guimaraes & Abberton, 2004; Hsiung, Pai, & Wang, 2002; Ohlsson & Dotevall, 2009; Siupsinskiene, 2002; Verdonck-de Leeuw et al. 2008; Woisard, Bodin, & Puech, 2004). The VHI consists of 30 voice-related questions divided into three subsections: functional, emotional and physical symptoms. The functional part (10 items) measures how well the voice works in daily activities. The physiological part (10 items) measures symptomatic behavior of the voice, i.e., vocal symptoms. The third section (10 items) measures emotional reactions and attitude according to the voice. The answers are given on a scale from 0 = *never* to 4 = *always*. The higher the score, the more severe perception of disability caused by the voice symptoms, with the highest possible score of 120. The VHI measures the subjective perception of disability due to a voice disorder, i.e., voice-related quality of life (Seifpanahi et al., 2015). The Finnish version of the VHI is translated and confirmed by Alaluusua and Johansson (2003), and this version has been used in the data collection for this thesis.

The VAPP (Ma & Yiu, 2001) is a self-assessment questionnaire that probes for information about limitations of activity and participation of persons with voice disorders. Developed by Ma and Yiu (2001), the VAPP questionnaire was shown to be a valid and

reliable tool to measure the voice-related quality of life in Chinese patients. The VAPP consists of 28 questions, which are answered on visual analog scales. The questions are divided into five subsections as follow: self-perceived severity of voice problem (1 item), effect on job (4 items), effect on daily communication (12 items), effect on social communication (4 items) and effect on emotion (7 items). The total score from the VAPP varies from 0 to 280, where high scores refer to the severity of the voice symptoms and their impact on individuals' activity and participation. A Finnish version of the VAPP has been validated by Sukanen et al. (2007) and this version was used in the present study.

Additional self-evaluative questionnaires screening for voice problems have been developed, such as The Voice Symptom Scale (VoiSS; Deary, Wilson, Carding, & MacKenzie, 2003), The Voice Related Quality of Life Measure (V-RQOL; Murry, Medrado, Hogikyan, & Aviv, 2004), The Vocal Performance Questionnaire (VPQ; Deary, Webb, MacKenzie, Wilson, & Carding, 2004). These questionnaires typically capture the type and severity of possible voice disorders and measure their effect on the quality of life. None of these questionnaires are currently available in Finnish.

1.3.2 Development of Screen11

The development and application of the Screen11 questionnaire is described in more detail in this section. The previous version of Screen11, originally coined as Screen6 (Ohlsson et al., 2012), has been conducted mostly on teachers or teacher students as individuals from these occupations typically seek help for voice problems more often compared to individuals working in other professions (Pekkarinen et al., 1992; Simberg et al., 2000; Simberg et al., 2001). Starting with a pioneering study by Simberg et al. (2000), the prevalence of voice disorders was examined among teacher students ($N = 226$) using an extensive research battery of questions related to voice use. Moreover, the participants' voices were assessed perceptually by a speech therapist, and clinically by a laryngologist. The results showed that 20% of the teachers had a voice disorder (i.e., reported two or more frequently occurring voice symptoms), and most of these voice disorders (19%) were of organic type. To further examine the validity of a reduced version of a possible voice screening questionnaire, Simberg et al. (2001) investigated how the questions were related to perceptual voice quality and laryngeal status as measured with indirect laryngoscopy. In other words, the participants' voices were assessed perceptually by a nurse and a speech and language pathologist besides the self-assessment voice questionnaire. Participants who

were evaluated to have an abnormal voice quality according to the perceptual assessments and those who reported two or more vocal symptoms were referred to further phoniatric examinations. The results showed that the participants who reported two or more vocal symptoms occurring daily or weekly, often also had visible changes on their vocal folds. The questionnaire consisted of eight voice-related statements: *1) Morning hoarseness 2) The voice gets strained or tired 3) The voice gets low or hoarse while talking 4) Voice-breaks while talking 5) To loosen voice for at least a couple of minutes while talking 6) Difficulty in being heard 7) Throat clearing or coughing while talking 8) Feel of pain, tension or a lump in throat.* The questionnaire gathered information from the past year and the options for occurred frequency were *1) Every day or most days 2) Weekly or most weeks 3) Monthly or most months 4) Less often than above 5) Only periodic symptoms 6) No symptoms*.

Based on the results from Simberg et al. (2000) and Simberg et al. (2001), the development of a voice screening questionnaire to discover and prevent voice disorders among students in an earlier stage was continued (Simberg et al., 2004). In line with the previous studies, a voice disorder was considered present if the participant reported to experience vocal symptoms on a weekly basis or more often. The results showed that the most frequently reported symptoms were morning hoarseness (28%) and throat clearing (20%) while none of the participants reported difficulty in being heard or loss of voice. The results and the realization of this screening study appeared to be effective, but the authors suggested that the questionnaire could be compressed as none of the subjects reported difficulties in being heard or loss of voice.

Using the same vocal symptoms as employed in Simberg et al. (2001), Ohlsson et al. (2012) investigated the prevalence of voice disorders among Swedish first-year teacher students. It turned out that an even more shortened version including only six of the questions exhibited the most adequate screening approach for a voice disorder. These six symptoms, coined as Screen6, were; *Does your voice become strained or tired? Does your voice become low or hoarse? Does your voice break? Do you have difficulties in being heard? Do you need to clear your throat or to cough? Do you have a sensation of pain or lump in the throat?* The results showed that 17% ($n = 208$) of the teacher students ($N = 1250$) reported two or more frequently occurring (daily or weekly) vocal symptoms in Screen6. Moreover, two or more frequently occurring voice symptoms were significantly associated with higher scores in the VHI. Following this, the Screen6 has been extensively used to investigate the prevalence of voice disorders in the general population (Holmqvist

et al., 2013) as well as in several occupational groups (e.g. Fellman & Simberg, 2017; Hagelberg & Simberg, 2015), and it is commonly completed together with other questionnaires concerning quality of life.

Although Screen6 has been widely used, especially in combination with other questionnaires, it has some shortcomings that should be improved and confirmed. Therefore, there is a need to evaluate its adequacy as a screening tool for voice disorders. For example, some of the items in Screen6 have double-barreled statements. Such statements are problematic as they have the potential to confuse participants, and in worst case, it might lead to response avoidance due to controversial inquires (Hellweg, Pfau, & Brydon, 1992). Hence, these items should be divided into single statements to achieve more specific results. Another shortcoming of the Screen6 is that it lacks a careful investigation of the internal consistency. This examination is important, as it measures whether the items intended to capture the prevalence of voice disorders actually produce similar scores. Moreover, it has been shown that the sensitivity of the Screen6 is somewhat weak: in a study by Simberg et al. (2001), it was shown that only nine of 19 teacher students who had abnormal voice quality were having a voice disorder according to the Screen6 criteria (i.e., two or more frequently occurring vocal symptoms). This raises the issue whether this cut-off score is the most optimal one for discriminating between those with a voice disorder and those without a voice disorder.

1.4 The aim of the thesis

The aim for this thesis was to investigate if a voice questionnaire, coined as Screen11, consisting of only eleven statements could be a viable screening method to detect possible voice disorders. Besides some slight statement modifications, Screen11 (occasionally titled Sceen6) relies heavily on the voice symptom battery used in several previous studies investigating the prevalence of voice disorders (e.g. Simberg et al., 2000; Simberg et al., 2001; Simberg et al., 2004; Ohlsson et al., 2012; Ohlsson et al., 2016).

The first specific research question for this thesis was: *Can Screen11 be used as a viable screening method to detect possible voice disorders?* The first research question was followed by two subquestions: *How well do the eleven items in Screen11 correlate with each other (internal consistency)?* and *How well does Screen11 correlate with VAPP, VHI and VHI's subsections (convergent validity)?* The second main research question was: *Which is the best cut-off score in Screen11 for differentiating the participants who seek*

help for voice disorders from those who do not? In case of a successful validation, Screen11 could be implemented as a screening method for voice disorders especially for persons who have professions that are demanding for their voice.

2 Method

This study was conducted in co-operation with the *Phoniatric Outpatient Clinic* of *Turku University Hospital* (TYKS). The study was approved by the *Clinical Research Centre* in Turku and *The Board for Reasearch Ethics* at *Åbo Akademi University.*

2.1 Participants

In October 2017, the recruitment of participants was initiated. The recruitment procedure was arranged with the permission of Turku University Hospital: clients who were attending their first phoniatric examination in the near future received a request by mail to participate in the study. In the request, they were asked to fill in the following questionnaire forms: Screen11 (Appendix A), the VHI (Appendix B) and the VAPP (Appendix C). In addition, the envelope contained information about the study, two copies of participation consent and one stamped return envelope. Practical arrangements of the data collection will be described subsequently in this chapter.

During the period from October 2017 to February 2018, a number of 54 envelopes were sent to the prospective participants, 29 of whom responded. One of the criteria for participation was to respond to all questionnaires and post them in a letter before arriving to the hospital. Exclusion criteria comprised clients under 18 years, persons who had received voice therapy during the previous two years and persons with upper airway malignancies. Only Finnish-speaking persons were included in the study.

Matched pairs were used as control participants for those with prospective voice disorder and the controls were recruited via convenient sampling. Thus, each help seeking participant of whom the complete data were compiled was matched with a participant without a diagnosed voice disorder. The control participants were matched to help seeking participants according to age, gender and profession.

2.1.1   Occupation, gender and age distribution

Altogether 57 participants (help seeking participants $n = 29$; control participants $n = 28$) were recruited to take part in the present study. However, nine of these participants (help seeking participants $n = 5$, control participants $n = 4$) did not provide complete responses in the questionnaires (i.e., missing data in one or more items from either Screen11, the VHI or the VAPP) and were thus excluded from the analyses. Thus, the final sample of this study included voluntary participants seeking help for their voice symptoms ($n = 24$) and control participants ($n = 24$) without any voice problems. The age distribution ranged from 20 years to 78 years with a mean age of 47 years across groups.

The help seeking participants consisted mostly of females (females $n = 20$, males $n = 4$), which is in line with previous research showing that voice disorders are more frequent among females than males (Williams, 2003). The gender distribution in the control group differed slightly from the help seeking group, consisting of 21 females and 3 males.[1]

The participants in the study represented different occupational areas such as kindergarten teachers, class teachers, designers, practical nurses, students, salespersons and managing directors. Following Sukanen et al. (2007), the participants were categorized into six groups according to the vocal demands of the occupation. This categorization made it more convenient to find matching pairs according to the participant's profession. The categorized professions were the following: *teaching professionals, health care professionals, child care professionals, pensioners, office workers* and *others*. The category named *others* consisted mostly of students and businessmen.

2.2 Practical arrangements of the data collection

The recruitment letter (consisting of an envelope including an informed consent form and the voice questionnaires, see Appendix D) was sent approximately two weeks before the person's phoniatric appointment. The envelopes were sent directly from the Phoniatric Outpatient Clinic for reasons of confidentiality. The recruitment letter included information about the study and a request to respond to the questionnaires as soon as possible and return them to the hospital before the phoniatric appointment.

When arriving to the phoniatric appointment, the participant was asked to respond to a retest questionnaire of the Screen11, which was identical to the one that the participant

---

[1] The difference in the gender distribution between groups was a result of the exclusions of participants due to missing data.

responded to earlier. The purpose of the retest form of the Screen11 questionnaire was to investigate its external reliability.[2] After the phoniatric examination of the participant, the doctor wrote down the results and the diagnosis on a specific field in the retest form. When the participant had completed all questionnaires (Screen11, the VHI, the VAPP, retest form of Screen11), and been given a diagnosis from the phoniatric examination, the complete data for one participant were compiled and taken to Åbo Akademi University for further analysis.

2.3 Measures

The eleven statements in Screen11 capture the occurrence of vocal symptoms during the past year. The vocal symptom statements in Screen11 are: *Does your voice become strained? Does your voice become tired? Does your voice become hoarse? Does your voice become low? Does your voice break? Do you have difficulties in being heard? Do you need to clear your throat? Do you need to cough? Do you have a sensation of tension in the throat? Do you have a sensation of lump in the throat? Do you have a sensation of pain in the throat?* The Screen11 questionnaire used in this thesis can be found in Appendix 1. The measurement in Screen11 is retrospective as it collects information about voice experiences from the past year. The given frequency alternatives are measured on a Likert scale ranging from *every day (3), every week (2), less often (1)* and *never (0)*, with a total score from 0 to 33. The higher the score, the more severe the subject's experience of vocal symptoms and possible voice disorder. Based on previous studies using Screen6 (Simberg et al., 2000; Simberg et al., 2001) vocal symptoms occurring every day or every week were regarded as frequently occurring vocal symptoms.

In order to examine the convergent validity of Screen11, it was compared against two validated screening questionnaires, the VHI (Jacobson et al., 1997) and the VAPP (Ma and Yiu, 2001). In the VHI, the 30 voice-related statements yield four different outcomes; an overall score consisting of the summed score (range 0–120) of each statement in the questionnaire, and three subsections (with a score range from 0 to 40 in each subsection) measuring emotional, physical, and functional issues. In this thesis, the VHI overall sum score variable and the three subsections as dependent variables were used (especially the

---

[2] In this thesis, the test-retest reliability of the Screen11 was not analyzed due to a small sample size (i.e., only 22 participants responded to the retest questionnaire at the clinic). The external reliability will thus be reported in a subsequent thesis where the sample size will be twice as large compared to the one included in the present thesis.

physical subsection was chosen as it was expected that physical symptoms were strongly correlated with Screen11). The VAPP questionnaire yields altogether five outcome measures capturing different issues related to the voice use: self-perceived severity of voice problem, effect on job, effect on daily communication, effect on social communication and effect on emotion. In this study, only the sum score of all the VAPP subsections was used as the dependent variable.

2.3.1   Sensitivity and specificity in diagnostic tests

Diagnostic tests that distinguish positive and negative outcome from individuals' health condition need a certain accuracy to be assumed as reliable. The term accuracy in this case refers to how well the cut-off score of the questionnaire separates the population into groups with and without the disorder in question. The typical approach when evaluating a diagnostic test, when the outcome is dichotomous, is to use *sensitivity* and *specificity* as measures of accuracy (Swets, 1979). Sensitivity indicates the true positive rate or the actual proportion of the positive outcomes that have been correctly identified with the questionnaire. Specificity, in contrast, stands for the true negative rate or the proportion of the negative outcome where the condition likewise in reality is negative. The purpose of these two values is to avoid false positive and false negative outcomes when using the questionnaire.

2.4 Statistical analyses

The data were analysed with The Statistical Package for the Social Sciences, *SPSS,* version 24. Demographic variables between the help seeking participants and the control participants were examined using independent samples *t*-test for the continuous variable age, and with χ^2 tests for the categorical variables gender and occupation. The internal consistency of the Screen11 questionnaire was probed with the item-total correlations (ITCs). A positive ITC means that a higher score on a given item predicts a higher score on the other items, while a low ITC indicates that the item does not covary with the rest of the items in the task. The internal consistency was also examined by investigating the Cronbach's alpha for the Screen11 questionnaire. The intercorrelations between Screen11 and the well-established voice questionnaires VHI and VAPP were examined using Pearsons's correlation coefficients. High intercorrelations (> .80) would indicate good

convergent validity, whereas low intercorrelations (> .50) would indicate poor convergent validity.

For examining the accuracy, sensitivity and specificity for the Screen11 questionnaire, a Receiver Operator Characteristic (ROC) curve (Zweig & Campbell, 1993) was used, where the value of specificity is interpreted as 1–Specificity. The value of accuracy, interpreted in a ROC curve as the area under the curve, has an excellent accuracy with values 0.9–1 and a failed accuracy with 0.5–0.6 according to a traditional academic point system. Thus, the ROC curve can be used to examine the diagnostic ability of a test, allowing the researcher to find the most optimal cut-off value for classifying the patient group from the control group (Kumar & Indrayan, 2011). As previous research implies that the Screen6 exhibits a good specificity, but poor sensitivity (e.g. Ohlsson et al., 2012), finding a cut-off score weighting both of these diagnostic estimates equally was aimed. For this analysis the coding scheme employed in previous studies was followed (e.g. Fellman & Simberg, 2017; Hagelberg & Simberg, 2015) so that symptoms that occurred daily or weekly were scored with 1 point whereas less frequently or not at all occurring symptoms were scored as 0 points separately for each of the eleven items. This yielded a score range of 0–11 with higher scores representing more frequently occurring voice symptoms.

3 Results

The questionnaires, including Screen11, the VAPP and the VHI, were answered by 57 participants of whom 29 were persons seeking help for voice problems and 28 were control persons considered to have a healthy voice. Nine participants did not respond to all questions in the questionnaires. These participants were thus excluded from the statistical analyses, yielding a final sample of 48 participants (help seeking participants $n = 24$, control participants $n = 24$). An overview of the participants and their demographic characteristics can be seen in Table 1.

Table 1

Demographical characteristics of the help seeking participants and the control participants

	Help seeking participants	Control participants
Gender (*n*)	24	24
Female	20	21
Male	4	3
Age in years		
Mean, (SD)	46.46 (16.20)	46.80 (16.46)
Range	20–78	20–67
Occupation (*n*)		
Teaching professionals	4	5
Health care professionals	2	1
Child care professionals	2	2
Pensioners	3	3
Office workers	2	1
Others	11	12

3.1 Age, gender, occupation and total score distributions

An independent two samples *t*-test was conducted to compare whether the two groups were comparable with regard to age (help seeking participants $M_{age} = 46.46$, $SD = 16.20$; control participants $M_{age} = 46.80$, $SD = 16.46$). The results showed that the age distribution of the participants in the two groups did not differ significantly from each other $t(45.99) = -0.071$, $p = .94$.

Because the variables gender and occupation are dichotomous, Pearsons's Chi-squared test was used to examine the group distributions in these two variables. The results showed that the groups were comparable in terms of gender, $\chi^2(1) = 0.00$, $p = 1.00$, with Yates' continuity correction, and occupation, $\chi^2(5) = 0.82$, $p = .98$, indicating that our control participants were well-matched to the help seeking participants in all key-dependent background variables.

Independent samples t-tests examining the total scores in Screen11 (range 0–33) as a function of group were conducted (help seeking participants: $M_{screen11}$ = 19.96, SD = 6.52; control participants: $M_{screen11}$ = 7.08, SD = 4.26. The results showed statistically significant differences between the groups $t(39.608) = 8.094$, $p < 0.001$, with the help seeking participants reporting higher scores as compared with the controls.

3.2 Internal consistency

Descriptive statistics of the 11 items in Screen11 can be found in Table 2. All items correlated well with each other (ITCs, r = .70–85). Moreover, each item showed sufficient skew and kurtosis levels, suggesting that the item responses in Screen11 were normally distributed. As regards the internal consistency, the results showed a Cronbach's alpha of α = .94, indicating that Screen11 exhibited an excellent internal consistency (see Table 3).

Table 2

Item descriptives and corrected item-total correlations for the Screen11 questionnaire (N = 48)

Item	*M*	*SD*	Skew	Kurtosis	ITC	Range Potential	Actual
Does your voice become strained?	1.52	0.90	0.20	−0.85	.80	0–3	0–3
Does your voice become tired?	1.46	0.92	0.12	−0.88	.85	0–3	0–3
Does your voice become hoarse?	1.60	1.05	0.06	−1.29	.82	0–3	0–3
Does your voice become low?	1.33	1.04	0.22	−1.17	.83	0–3	0–3
Does your voice break?	1.00	1.01	0.85	−0.36	.82	0–3	0–3
Do you have difficulties in being heard?	1.06	1.02	0.59	−0.83	.75	0–3	0–3
Do you need to clear your throat?	1.54	0.97	0.02	−1.03	.68	0–3	0–3
Do you have a sensation of lump in the throat?	0.85	0.87	0.65	−0.56	.74	0–3	0–3
Do you have a sensation of pain in the throat?	0.77	0.90	1.13	0.53	.70	0–3	0–3

Note. ITC = Item-Total Correlation.

Table 3

Item-Total Statistics in Screen11(N = 48)

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
Item 1	12.00	60.04	.80	.88	.94
Item 2	12.06	59.12	.85	.88	.94
Item 3	11.92	57.91	.82	.80	.94
Item 4	12.19	57.90	.83	.79	.94
Item 5	12.52	58.34	.82	.83	.94
Item 6	12.46	59.28	.75	.70	.94
Item 7	11.98	60.91	.68	.64	.94
Item 8	12.23	63.71	.56	.58	.95
Item 9	12.44	58.93	.76	.78	.94
Item 10	12.67	61.16	.74	.76	.94
Item 11	12.75	61.26	.70	.76	.94

Note. Item 1: Does your voice become strained? Item 2: Does your voice become tired? Item 3: Does your voice become hoarse? Item 4: Does your voice become low? Item 5: Does your voice break? Item 6: Do you have difficulties in being heard? Item 7: Do you need to clear your throat? Item 8: Item 9: Do you have a sensation of tension in the throat? Item 10: Do you have a sensation of lump in the throat? Item 11: Do you have a sensation of pain in the throat?

3.3 Convergent validity

The convergent validity of Screen11 was examined by investigating its intercorrelations with VHI (using both the sum score and the three different subsection scores as outcome variables) and VAPP (see Table 4). The intercorrelations between the total scores of VAPP, VHI and Screen11 were statistically significant ($p < .01$). VAPP correlated strongly with the VHI ($r = .89$) and moderately with Screen11 ($r = .69$), whereas Screen11 and VHI showed a strong correlation with each other ($r = .82$). The correlation between VHI's subsection of physical voice symptoms and Screen11 was strong and statistically significant ($r = .90$).

Table 4

Summary of Intercorrelations, Means and Standard Deviations for Scores on Screen11,
VAPP, VHI and the subsections of VHI (N = 48)

Measure	1.	2.	3.	4.	5.	6.	*M*	*SD*	*Cronbach's alpha*
1.Screen11 total score	-	.692**	.820**	.901**	.695**	.728**	13.52	8.49	.94
2.VAPP total score		-	.886**	.811**	.859**	.865**	63.79	67.38	.98
3.VHI total score			-	.943**	.956**	.957**	28.56	22.86	.97
4.VHI physical symptoms				-	.840**	.839**	12.71	8.71	.94
5. VHI functional symptoms					-	.901**	8.25	7.52	.93
6. VHI emotional symptoms						-	7.60	7.79	.93

Note. ** indicates *p* < .01. *M* and *SD* are used to represent Mean and Standard Deviation
for respective scores.

3.4 Accuracy of Screen11

Next, the most optimal cut-off point in Screen11 was examined for discriminating the help
seeking participants from the vocally healthy controls. The results from the ROC curve
analysis showed that the area under the curve was 0.94 (*p* < .001), which stands for an
excellent accuracy for a test (see Figure 1 for an illustration). Moreover, the results showed
that the most optimal cut-off score was four or more frequently occurring (weekly or more
often) voice symptoms when specificity and sensitivity was equally weighted. The
sensitivity of Screen11 was 0.875 when the cut-off score was set to four, indicating that
87.5% of the help seeking participants reported four or more voice symptoms occurring
weekly. Using the same cut-off score, the specificity showed a true negative rate of 0.912,
implying that 91.2% of the control group participants reported three or fewer occurring
voice symptoms per week (see also Table 5 depicting the frequency distribution of
occurred voice symptoms in the two groups using four or more voice symptoms as the cut-
off for having a voice disorder). Thus, the results indicated that the cut-off score of four or
more frequently occurring voice symptoms weekly in Screen11 was the most optimal

threshold for discriminating the help seeking participants from the healthy control participants.
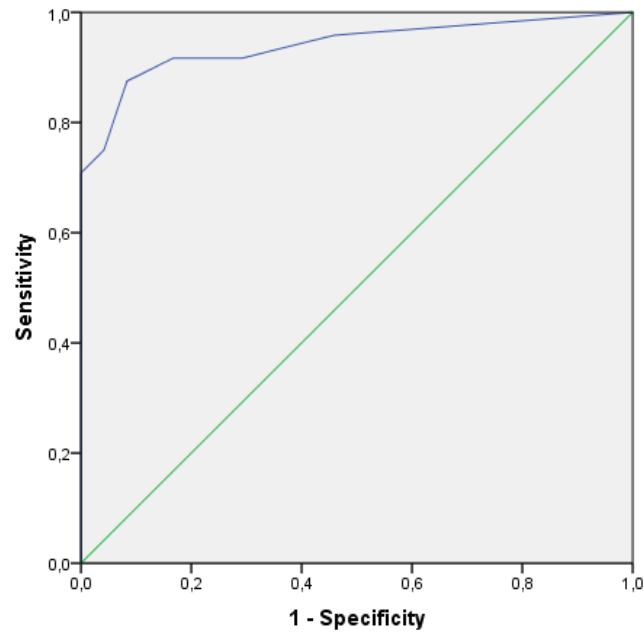
Figure 1

Receiver Operator Characteristic (ROC) curve of Screen11.

Table 5

Group distribution as a function of frequency of vocal symptoms (N = 48)

	Group		Total
	help seeking participants	control participants	
Less frequently occurring symptoms	3	22	25
Four or more frequently occurring symptoms	21	2	23
Total	24	24	48

Note. Frequently occurring = voice symptom occurring daily or weekly.

4 Discussion

Voice disorders are a common health issue especially among persons with voice demanding occupations (Verdolini & Ramig, 2001), complicating several aspects in life and therefore causing increased sick leaves (Van Houtte et al., 2010), involuntary career changes (Roy, Merrill, Thibeault, Gray, & Smith, 2004), and reduced quality of life (Verdolini & Ramig, 2001). Thus, vocal symptoms should be observed and detected in an early stage to avoid them from developing to voice disorders. A fast and easily-administered screening method developed for capturing early signs of a possible voice disorder could have the potential to prohibit more severe voice-related issues.

The aim of this study was to validate an updated voice screening questionnaire (titled Screen11) consisting of only 11 questions concerning voice-related issues. The eleven items (i.e. questions) in Screen11 rely on a voice question battery that has been used in several previous studies investigating the prevalence of voice problems (e.g. Simberg et al., 2000; Simberg et al., 2001). However, the items used in the previous studies have been noticed for several shortcomings, such as including items with double-barreled questions and a lack of a careful investigation of the internal consistency of the items. In this validation procedure, such shortcomings have been considered, suggesting that Screen11 is a more precise screening tool with its exact statements focusing on vocal symptoms, without misunderstandings or over-reporting.

The internal consistency, convergent validity and accuracy of the Screen11 were investigated to further explain the precision of Screen11. Moreover, the best cut-off score in Screen11 was determined for differentiating the participants who sought help for voice disorders from those who did not.

As method for this thesis, self-evaluation voice questionnaires, including Screen11, were answered by two different populations and the answers were compared with statistical analyses. The demographic variables age, gender, and occupation were comparable between the voice patient group and the control group, indicating that our matching procedure was successful. Thus, it can be concluded that the demographical characteristics between the groups were similar and did not have any effect on the results. The occupation distribution was similar in both groups; however, the number of participants was highest in the group named Others. The division of occupation categories was made based on the study by Sukanen et al. (2007) which also included an occupation category named Priests and lecturers, but there was no need for such group in the present

study. The age distribution included a wide range in both groups, indicating that the persons who seek help for their voice problems are people in different ages and life situations. As expected, most of the help seeking population were females, which is in line with previous studies (Williams, 2003).

4.1 Statistical strengths of Screen11

The eleven statements of Screen11 appeared to have an excellent internal consistency, indicating that they correlated well with each other. Importantly, it confirms that none of the statements were emphasizing or distorting the total result. The results also indicate that each of the eleven statements (i.e., the different vocal symptoms measured) are essential for the Screen11 total score, with none of them being redundant. The relevant statistics when exploring the internal consistency is ITCs, which in the case of Screen11 showed a very good distribution. The test is used to probe if any of the eleven statements (items) in Screen11 disagree or stand out with each other, and therefore it is used to decide whether an item should be included in the pool of questions. According to the results of Screen11's ITC, every single statement in the questionnaire is necessary for the total score and therefore for the distinguishing in scores between persons with potential voice disorders and vocally healthy persons. The previous versions of Screen11 lack an examination of the internal consistency across items, suggesting that the present version provides more reliable results as compared to its predecessors.

The results showed that Screen11 correlated positively with the two commonly used voice questionnaires, VHI and VAPP, indicating that Screen11 exhibited good convergent validity. As both VHI and VAPP are validated and well-established voice screening questionnaires, their moderate-to-strong intercorrelations with Screen11 indicates that the latter one captures voice-related issues as well. However, the strength between the intercorrelations of these three questionnaires varies slightly. Screen11 showed the strongest intercorrelation with VHI's scores from the physical subsection. This was in line with the hypothesis, potentially stemming from the fact that Screen11 consists of only physical statements in its entirety. Therefore, it is not surprising that the intercorrelation was stronger with only comparing the physical domain as when comparing Screen11 with the total scores of VHI. In contrast to the Screen11, the aim of VHI total score is to evaluate the perceived limitation the vocal symptoms cause.

Another notion regarding the intercorrelations was that VHI and VAPP had stronger intercorrelations than Screen11 and VAPP. This is expected and could be explained by the fact that VHI and VAPP evaluate aspects concerning perception of limitations and participation including emotional and psychosocial sections in their pool of questions, which Screen11 is lacking from. The purpose with Screen11 itself is to concentrate solely on the physical symptoms. If the VoiSS (Deary et al., 2003) had been available in Finnish, it might have been informative to use this questionnaire when analyzing the convergent validity of the Screen11. VoiSS includes 44 items focusing mostly on the vocal symptoms, but psychosocial questions are included as well, yet to a lesser extent (Deary et al., 2003).

The precision of Screen11 (i.e., the accuracy of the questionnaire) was excellent in terms of capturing those individuals seeking help for their voice problems from those individuals without voice problems. Using a ROC curve analysis, the results showed that four or more frequently occurring voice symptoms was the most adequate cut-off score, indicating that when a person suffers from four or more frequently occurring (i.e. weekly or more often) voice symptoms, it refers to a possible voice disorder. When the score is less than four, the probability of a voice disorder is less likely. In previous studies (Simberg et al., 2000; Simberg et al., 2001), the definition of voice disorder has been a cut-off score of two vocal symptoms occurring weekly or more often. Although this criterion has not been investigated in any statistical way, it has been compared to scores from other questionnaires and the laryngeal status of the participants (Simberg et al, 2000; Simberg et al. 2001). The different cut-off score in the previous study might also depend on the fewer number of items. Concerning the accuracy, the sensitivity of Screen11 showed to be stronger than the sensitivity of Screen6. Altogether 87.5% of the participants who reported four or more vocal symptoms weekly or more often belonged to the voice patient group. In case of using Screen11 as a mean for screening for voice disorders, this would indicate that 87.5% of those reporting four or more frequently occurring symptoms actually are in risk for developing voice disorders.

4.2 Screen11 in comparison with previous studies

Several previous studies about voice disorders and vocal symptoms have been made with participants only from teaching occupations (Sapir et al., 1993; Smith et al., 1998; Roy et al., 2004; Ohlsson et al., 2012). The sample of participants in this particular study represents a wide range of different occupations and therefore different voice demands,

which in the future allows a more extensive use of Screen11. Because of the variation of characteristics and voice symptoms in the sample, the results are more suitable for generalization in the general population.

The convergent validity analysis showed a very strong correlation between Screen11 and the VHI's subsection for physical symptoms. This result leads to a consideration of using only the physical subsection of VHI in charting for vocal symptoms, because it would be shorter and therefore less time consuming. In fact, Rosen, Lee, Osborne, Zullo and Murry (2004) designed a shortened version, *VHI 10*, from the original 30-item VHI form. The VHI 10 consists of ten items from all the three subsections for functional, physical and emotional symptoms, and a total score of 11 (out of 40) or higher should be considered as an abnormal voice (Arffa, Krishna, Gartner-Schmidt, & Rosen, 2011). Although this shortened version of VHI is developed, the original 30-item VHI is commonly used in clinics. Another shorter voice questionnaire based on the VHI has been developed by Lyberg-Åhlander, Rydell, Eriksson and Schalén (2010), where the focus is on throat symptoms. This version is coined as *VHI-Throat* or *VHI-T* and consists of 10 throat-related questions, i.e. sensation of globus and irritated throat. The VHI-Throat is recommended to be used as a complementary questionnaire with the VHI, because it allows a better identification of the actual voice disorder.

Inspecting this variety of different shortened versions of the VHI and other voice questionnaires, the advantages of Screen11 is debatable. Screen11 is developed as an entity and it does not necessarily have to be used as a complement to other voice questionnaires. Potentially, it could be followed by the VHI and the VAPP if the scores would refer to a possible voice disorder, and with these more extensive questionnaires the voice disorder could be more charted. Screen11 questionnaire, however, aims to be a screening method.

4.3 Limitations of the study

Already during the data collection for this study, some improvement remarks were noticed, such as the necessity for even more precise instructions concerning how to respond to the questions in Screen11. The improvements have been implemented in the study followed by this thesis. These improvements should minimize the amount missing data and, allowing for a more extensive sample size.

There are no general criteria for sample size in a validation study, but a common recommendation for a questionnaire study such as this present study, is 50–100 (Statistics

Solutions, 2018). The sample size for this study was 48, with the actual test group (i.e., those individuals seeking help for voice problems) consisting of 24 participants. The aim in validation studies is to gather a sample as heterogeneous as possible (Statistics Solutions, 2018) which can be more difficult to achieve in a small sample size. However, the data in this present study included a large range of demographic characteristics. Although the statistical analyses indicated good psychometric properties in Screen11, the results could be criticized due to the limited number of participants.

In the present study, altogether nine participants (help seeking participants $n = 5$, control participants $n = 4$) had missing data in the questionnaires. Thus, 5% of the participants were excluded from the analyses. The reason for the missing data points were multifactorial. For instance, pensioners and students might have been ignoring work-related questions for not being relevant to them, whereas other participants might have missed some questions only by oversight. Eventually, the two groups included an equal number of participants. However, the missing data resulted in some pairs not being matched on gender and exact age.

In comparison with other self-reporting voice questionnaires, Screen11 is the only one which gathers information about voice symptoms from the past year. In one year, the symptoms can vary greatly and occur in different frequencies. Some of the participants in the data collection phase had made notes on the questionnaire explaining that the given answers reflect their present voice situation, and not the situation a year ago. The retrospective factor is probably enrichening the setup of Screen11, but the time period could possibly be shortened from one year to a few months.

The Screen6 questionnaire comprises more answer options than Screen11. Screen6 as well as Screen11 gathers information from the past year but the options for occurred frequency are more encompassing: *1) Every day or most days 2) Weekly or most weeks 3) Monthly or most months 4) Less often than above 5) Only periodic symptoms 6) No symptoms*, compared with Screen11's alternatives: *1) Every day 2) Every week, 3) Less often 4) Never*. When the time period is as extensive as one year in Screen11, an option as *Only periodic symptoms* could be necessary to insert in the pool of answers.

4.4 Conclusions and future research

This thesis sought to validate a potential voice screening tool, titled Screen11, consisting of only 11 voice-related questions. The method in the study comprised self-evaluation questionnaires and statistical analyses of the data.

The results indicated that the eleven statements of Screen11 appeared to have an excellent internal consistency and additionally that Screen11 has a good convergent validity in correlation with VHI and VAPP. Moreover, a cut-off score of four or more frequent vocal symptoms occurring weekly or more often was explored. The cut-off score distinguished help seeking participants from control participants.

Future research should yet complete a validation of Screen11 with a larger sample. Screen11 could be a fast and easily-administered method among persons with voice demanding occupations, and it could be conducted clinically by health care professionals or answered by the potential client right before a meeting with a speech and language pathologist.

Swedish summary - Svensk sammanfattning

Validering av ett snabbt och enkelt screeningformulär

Introduktion

Röstsymtom och röststörningar är hälsoproblem som har belastande inverkan på människors vardag och livskvalitet (Verdolini & Ramig, 2001). Ifall röstsymtom inte identifieras i tid kan de leda till sjukledigheter och därmed också öka samhällskostnader (Van Houtte m.fl., 2010). För att hindra röstsymtom från att utvecklas till röststörningar borde de upptäckas och kartläggas så tidigt som möjligt, vilket kunde verkställas till exempel genom screening av dem.

Screening är en procedur där målet är att upptäcka eventuella sjukdomar, fysiska eller psykiska störningar eller missbildningar i ett tidigt skede bland en population som befinner sig i riskzonen för det som screenas (National Screening Unit, 2018). I bästa fall är screeningen snabb, enkel och billig att genomföra (ASHA, 2018). Screening av röststörningar förverkligas vanligtvis genom självskattningsformulär. *Voice Handicap Index* (VHI; Jacobson m.fl., 1997) och *Voice Activity and Participation Profile* (VAPP; Ma & Yiu, 2001) är validerade frågeformulär som används kliniskt för att kartlägga den subjektiva uppfattningen av individens egna röstproblem. VHI innefattar 30 frågor som är indelade i tre sektioner som behandlar frågor om emotionella, funktionella och fysiska symtom (Jacobson m.fl., 1997). VAPP består av 28 röstrelaterade frågor om aktivitet och delaktighet som besvaras på en visuell analog skala. Frågorna i VAPP är indelade i fem sektioner som berör den egna uppfattningen om röstproblemets svårighet, dess inverkan på arbetsförmåga, dess påverkan på daglig kommunikation, dess påverkan på sociala relationer och det emotionella livet (Ma & Yiu, 2001). Trots att dessa två röstfrågeformulär är omfattande och mycket använda så finns det behov för ett snabbt och enkelt genomförbart frågeformulär, där endast de mest relevanta frågorna kring röstsymtomen skulle inkluderas. Ett nytt frågeformulär, som består av elva symtomrelaterade påståenden, har undersökts noggrannare i den här avhandlingen. Syftet var att utreda ifall frågeformuläret, som kallas för Screen11, kunde användas som ett lämpligt screeningverktyg för eventuella röststörningar.

Inom litteraturen åtskiljs ofta termerna *röstsymtom* och *röststörning* från varandra. Den distinktionen har använts också i den här avhandlingen. Termen röstsymtom syftar på temporära symtom i halsen, rösten eller röstorganen medan röststörning förekommer när en individ upplever sin röst avvikande eller att rösten inte klarar av de krav som ställs på

den i vardagen. Det råder relativt stora skillnader mellan resultaten i prevalensstudier om röststörningar, vilket kan bero på skillnader i hur röststörningar är definierade eller operationaliserade. I en nyligen utförd svensk studie på den allmänna befolkningen ($n >$ 100 000) var prevalensen av röststörningar 16,9 % (Lyberg-Åhlander m.fl., 2018). Amerikanska studier påvisar att prevalensen av röststörningar är mellan 3 % och 9 % i den amerikanska befolkningen (Ramig & Verdolini, 1998; Roy m.fl., 2005). Resultaten skiljer sig alltså något mellan länder.

Resultat från flertalet studier visar att individer som jobbar inom röstkrävande yrken löper större risk att få röststörningar (Hocevar-Boltezar, 2009; Linnasalo, 1990; Roy m.fl., 2004). Röstkrävande yrken är bland annat lärare, barnträdgårdslärare, präster och skådespelare. Personer som arbetar inom röstkrävande yrken utsätter rösten och röstorganen för större belastning, vilket ökar risken för rösttrötthet och röstöveransträngning (Vilkman, 2004; Whitling m.fl., 2015). Könsskillnader har också en inverkan på förekomsten av röststörningar. Jämfört med män, drabbas kvinnor oftare av röstsymtom och röststörningar, vilket möjligen kan bero på laryngeala strukturskillnader mellan könen (Cohen m.fl., 2012; Martins m.fl., 2016; Roy m.fl., 2005). Kvinnor har bland annat kortare stämband än män (Titze, 1994) och därför också en högre stämbandsfrekvens, vilket leder till större belastning på stämbanden (Butler m.fl., 2001; Hunter m.fl., 2011). En ytterligare riskfaktor för röststörningar är ålder. Åldrandet förorsakar anatomiska förändringar i laryngeala strukturer och ökar risken för röstsymtom och röststörningar (Martins m.fl., 2016).

Simberg med flera (2001) har initierat och utforskat ett röstrelaterat frågeformulär som påbörjades med en studie (Simberg m.fl., 2000) i vilken prevalensen av röststörningar hos finska lärarstuderanden ($N = 226$) undersöktes. Deltagarna fick besvara ett röstrelaterat formulär och deras röster blev perceptuellt bedömda av en talterapeut och foniatriskt undersökta av en läkare. Resultaten visade att 20 % av studerandena hade röststörningar, varav 19 % var organiska röststörningar. I den påföljande studien (Simberg m.fl., 2001) reducerades antalet frågor till åtta stycken där resultaten visade att när deltagaren rapporterade två eller flera röstsymtom i veckan tydde det på en möjlig röststörning. Simberg med flera (2004) fortsatte att undersöka prevalensen av röststörningar på finska lärarstudeanden genom att använda frågeformuläret som screeningverktyg. Resultaten påvisade bland annat att formuläret kunde reduceras till ännu färre frågor. Baserat på dessa resultat utvecklade Ohlsson med flera (2012) en förkortad version av frågeformuläret som bestod av endast sex påståenden och kallades därför för Screen6.

Screen6 har därefter använts bland annat inom röstforskning, men det finns en del brister i den och dess användning. Flera av påståendena i Screen6 är tvådelade (till exempel *Rösten blir trött och ansträngd),* vilket kan förorsaka missförstånd eller ge vaga svar. Dessutom har den interna konsistensen av Screen6 inte undersökts, vilket betyder att frågeformulärets interna reliabilitet tillsvidare är okänd. Det har även påvisats att sensitiviteten av Screen6 är något svag (Ohlsson m.fl., 2012), vilket väcker följdfrågan ifall det nuvarande gränsvärdet för röststörning (två eller flera återkommande röstsymtom i veckan) är den mest optimala för att urskilja personer som har röststörningar från dem som inte har det. Den uppdaterade versionen av frågeformuläret, Screen11, består av elva symtomrelaterade påståenden. Sociala eller emotionella områden kartläggs inte med Screen11, eftersom den endast fokuserar på röstsymtomen.

Syftet med avhandlingen var att utforska ifall frågeformuläret Screen11 kan användas som ett screeningverktyg för röststörningar. De specifika syftena var att kontrollera den interna konsistensen och konvergenta validiteten av formuläret samt bestämma ett gränsvärde för återkommande röstsymtom som kunde tyda på en röststörning.

Metod

Den här studien har utförts i samarbete med Åbo Centralsjukhus (ÅUCS) foniatriska poliklinik. Forskningsplanen har godkänts av *Clinical Research Center* i Åbo och Åbo Akademis forskningsetiska nämnd.

I oktober 2017 påbörjades rekryteringen av deltagare till studien. Personer som var i kö för foniatrisk mottagning vid ÅUCS foniatriska poliklinik fick ett informationsbrev med forskningsmaterialet och samtycket hemskickat med posten. Ett antal på 29 besvarade material returnerades till polikliniken, varav fem deltagare exkluderades på grund av ofullständigt ifyllda frågeformulär. Dessa 24 deltagare var personer som sökte hjälp för sina röstproblem. De matchades med röstfriska kontrollpersoner ($n = 24$) för egenskaperna ålder, kön och yrke. Det slutgiltiga samplet som inkluderades i analyserna bestod av 48 deltagare varav hälften var personer som sökte hjälp för sina röstproblem och den andra hälften röstfriska kontrollpersoner som hade rekryterats genom bekvämlighetsurval. Majoriteten av deltagaren var kvinnor ($n = 41$) och genomsnittsålderåldern var 47, med en variationsvidd mellan 20 och 78 år. Yrkesvariabeln var indelad i sex yrkesgrupper: lärare, hälsovårdspersonal, barnvårdspersonal, pensionärer, kontorsarbetare och övriga.

Datainsamlingen för varje hjälpsökande deltagare innebar att fylla i tre röstfrågeformulär (Screen11, VHI och VAPP) som skickats hem med posten ungefär två veckor innan mottagningen på den foniatriska polikliniken samt den foniatriska undersökningen. Kontrollpersonerna besvarade endast frågeformulären.

Screen11 bestod av följande elva röstrelaterade påståenden: *1) Rösten blir ansträngd 2) Rösten blir trött 3) Rösten blir hes 4) Rösten blir låg 5) Rösten bryter vid tal 6) Svårigheter med att få rösten hörd 7) Behov av att harkla sig 8) Behov av att hosta 9) Känsla av spänning kring halsen 10) Känsla av en klump i halsen 11) Känsla av smärta i halsen.* Svarsalternativen till påståendena var följande: *0 = aldrig, 1 = sällan, 2 = varje vecka, 3 = dagligen.*

Data analyserades med programmet The Statistical Package for the Social Sciences, *SPSS,* version 24. Den demografiska variabeln ålder utforskades genom oberoende t-test mellan de två grupperna (hjälpsökande personer och kontrollpersoner) och med Chi-två-test för de kontinuerliga variablerna kön och yrke. Den interna konsistensen av Screen11, det vill säga hur väl de elva påståendena i formuläret mäter samma sak, undersöktes genom item-total korrelationer (ITCs) mellan påståendena och genom att undersöka värdet på Cronbachs alfa (α) över dem. För att analysera den konvergenta validiteten, det vill säga hur starkt Screen11 korrelerar med de tidigare röstrelaterade frågeformulären VHI och VAPP, användes Pearsons korrelationskoefficienter. Höga koefficienter mellan Screen11 och de andra röstformulären pekar på stark konvergent validitet medan låga koefficienter tyder på svag konvergent validitet. Precisionen av Screen11, det vill säga dess noggrannhet, utvärderades genom en *Receiver Operator Characteristic Curve* (ROC-kurvanalys). Precisionen tolkas genom arean under ROC-kurvan, där 0,9–1 betecknar utmärkta värden och 0,5–0,6 tyder på misslyckad noggrannhet. Med hjälp av ROC-kurvan kan man bestämma ett gränsvärde mellan två olika grupper genom att kartlägga det gränsvärde som har högst specificitet och sensitivitet. Sensitiviteten är värdet för de sanna positiva testresultaten (de deltagare som korrekt kategoriserades till den hjälpsökande gruppen) och specificitet är värdet för sanna negativa testresultat (de deltagare som korrekt kategoriserades till den röstfriska gruppen).

Resultat

Resultaten visade att åldern i de två grupperna (hjälpsökande deltagare $M_{ålder} = 46{,}46$, $SD = 16{,}20$; kontrolldeltagare $M_{ålder} = 46{,}80$, $SD = 16{,}46$) var jämförbara $t(45{,}99) = -0{,}071$, $p = 0{,}94$. Så var också de dikotoma demografiska variablerna för kön och yrke, som testades

genom Pearsons Chi-två-analys: kön $\chi^2(1) = 0,00$, $p = 1,00$ och yrke: $\chi^2(5) = 0,82$, $p = 0,98$. Skillnaden mellan totalpoängen från Screen11 var signifikant mellan de två grupperna (hjälpsökande delatagare: $M_{screen11} = 19,96$, $SD = 6,52$; kontrolldeltagare: $M_{screen11} = 7,08$, $SD = 4,26$, vilken också utfördes genom oberoende t-test: $t(39.608) = 8,094$, $p < 0,001$.

Den interna konsistensen av Screen11 kontrollerades med hjälp av ITC mellan påståendena i Screen11 (ITCs, $r = 0,70–0,85$) och med Cronbachs alfa $\alpha = 0,94$, vilket pekar på en stark intern konsistens. Den konvergenta validiteten av Screen11 utforskades genom interkorrelationer mellan totalpoängen av VAPP, VHI och VHIs subskalor. Interkorrelationerna mellan alla dessa jämförda variabler var statistiskt signifikanta ($p < 0,01$.) VAPP korrelerade starkast med VHI ($r = 0,89$) och måttligt med Screen11 ($r = 0,69$), medan Screen11 och VHI hade en stark korrelation ($r = 0,82$). Korrelationen mellan VHIs fysiologiska subskala och Screen11 hade också en stark och statistiskt signifikant korrelaton ($r = 0,90$).

Precisionen av Screen11 utforskades med en ROC-kurva, där området under kurvan var 0,94 ($p < 0,001$), vilket antyder en utmärkt precision. Samtidigt bestämdes ett optimalt gränsvärde för antalet upplevda röstsymtom som urskiljer hjälpsökande personer från röstfriska kontrollpersoner. Gränsvärdet visade sig vara fyra eller flera återkommande röstsymtom i veckan eller dagligen då både sensitivitet och specifitet togs i beaktande likvärdigt. Värdet för sensivitet vid gränsvärdet fyra var 0,875 vilket innebär att 87,5 % av de hjälpsökande personerna hade rapporterat fyra eller flera röstsymtom i veckan. Värdet för specificitet vid gränsvärdet fyra var 0,912 vilket däremot innebär att 91,2 % av de röstfriska kontrollpersonerna rapporterade tre eller färre röstsymtom i veckan.

Diskussion

Syftet med den här avhandlingen var att validera ett uppdaterat röstrelaterat frågeformulär (Screen11) som eventuellt kunde användas som ett screeningverktyg för röststörningar. I studien validerades Screen11 genom att undersöka den interna konsistensen och den konvergenta validiteten samt genom en fastställning av ett gränsvärde som bäst åtskiljer personer med eventuell röststörning från röstfriska personer. Resultaten visade att den interna konsistensen och den konvergenta validiteten av Screen11 visade sig vara starka, vilket tyder på en lyckad validering. ROC-kurvan och värdena för sensitivitet och specificitet visade att fyra eller flera återkommande röstsymtom i veckan var det mest optimala gränsvärdet för att åtskilja deltagare som sökte hjälp för sina röstproblem från de röstfriska deltagarna.

Screen11 och VHIs subskala för de fysiska symtomen korrelerade starkt. Detta resultat pekar på att subskalan eventuellt kunde användas separat vid kartläggning av röstsymtom, eftersom den skulle vara tidsmässigt kortare att genomföra. Det bör noteras att förkortade versioner av röstscreeningformulär till viss del redan undersökts. Rosen med flera (2004) har bland annat utvecklat en förkortad version av VHI som består av 10 påståenden om funktionella, fysiska och emotionella röstrelaterade symtom, och kallas därför för *VHI 10*. VHI med de 30 påståendena används ändå mera kliniskt än VHI 10. Lyberg-Åhlander med flera (2010) har utvecklat en ytterligare förkortad version av VHI som fokuserar på symtom kring halsen och strupen. Versionen kallas för *VHI Throat* eller *VHI T* och innehåller också 10 påståenden. VHI T rekommenderas som en komplettering vid utförandet av VHI. Screen11, till skillnad från de ovanförnämnda förkortade versionerna av VHI, är utvecklat som ett separat verktyg för kartläggning av röstsymtom och identifiering av eventuella röststörningar och behöver nödvändigtvis inte kompletteras med övrigt material.

En del metodologiska begränsningar bör noteras i föreliggande studie. Storleken på studiens sampel var litet. Flera potentiella deltagare exkluderades för att de hade missuppfattat instruktionerna eller inte svarat på alla påståenden. Anvisningarna kunde alltså vara ännu tydligare för att undvika missförstånd och därmed exkludering av deltagare.

Retrospektivt insamlat data ger information om förfluten tid och anses ofta som en berikande egenskap för samplet. I den här studien samlades information kring upplevda röstsymtom sedan ett år tillbaka, vilket kan upplevas som en för lång tidsperiod att erinra sina röstsymtom. Tidsperioden kunde förkortas till ett antal månader. Dessutom kunde svarsalternativen i Screen11 vara flera än fyra (*0 = aldrig, 1 = sällan*, 2 = *varje vecka, 3 = dagligen*), liksom i Screen6 där det också finns alternativet *Periodvis.*

Syftet med den här avhandlingen var att validera Screen11, som är ett frågeformulär bestående av elva symtomrelaterade påståenden. Resultaten visade stark intern konsistens mellan påståendena i Screen11 och utmärkt konvergent validitet med de röstrelaterade frågeformulären VHI och VAPP. Ett gränsvärde av fyra eller flera återkommande röstsymtom i veckan visade sig vara det mest optimala värdet för att åtskilja röstfriska personer från personer med eventuell röststörning. En replikering av den här studien, utförd med ett större sampel, vore önskvärt i framtida forskning. Screen11 kunde i framtiden användas kliniskt som ett snabbt och enkelt screeningsverktyg för röststörningar.

References

Alaluusua, S., & Johansson, M. (2003). Äänihäiriöiden aiheuttama psykososiaalinen haitta ja sen kuntoutuminen. Voice Handicap Index'n suomennoksen   kokeilututkimus. *Logopedian pro gradu tutkielma.* Helsingin yliopisto.

American Speech-Language-Hearing Assosiation, (2018). Retrieved from https://www.asha.org/practice-portal/clinical-topics/voice-disorders/

Amir, O., Ashkenazi, O., Leibovitzh, T., Michael, O., Tavor, Y., & Wolf, M. (2006). Applying the Voice Handicap Index (VHI) to dysphonic and nondysphonic Hebrew speakers. *Journal of Voice, 20*, 318–324. doi: 10.1016/j.jvoice.2005.08.006

Arffa, R. E., Krishna, P., Gartner-Schmidt, J., & Rosen, C. A. (2012). Normative values for the voice handicap index-10. *Journal of Voice, 26*, 462–465. doi: 10.1016/j.jvoice.2011.04.006

Aronson, A. E. (1990). *Clinical voice disorders: an interdisciplinary approach*. New York: Thieme.

Buekers, R. (2001). Voice dosimetry. *Occupational voice: Care and Cure*, 21.

Butler, J. E., Hammond, T. H., & Gray, S. D. (2001). Gender-related differences of hyaluronic acid distribution in the human vocal fold. *Laryngoscope*, *111*, 907–911. doi: 10.1097/00005537-200105000-00029

Cohen, S. M., Kim, J., Roy, N., Asche, C., & Courey, M. (2012). Prevalence and causes of dysphonia in a large treatment-seeking population. *Laryngoscope, 122*, 343–348. doi: 10.1002/lary.22426

Cutiva, L. C. C., Vogel, I., & Burdorf, A. (2013). Voice disorders in teachers and their associations with work-related factors: a systematic review. *Journal of Communication Disorders*, *46*, 143–155. doi: 10.1016/j.jcomdis.2013.01.001

Deary, I. J., Webb, A., Mackenzie, K., Wilson, J. A., & Carding, P. N. (2004). Short, self-report voice symptom scales: psychometric characteristics of the voice handicap index-10 and the vocal performance questionnaire. *Otolaryngology—Head and Neck Surgery*, *131*, 232–235. doi: 10.1016/j.otohns.2004.02.048

Deary, I. J., Wilson, J. A., Carding, P. N., & MacKenzie, K. (2003). VoiSS: a patient-derived voice symptom scale. *Journal of Psychosomatic Research*, *54*, 483–489. doi: 10.1016/S0022-3999(02)00469-5

Dietrich, M., Abbott, K. V., Gartner-Schmidt, J., & Rosen, C. A. (2008). The frequency of perceived stress, anxiety, and depression in patients with common pathologies affecting voice. *Journal of Voice*, *22*(4), 472–488. doi: 10.1016/j.jvoice.2006.08.007

Fellman, D., & Simberg, S. (2017). Prevalence and risk factors for voice problems among soccer coaches. *Journal of Voice*, *31*, 121–129. doi: 10.1016/j.jvoice.2016.02.003

Fritzell, B. (1996). Voice disorders and occupations. *Logopedics Phoniatrics Vocology*, *21*, 7–12. doi: 10.3109/14015439609099197

Guimaraes, I., & Abberton, E. (2004). An investigation of the Voice Handicap Index with speakers of Portuguese: preliminary data. *Journal of Voice, 18*, 71–82. doi: 10.1016/j.jvoice.2003.07.002

Hagelberg, A. M., & Simberg, S. (2015). Prevalence of voice problems in priests and some risk factors contributing to them. *Journal of Voice*, 29, 389–411. doi: 10.1016/j.jvoice.2014.08.015

Hellweg, S. A., Pfau, M., & Brydon, S. R. (1992). *Televised presidential debates: Advocacy in contemporary America.* Praeger Publishers.

Hemler R.J.B., Wieneke, G. H., Lebacq J., Dejonckere, P.H. (2001). Laryngeal mucosa elasticity and viscosity in high and low relative humidity. *European Archives of Oto-Rhino Laryngology* , 258, 125–129.

Hocevar-Boltezar, I. (2009). Prevalence and risk factors for voice problems in priests. *Wiener Klinische Wochenschrift, 121,* 276–281. doi: 10.1007/s00508-009-1163-1

Holmqvist, S., Santtila, P., Lindström, E., Sala, E., & Simberg, S. (2013). The association between possible stress markers and vocal symptoms. *Journal of Voice*, 27, 787–e1. doi:10.1016/j.jvoice.2013.06.012

Hsiung, M. W., Pai, L., & Wang, H. W. (2002). Correlation between voice handicap index and voice laboratory measurements in dysphonic patients. *European Archives of Oto-Rhino Laryngology, 259*, 97–99.

Hunter, E. J., Tanner, K., & Smith, M. E. (2011). Gender differences affecting vocal health of women in vocally demanding careers. *Logopedics Phoniatrics Vocology*, 36, 128 –136. doi:10.3109/14015439.2011.587447

Jacobson, B. H., Johnson, A., Grywalski, C., Silbergleit, A., Jacobson, G., Benninger, S., & Newman, C.W. (1997). The voice handicap index (VHI): development and validation. *American Journal of Speech-Language Pathology, 6,* 66–70. doi: 058-0360/97/0603-006

Kooijman, P. G. C., De Jong, F. I. C. R. S., Thomas, G., Huinck, W., Donders, R., Graamans, K., & Schutte, H. K. (2006). Risk factors for voice problems in teachers. *Folia Phoniatrica et Logopaedica*, 58, 159–174. doi: 10.1159/000091730

Kumar, R., & Indrayan, A. (2011). Receiver operating characteristic (ROC) curve for medical researchers. *Indian Pediatrics, 48*, 277–287. doi:10.1007/s13312-011-0055-4

Linnasalo, A-L. (1990). The number and earlier treatment of speech difficulties experienced by students and the number and background of voice disorder symptoms. (In Finnish). Research reports, Department of Speech Education, University of Jyväskylä.

Lyberg-Åhlander, V., Rydell, R., Eriksson, J., & 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. doi:10.1186/1472-6815-10-5

Lyberg-Åhlander, V., Rydell, R., Fredlund, P., Magnusson, C., & Wilén, S. (2015). Prevalence of voice disorders in the general population in Sweden. In *PAN EUROPEAN VOICE CONFERENCE ABSTRACT BOOK* (p. 169).

Lyberg-Åhlander, V., Rydell, R., Fredlund, P., Magnusson, C., & Wilén, S., (In press). Prevalence of voice disorders in the general population, based on the Stockholm public health cohort.

Ma, E. P-M., & Yiu, E. M-L. (2001). Voice activity and participation profile: assessing the impact of voice disorders on daily activities. *Journal of Speech, Language and Hearing Research, 44,* 511–524.

Martins, R. H. G., Do Amaral, H. A., Tavares, E. L. M., Martins, M. G., Gonçalves, T. M., & Dias, N. H. (2016). Voice disorders: etiology and diagnosis. *Journal of Voice*, *30*, 761–e1. doi: 10.1016/j.jvoice.2015.09.017

Morrison, M. D., & Rammage, L. A. (1993). Muscle misuse voice disorders: description and classification. *Acta Oto-Laryngologica, 3,* 428–434.

Morton, V., & Watson, D. R. (1998). The teaching voice: problems and perceptions. *Logopedics Phoniatrics Vocology, 23,* 133–139.

Murry, T., Medrado, R., Hogikyan, N. D., & Aviv, J. E. (2004). The relationship between ratings of voice quality and quality of life measures. *Journal of Voice*, *18*, 183–192. doi: 10.1016/j.jvoice.2003.11.003

National Screening Unit. (2018). Retrieved from https://www.nsu.govt.nz/about-us-national-screening-unit/what-screening

Ohlsson, A. C., Andersson, E. M., Södersten, M., Simberg, S., & Barregård, L. (2012). Prevalence of voice symptoms and risk factors in teacher students. *Journal of Voice*, *26*, 629–634. doi: 10.1016/j.jvoice.2011.11.002

Ohlsson, A. C., Andersson, E. M., Södersten, M., Simberg, S., Claesson, S., & Barregård, L. (2016). Voice disorders in teacher students—A prospective study and a randomized controlled trial. *Journal of Voice*, *30*(6), 755–e13. doi: 10.1016/j.jvoice.2015.09.004

Ohlsson, A. C., & Dotevall, H. (2009). Voice handicap index in Swedish. *Logopedics Phoniatrics Vocology, 34*, 60–66. doi: 10.1080/14015430902839185

Pekkarinen, E., Himberg, L., & Pentti, J. (1992). Prevalence of vocal symptoms among teachers compared with nurses: a questionnaire study. *Scandinavian Journal of Logopedics and Phoniatrics*, *17*, 113–117. doi: 10.3109/14015439209098721

Ramig, L. O., & Verdolini, K. (1998). Treatment efficacy: voice disorders. *Journal of Speech Language and Hearing Research, 41,* 101–116.

Rantala, L. M., Hakala, S., Holmqvist, S., & Sala, E. (2013). Connections between voice ergonomic risk factors in classrooms and teachers' voice production. *Folia Phoniatrica et Logopaedica, 64,* 278–282. doi: 10.1159/000346864

Remacle, A., Petitfils, C., Finck, C., & Morsomme, D. (2017). Description of patients consulting the voice clinic regarding gender, age, occupational status, and diagnosis. *European Archives of Oto-Rhino-Laryngology*, *274*, 1567–1576. doi: 10.1007/s00405-016-4332-7

Rosen, C. A., Lee, A. S., Osborne, J., Zullo, T., & Murry, T. (2004). Development and validation of the voice handicap index-10. *Laryngoscope, 114,* 1549–1556. doi: 0.1097/00005537-200409000-00009

Roy, N., Gray, S. D., Simon, M., Dove, H., Corbin-Lewis, K., & Stemple, J. C. (2001). An evaluation of the effects of two treatment approaches for teachers with voice disorders: a prospective randomized clinical trial. *Journal of speech, language and hearing research, 44,* 286–296. doi: 1092-4388/01/4402-0286

Roy, N., Merrill, R. M., Gray, S. D., & Smith, E. M. (2005). Voice disorders in the general population: prevalence, risk factors, and occupational impact. *The Laryngoscope*, *115*, 1988–1995. doi: 10.1097/01.mlg.0000179174.32345.41

Roy, N., Merrill, R. M., Thibeault, S., Gray, S. D., & Smith, E. M. (2004). Voice disorders in teachers and the general population: effects on work performance, attendance, and future career choices. *Journal of Speech, Language, and Hearing Research, 47*, 542–551. doi: 10.1044/1092-4388(2004/042)

Roy, N., Merrill, R. M., Thibeault, S., Parsa, R. A., Gray, S. D., & Smith, E. M. (2004). Prevalence of voice disorders in teacher and the general population. *Journal of Speech, Language and Hearing Research, 47*, 281–293. doi: 10.1044/1092-4388(2004/023)

Sapir, S., Keidar, A., & Mathers-Schmidt, B. (1993). Vocal attrition in teachers: survey findings. *European Journal of Disorders of Communication*, *28*(2), 177–185. doi: 10.3109/13682829309041465

Seifpanahi, S., Jalaie, S., Nikoo, M. R., & Sobhani-Rad, D. (2015). Translated versions of Voice Handicap Index (VHI)-30 across languages: a systematic review. *Iranian Journal of Public Health*, *44*, 458–469.

Simberg, S., Laine, A., Sala, E., & Rönnemaa, A. M. (2000). Prevalence of voice disorders among future teachers. *Journal of Voice*, *14*, 231–235. doi: 10.1016/S0892-1997(00)80030-2

Simberg, S., Sala, E., & Rönnemaa, A.M. (2004). A comparison of the prevalence of vocal symptoms among teacher students and other university students. *Journal of Voice, 18,* 363–368. doi: 10.1016/j.jvoice.2003.12.005

Simberg, S., Sala, E., Laine, A., & Rönnemaa, A. M. (2001). A fast and easy screening method for voice disorders among teacher students. *Logopedics Phoniatrics Vocology*, *26*, 10–16. doi: 10.1080/14015430119481

Simberg, S., Santtila, P., Soveri, A., Varjonen, M., Sala, E., & Sandnabba, N. K. (2009). Exploring genetic and environmental effects in dysphonia: a twin study. *Journal of Speech, Language and Hearing Reasearch, 52,* 153–163. doi: 10.1044/1092-4388(2008/07-0095)

Siupsinskiene, N. (2002). Importance of voice quality evaluation in the assessment of treatment outcome after endolaryngeal microsurgery. *Medicina (Kaunas, Lithuania), 38*, 1183–1191.

Smith, E., Lemke, J., Taylor, M., Kirchner, H. L., & Hoffman, H. (1998). Frequency of voice problems among teachers and other occupations. *Journal of Voice*, *12*(4), 480–488.

Statistics Solutions, 2018. Retrieved from https://www.statisticssolutions.com/sample-size-formula/

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

Sukanen, O., Sihvo, M., Rorarius, E., Lehtihalmes, M., Autio, V., & Kleemola, L. (2007). Voice activity and participation profile (VAPP) in assessing the effects of voice disorders on patients' quality of life: validity and reliability of the Finnish version of VAPP. *Logopedics Phoniatrics Vocology, 32,* 3–8. doi: 10.1159/000316311

Swets, J. A. (1979). ROC analysis applied to the evaluation of medical imaging techniques. *Investigative Radiology, 14,* 109–121. doi: 10.1097/00004424-197903000-00002

Timmermans, B., De Bodt, M.S., Wuyts, F.L., Boudewijns, G., Clement, G., Peeters, A., & Van de Heyning, P.H. (2002). Poor voice quality in future elite vocal performers and professional voice users. *Journal of Voice, 16,* 372–382. doi: 10.1016/S0892-1997(02)00108-X

Titze, I. R. (1994). Mechanical stress in phonation. *Journal of Voice*, *8*, 99–105. doi: 10.1016/S0892-1997(05)80302-9

Van Houtte, E., Claeys, S., Wuyts, F., & Van Lierde, K. (2011). The Impact of Voice Disorders Among Teachers: Vocal Complaints, Treatment-Seeking Behavior, Knowledge of Vocal Care, and Voice-Related Absenteeism. *Journal of Voice, 25,* 570–575. doi: 10.1016/j.jvoice.2010.04.008

Van Houtte, E., Van Lierde, K., D'haeseleer, E., & Claeys, S. (2010). The prevalence of laryngeal pathology in a treatment-seeking population with dysphonia. *Laryngoscope*, *120*, 306–312. doi: 10.1002/lary.20696

Verdolini, K., & Ramig, L. O. (2001). Review: occupational risks for voice problems. *Logopedics Phoniatrics Vocology, 26*, 37–46. doi: 10.1080/140154301300109125

Verdolini, K., Rosen, C. A., & Branski, R. C. (2006). Classification Manual for Voice Disorders-I.

Verdonck-de Leeuw, I. M., Kuik, D. J., De Bodt, M., Guimaraes, I., Holmberg, E. B., Nawka, T., ... & Woisard, V. (2008). Validation of the voice handicap index by assessing equivalence of European translations. *Folia Phoniatrica et Logopaedica, 60*, 173–178. doi: 10.1159/000200767

Vilkman, E. (2000). Voice problems at work: a challenge for occupational safety and health arrangement. *Folia Phoniatrica et Logopaedica*, *52*, 120–125. doi: 10.1159/000021519

Vilkman, E. (2004). Occupational safety and health aspects of voice and speech professions. *Folia Phoniatrica et Logopedica, 56,* 220–253. doi: 10.1159/000078344

Whitling, S., Rydell, R., & Åhlander, V. L. (2015). Design of a clinical vocal loading test with long-time measurement of voice. *Journal of Voice*, *29*, 261–e13. doi: 10.1016/j.jvoice.2014.07.012

Williams, N. R. (2003). Occupational groups at risk of voice disorders: a review of the literature. *Occupational Medicine*, *53*, 456–460. doi: 10.1093/occmed/kqg113

Winkworth, A., & McCabe, P. (2001). The vocal process of becoming an actor: Vocal risk profiles of first year acting students. Paper presented at the fourth Pan European Voice Conference: Stockholm, Sweden.

Woisard, V., Bodin, S., & Puech, M. (2004). The Voice Handicap Index: impact of the translation in French on the validation. *Revue de laryngologie-otologie-rhinologie, 125*, 307–312.

Zweig, M. H., & Campbell, G. (1993). Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clinical Chemistry, 39*, 561–57.

SCREEN-11

Päivämäärä: _____

Syntymäaika: _____

Sukupuoli: _____

Ammatti: _____

Arvioi, kuinka usein Sinulla on <u>viimeisten kahdentoista kuukauden aikana</u> esiintynyt seuraavia äänioireita, rastittamalla [x] ruutu.

Äänioire	Ei koskaan	Harvoin	Viikoittain	Päivittäin
1. Ääni rasittuu				
2. Ääni väsyy				
3. Ääni käheytyy				
4. Ääni madaltuu puhuessa				
5. Ääni pettää puhuessa				
6. Ääntä on vaikea saada kuuluviin				
7. Puhuessa tulee tarve rykiä				
8. Puhuessa tulee tarve yskiä				
9. Kurkunpään seudulla tuntuu jännittyneisyyttä				
10. Palan tunnetta kurkussa				
11. Kurkunpään seudulla tuntuu kipua				

Foniatrisen tutkimuksen tulokset (lääkäri täyttää):

_____
_____

Voice Handicap Index - VHI   Päivämäärä:_____

On monia sanontoja, jolla ihmiset kuvailevat ääntään ja äänen vaikutusta elämään.
Laita rasti [x] kohtaan, joka kuvaa kuinka usein koet samaa.

Oire			Ei koskaan	Ei juuri koskaan	Joskus	Melkein aina	Aina
1	F1	Ääneni vuoksi ihmisten on vaikea kuulla minua.					
2	P2	Minulta loppuu ilma kun puhun.					
3	F3	Ihmisten on vaikea ymmärtää puhettani meluisassa huoneessa.					
4	P4	Ääneni vaihtelee päivän mittaan.					
5	F5	Perheelläni on vaikeuksia kuulla minua, kun puhun heille asunnon toiselta puolelta.					
6	F6	Käytän puhelinta harvemmin kuin haluaisin.					
7	E7	Olen puhuessani jännittynyt ääneni vuoksi.					
8	F8	Välttelen ryhmätilanteita ääneni vuoksi.					
9	E9	Ääneni tuntuu ärsyttävän ihmisiä.					
10	P10	Minulta kysellään: "Mikä ääntäsi vaivaa?"					
11	F11	Ääneni vuoksi puhun harvemmin ystävien, naapureiden tai sukulaisten kanssa.					
12	F12	Ihmiset pyytävät minua toistamaan, kun puhumme kasvotusten.					
13	P13	Ääneni kuulostaa käheältä tai narisevalta.					

Oire			Ei koskaan	Ei juuri koskaan	Joskus	Melkein aina	Aina
15	E15	Mielestäni muut ihmiset eivät ymmärrä ääniongelmaani.					
16	F16	Ääniongelmani rajoittavat henkilökohtaista ja sosiaalista elämääni.					
17	P17	Ääneni selkeyttä on vaikea ennustaa.					
18	P18	Yritän saada ääneni kuulostamaan erilaiselta.					
19	F19	Tunnen jääväni keskustelujen ulkopuolelle ääneni vuoksi.					
20	P20	Puhuminen on minulle työlästä.					
21	P21	Ääneni on huonompi iltaisin.					
22	F22	Ääniongelmani vuoksi menetän tuloja.					
23	E23	Ääniongelmani huolestuttaa minua.					
24	E24	En ole kovin seurallinen ääniongelmani vuoksi.					
25	E25	Ääneni saa minut tuntemaan itseni vajavaiseksi.					
26	P26	Ääneni pettää kesken puhumisen.					
27	E27	Minua vaivaa, kun ihmiset pyytävät toistamaan.					
28	E28	Kiusaannun, kun ihmiset pyytävät minua toistamaan.					
29	E29	Ääneni vuoksi tunnen itseni kyvyttömäksi.					
30	E30	Häpeän ääniongelmaani.					

Voice Activity and Participation Profile – VAPP

Päivämäärä: _____

POTILAAN OMAT KOKEMUKSET ÄÄNIHÄIRIÖSTÄÄN

Vastaa laittamalla viivaan rasti (X) sille kohdalle, joka vastaa kokemustasi.
Rasti vasemmalla merkitsee, että äänesi on aina hyvä,
rasti oikealla merkitsee, että äänesi on koko ajan huono.

1. Millainen äänesi on nyt?

Normaali ├──┤ Hyvin huono

ÄÄNIHÄIRIÖN VAIKUTUS TYÖNTEKOON

2. Vaikuttaako äänihäiriö työhösi?

Ei ├──┤ Suuressa määrin

3. Oletko ajatellut viimeksi kuluneen puolen vuoden aikana työn vaihtamista
ääniongelmasi takia?

En ├──┤ Koko ajan

4. Lisääkö ääniongelma työpaineitasi?

Ei ├──┤ Koko ajan

5. Onko äänesi huonous vaikuttanut tulevaisuuden urasuunnitelmiisi viimeksi kuluneen puolen vuoden aikana?

Ei ├─────────────────────────────────────┤ Koko ajan

ÄÄNIHÄIRIÖN VAIKUTUS PÄIVITTÄISEEN KOMMUNIKOINTIIN

6. Pyydetäänkö sinua äänesi huonouden takia toistamaan, mitä sanoit?

Ei ├─────────────────────────────────────┤ Koko ajan

7. Oletko kertaakaan viimeksi kuluneen puolen vuoden aikana huonon äänesi takia vältellyt puhumista toisille?

En ├─────────────────────────────────────┤ Hyvin usein

8. Onko ihmisten äänesi huonouden takia vaikea ymmärtää puhettasi puhelimessa?

Ei ├─────────────────────────────────────┤ Hyvin usein

9. Oletko viimeksi kuluneen puolen vuoden aikana äänesi huonouden takia vähentänyt puhelimen käyttöä?

En ├─────────────────────────────────────┤ Hyvin usein

10. Vaikuttaako äänihäiriösi kommunikointiin hiljaisessa ympäristössä?

Ei ├─────────────────────────────────────┤ Suuressa määrin

11. Oletko viimeksi kuluneen puolen vuoden aikana vältellyt äänesi huonouden takia keskustelua hiljaisessa ympäristössä?

En ├───┤ Hyvin usein

12. Vaikuttaako äänihäiriösi kommunikointiisi meluisassa ympäristössä?

Ei ├───┤ Suuressa määrin

13. Oletko viimeksi kuluneen puolen vuoden aikana vältellyt äänesi huonouden takia keskustelua meluisassa ympäristössä?

En ├───┤ Suuressa määrin

14. Vaikuttaako äänihäiriösi viestisi ymmärrettävyyteen ryhmätilanteissa?

Ei ├───┤ Suuressa määrin

15. Oletko viimeksi kuluneen puolen vuoden aikana äänesi huonouden takia vältellyt ryhmäkeskusteluja?

En ├───┤ Suuressa määrin

16. Vaikuttaako ääniongelma viestisi perille menoon?

Ei ├───┤ Suuressa määrin

17. Oletko viimeksi kuluneen puolen vuoden aikana koskaan vältellyt puhumista äänesi huonouden takia?

En ├──────────────────────────────────────┤ Suuressa määrin

ÄÄNIHÄIRIÖN VAIKUTUS SOSIAALISEEN KOMMUNIKOINTIIN

18. Vaikuttaako ääniongelmasi sosiaaliseen aktiivisuuteesi?

Ei ├──────────────────────────────────────┤ Suuressa määrin

19. Oletko viimeksi kuluneen puolen vuoden aikana vältellyt sosiaalista toimintaa äänesi huonouden takia?

En ├──────────────────────────────────────┤ Suuressa määrin

20. Ärsyttääkö ääniongelmasi perhettäsi, ystäviäsi tai työtovereitasi?

Ei ├──────────────────────────────────────┤ Suuressa määrin

21. Oletko viimeksi kuluneen puolen vuoden aikana äänesi huonouden takia vältellyt keskusteluja perheesi, ystäviesi tai työtovereidesi kanssa?

En ├──────────────────────────────────────┤ Suuressa määrin

Alma Zenger   5

ÄÄNIHÄIRIÖIDEN VAIKUTUS TUNTEISIIN

22. Oletko ääniongelman takia hermostunut?

En ├───────────────────────────────┤ Suuressa määrin

23. Oletko ääniongelman takia ahdistunut?

En ├───────────────────────────────┤ Suuressa määrin

24. Onko itsetuntosi ääniongelman takia huono?

Ei ├───────────────────────────────┤ Suuressa määrin

25. Oletko ääniongelmasi takia huolestunut?

En ├───────────────────────────────┤ Suuressa määrin

26. Oletko ääniongelman takia tyytymätön?

En ├───────────────────────────────┤ Suuressa määrin

27. Vaikuttaako ääniongelmasi persoonallisuuteesi?

Ei ├───────────────────────────────┤ Suuressa määrin

28. Vaikuttaako ääniongelma kuvaan, joka sinulla on itsestäsi?

En ├───────────────────────────────┤ Suuressa määrin

APPENDIX D

Pyyntö ja tiedote tutkimukseen osallistuvalle

Sinua pyydetään mukaan tutkimukseen, jonka tarkoituksena on kehittää uusi äänioirekysely, jota tulevaisuudessa voitaisiin käyttää luotettavana äänihäiriöiden kartoitusmenetelmänä. Aloitteen tutkimukselle on tehnyt Åbo Akademin logopedian yksikkö. Kyselylomakkeen vahvistamiseksi on varmistettava sen luotettavuus, johon tarvitsisimme Sinun apuasi. Osallistuminen on täysin vapaaehtoista, eikä vaadi muuta kuin ohessa tulleiden lomakkeiden täyttämisen. **Lomakkeet tulisi täyttää huolellisesti heti ne saatuasi ja täytön jälkeen postittaa takaisin mukana tulleessa palautuskuoressa**. Vastaanotolle tultuasi saat täyttää yhden lomakkeista uudelleen. Lomakkeiden täyttämisen tarkoituksena on varmistaa, että niissä olevat väittämät mittaavat olennaisia osa-alueita äänioireista.

Tämä tutkimus toteutetaan yhteistyössä Turun yliopistollisen keskussairaalan kanssa. Olemme lähettäneet tämän osallistumispyynnön henkilöille, jotka ovat foniatrisen poliklinikan jonotuslistalla. Kun lomakkeet on täytetty ja palautettu, niihin liitetään foniatrisen tutkimuksen tulokset. Henkilötiedot muunnetaan nimettömiksi koodeiksi ja toimitetaan Åbo Akademin logopedian yksikölle, jossa tutkimusta kehitetään. Henkilötietoja ei siis luovuteta sairaalan ulkopuolelle missään vaiheessa, ja niitä säilytetään tietosuojalain mukaisissa olosuhteissa ja käsitellään luottamuksellisesti.

Osallistuvan tulisi siis täyttää ohessa tulleet lomakkeet sekä tulla foniatrin vastaanotolle vastaanottokutsun saapuessa. Tutkimuksesta ei koidu Sinulle minkäänlaisia taloudellisia kuluja tai terveydellisiä vaikutuksia. Tutkimukseen osallistumisen voi myös tarvittaessa keskeyttää milloin vain.

Olisimme kiitollisia, jos osallistuisit tutkimukseemme. Jokaisen osallistuvan kesken arvotaan elokuvalippuja. Jos haluat osallistua arvontaan, muista täyttää sähköpostiosoitteesi suostumusasiakirjaan.

Halutessasi saat lisätietoa tutkimuksesta vastaavilta henkilöiltä:

Alma Zenger, Åbo Akademin logopedian yksikkö, sähköposti: azenger@abo.fi
Daniel Fellman, Åbo Akademin psykologian yksikkö, sähköposti: dfellman@abo.fi
Sofia Holmqvist Jämsén, Åbo Akademin logopedian yksikkö, sähköposti: soholmqv@abo.fi

Suostumus tutkimukseen osallistumisesta

Minua on pyydetty osallistumaan tutkimukseen *Retrospektiivisen äänioirekyselyn validointi*.

Olen tietoinen siitä, että täyttämäni lomakkeet säilytetään tietosuojaturvalain mukaisissa olosuhteissa ja niitä käsitellään luottamuksellisesti. Henkilötietojani ei luovuteta missään vaiheessa Tyksin ulkopuolelle, vaan tiedot muunnetaan ennen Åbo Akademin tutkijoille lähettämistä tunnistekoodeiksi, joista ei voi päätellä henkilöllisyyttäni.

Olen perehtynyt edellä olevaan selvitykseen ja saanut riittävästi tietoa tutkimuksesta ja sen yhteydessä suoritettavasta tietojen keräämisestä, käsittelystä ja luovuttamisesta. Minulle on selvitetty, että tarvittaessa voin pyytää lisätietoja tiedotteessa nimetyltä tutkijalta. Suostun osallistumaan tutkimukseen ja teen sen vapaaehtoisesti. Allekirjoituksellani vahvistan, että osallistun tässä asiakirjassa kuvattuun tutkimukseen ja suostun luovuttamaan kyselylomakkeet tutkijoille. Suostumus allekirjoitetaan kaksin kappalein niin, että yksi jää tutkijalle ja toinen tutkittavalle.

Annan suostumukseni osallistumiselleni tähän tutkimukseen.

□ kyllä

□ ei

_____
Tutkittavan allekirjoitus ja nimen selvennys

_____
Paikka ja aika

_____
Suorittavan tutkijan allekirjoitus ja nimen selvennys

□ Haluan osallistua elokuvalippujen arvontaan.
Sähköpostiosoitteeni on: _____

□ En halua osallistua elokuvalippujen arvontaan.

Alma Zenger

PRESS RELEASE - PRESSMEDDELANDE

Validering av ett snabbt och enkelt screeningformulär för röststörningar

Pro gradu-avhandling i logopedi
Fakulteten för humaniora, teologi och psykologi, Åbo Akademi

Resultaten i en pro gradu-avhandling vid Åbo Akademi visar att en lyckad validering av ett potentiellt framtida screeningverktyg för röststörningar har utförts. Alma Zenger har undersökt det nya röstrelaterade frågeformuläret med namnet Screen11. Namnet på screeningformuläret hänvisar till de elva röstrelaterade påståendena som formuläret består av. De specifika syftena för valideringen var att undersöka den interna reliabiliteten mellan de elva påståendena, den konvergenta validiteten genom att undersöka interkorrelationerna mellan Screen11 och två andra väletablerade röstformulär samt att bestämma det gränsvärde som bäst åtskiljer röstfriska personer från personer som eventuellt har en röststörning.

Resultaten visade att Screen11 innehade en god intern reliabilitet. Dessutom tydde korrelationerna mellan Screen11 och de två andra röstrelaterade formulären på att den konvergenta validiteten var stark. Fyra eller fler återkommande röstsymtom i veckan visade sig vara mest optimala gränsvärdet för eventuell röststörning.

Den lyckade valideringen av Screen11 betyder att den i framtiden kunde användas som ett snabbt och enkelt screeningformulär för röststörningar.

Studien förverkligades i samarbete med Åbo Universitets Centralsjukhus foniatriska poliklinik. Ett antal på 24 personer som sökte hjälp för sina röstproblem rekryterades via polikliniken. De hjälpsökande deltagarna matchades genom ett bekvämlighetsurval av 24 röstfriska personer enligt ålder, kön och yrke. Data samlades in med röstrelaterade frågeformulär.

Trots en lyckad validering av Screen11 var sampelstorleken i studien liten. Därför kommer en replikering av valideringen med ett större sampel att fullföljas.

Alma Zenger
050 491 2913
alma.zenger@abo.fi

Alma Zenger

Sofia Holmqvist-Jämsén
Doktor i logopedi, legitimerad talterapeut
046 921 6756
sofia.holmqvist@abo.fi

Daniel Fellman
Doktor i psykologi, legitimerad talterapeut
050 326 7739
daniel.fellman@abo.fi