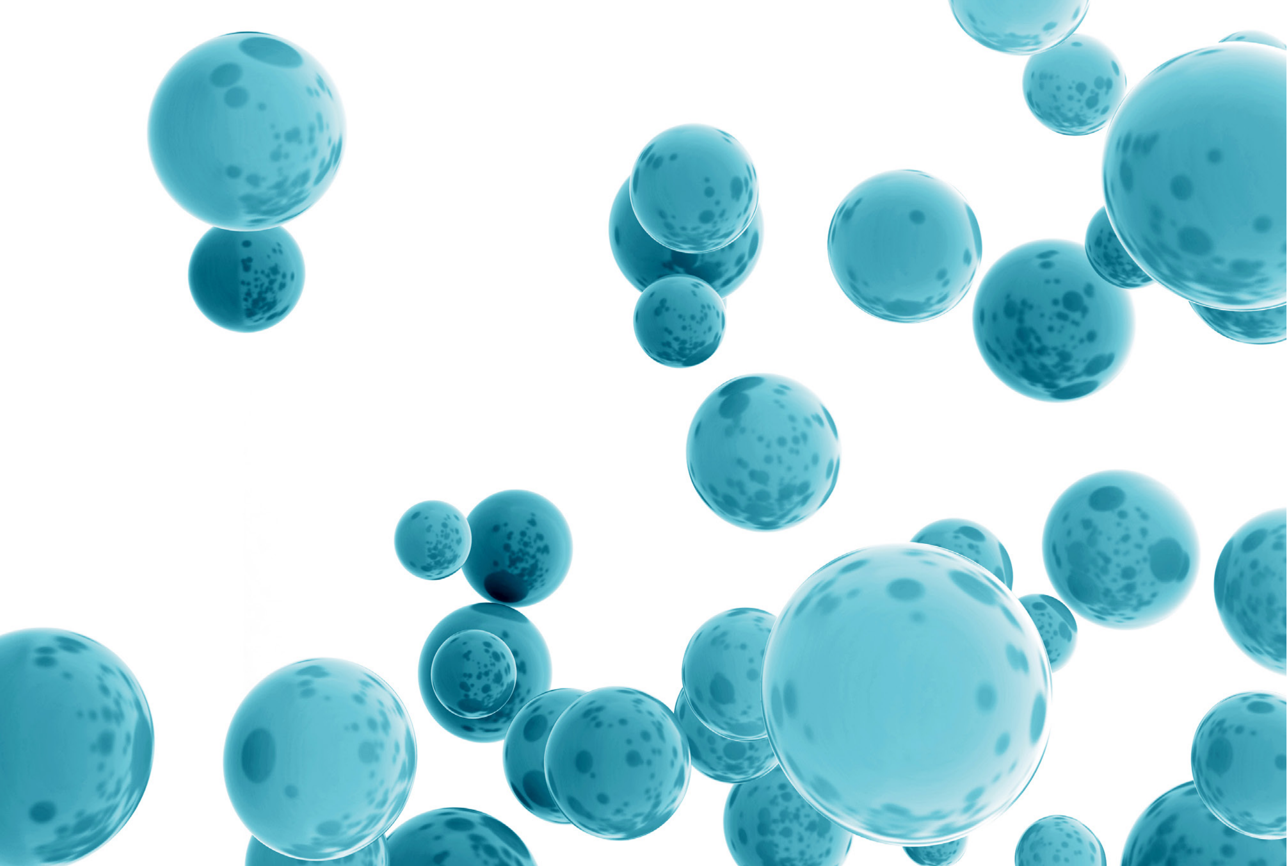
Acta Universitatis Tamperensis 2274

KATI JÄRVINEN

Voice Characteristics in Speaking a Foreign Language

A study of voice in Finnish and English as L1 and L2



KATI JÄRVINEN

Voice Characteristics in Speaking a Foreign Language

A study of voice in Finnish and English as L1 and L2

■

ACADEMIC DISSERTATION
To be presented, with the permission of
the Faculty council of the Faculty of Education
of the University of Tampere,
for public discussion in the lecture hall Linna K 103,
Kalevantie 5, Tampere,
on 12 May 2017, at 12 o'clock.

UNIVERSITY OF TAMPERE

KATI JÄRVINEN

Voice Characteristics in Speaking a Foreign Language

A study of voice in Finnish and English as L1 and L2

Acta Universitatis Tamperensis 2274
Tampere University Press
Tampere 2017

UNIVERSITY
OF TAMPERE

ACADEMIC DISSERTATION
University of Tampere
Faculty of Education
Finland

The originality of this thesis has been checked using the Turnitin OriginalityCheck service in accordance with the quality management system of the University of Tampere.

Cover design by
Mikko Reinikka

Acta Universitatis Tamperensis 2274
ISBN 978-952-03-0423-2 (print)
ISSN-L 1455-1616
ISSN 1455-1616

Acta Electronica Universitatis Tamperensis 1776
ISBN 978-952-03-0424-9 (pdf)
ISSN 1456-954X
http://tampub.uta.fi

Suomen Yliopistopaino Oy – Juvenes Print
Tampere 2017

To Vili, Aku and Onni

"You must do the things you think you cannot do."

—Eleanor Roosevelt

PREFACE

I have always been interested in languages. I have started my studies at the Tampere University in Nordic Philology. I think I have been aware that something changes when I speak another language than Finnish, but I have not been able to pinpoint the changes. Whenever I discussed the topic of my research with my friends and acquaintances, they used to reply that they had noticed something changing as well, but they were not able to locate the changes.

When I studied at the department of Speech Communication and Voice Research at the University of Tampere I finally understood that voice research was the answer to my questions. And this is how my research began. It has been a journey to voice, vocology, languages, phonetics, and me as a researcher. A journey not always smooth. I may have sometimes tripped along the way, but I have always been able to get back up.

World known soprano Karita Mattila has said: “I trust the original language. I haven´t sung translations since my debuting times. I agree with Esa-Pekka Salonen, as he says that different languages have different temperaments. As I change from one language to another, my voice changes, as well as my whole personality. (Helsingin Sanomat 6.12.2003).” This has been an inspiration throughout the years of this research.

ACKNOWLEDGEMENTS

First, I want to thank Professor Anne-Maria Laukkanen whose tireless guidance and encouragement over the years has been basis for the research, first by raising the inspiration for voice research and then helping me make my research move forward. I thank, also, Professor Olli Aaltonen for his precious comments and guidance over the years throughout the process. I owe you both my warmest gratitude.

I wish to thank Professor Paavo Alku and Professor Erkki Vilkman for their valuable comments and advice as pre-examiners of my work. Also, I want to thank MD Ahmed Geneid who taught me how to use TKK Aparat for inverse filtering and gave his persistent help in finishing Article III, and PhD Michael O´Dell for his teamwork and untiring guidance in the field of phonetics. I also want to thank PhD Irma Ilomäki as well as all the teachers in former Department of Speech Communication and Voice Research who were responsible that I became interested in voice research and speech communication in the first place. Also, special thanks I owe to special laboratory technician Jarmo Helin, who made the recording facilities usable and always had time to share a joke in the lab. MSc Jyrki Ollikainen helped with statistical analyses in Study III, for which I owe him a warm thanks. MA Virginia Mattila has checked the language of the English questionnaire, thank you.

I want to thank my colleagues PhD Elina Kankare, MA Päivi Lukkarila, MA, PhD student Jaana Tyrmi, and PhD Teija Waaramaa. They gave their opinions on my research, and noticed usually things that I would probably have overlooked. And they were an important support group whenever I needed to talk.

Also, I want to thank every person who participated in my study. The people who gave their voices, and the people who participated in the perceptual analyses. There is only so much a researcher can do by herself.

I want to thank University of Tampere, Kansan Sivistysrahasto (People´s Cultural Fund), Tampereen kaupungin tiederahasto (the scientific fund of city of Tampere), and Emil Aaltosen säätiö (Emil Aaltonen Foundation) for the financial support over the years.

My deepest gratitude I owe to my mother, Marjatta Rintala. She was the most important person for making the research possible, she took care of my children

when I needed time for myself or for my studies. She never had any doubt even when I did.

And last, but not least my dearest, Vili, Aku and Onni. Your mother is now done. Thank you for keeping my priorities straight, and for being so understanding, though there were days I forgot to make lunch or dinner, or both. As Moominpappa has said: "Life becomes unbearable if we don´t know whether it´s Sunday or Wednesday. One can´t live like that. "

Tampere 14.3.2017 *Kati Järvinen*

ABSTRACT

In multicultural societies and international connections, it is essential that people can speak foreign languages being able to communicate with people from other language and cultural backgrounds. However, how the voice or use of voice changes between native and foreign languages has not been previously vastly studied. This research aims to clarify how vocal parameters change when speaking a foreign language compared to speaking the native one. The study consists of acoustical and perceptual analyses of voice and subjects´ subjective notions of how the voice and voice use change in the language shift and, also, how it affects vocal fatigue. The subjective notions have been gathered by a questionnaire.

This research consists of three studies. First of the studies concentrated on fundamental frequency changes, the second study on vocal fatigue and vocal loading in native and foreign languages by subjective notions and analyses of the long term average spectrum and measures of vocal doses. The third study attends to changes in voice quality between native and foreign languages by inverse filtering and perceptual analyses.

Twenty native Finnish and 23 native English speakers participated in the research. The foreign languages they used were either English or Finnish. In Study I, sixteen Finnish and 14 English subjects were analyzed, in Study II twenty Finnish and 23 English. At the time, the results of Study I were presented, thirty subjects in total had been recorded, and the additional thirteen subjects came in later. To the Study III 24 subjects were chosen. Twelve of them reported more vocal fatigue and 12 who did not report to feel more vocal loading in the foreign language than the native one.

The speech samples were recorded in a well damped studio at the University of Tampere. The subjects read a text in the native and then in the foreign language. Spontaneous speech samples were also recorded. For spontaneous speech samples, the subjects were given a comic strip which they described first in the native language and then in the foreign one. The speech samples were calibrated for the analyses of sound pressure level.

According to the results, voice and voice use are affected by the language shift from the native language to the foreign one. Fundamental frequency was higher in

the foreign language (Study I), especially this was the case with the Finnish subjects, the English subjects change was not as clear. Almost eighty percent of the subjects considered that speaking a foreign language changes their voices, but acoustically measured rise in the mean fundamental frequency did not necessarily actualize in the subjective notion of rising of fundamental frequency. Majority of the subjects (70 %) reported that the voice tires faster when speaking the foreign language. Vocal dose measurements (Study II) did not, however, show unambiguously that speaking the foreign language was more loading. This is due to the fact that these dose measurements base on fundamental frequency, sound pressure level and the amount of voiced speech. When speaking a foreign language, the amount of the voiced speech tends to lower compared to that of the native language, and therefore, the doses are not sufficient as such in analyzing vocal loading in two languages. However, when the doses were normalized, that is calculated the doses per second, a trend that speaking a foreign language can be more loading than speaking the native one was found. Also, the acoustical analyses based on the long term average spectrum showed a similar kind of trend. In the study of voice quality (Study III) it was found that the subjects who reported more vocal fatigue were perceptually evaluated to have poorer voice quality, more pressed phonation and higher pitch in both native and foreign languages than the subjects who did not report more vocal fatigue in the foreign language. The changes obtained by inverse filtering showed that voice production was more pressed in the foreign language than in the native one, which may cause vocal overloading.

Vocal doses are not enough but voice quality should be added in the research of vocal fatigue and vocal loading. Also, people who need to use foreign languages extensively would benefit from training in speech technique in the foreign language, since even when the voice use was optimal in the native language it does not necessarily shift to voice use in the foreign one.

TIIVISTELMÄ

Kansainvälistyvässä maailmassa ja monikulttuurisissa yhteiskunnissa on vieraiden kielten osaamisella merkittävä tehtävä ihmisten välisessä kommunikoinnissa. Kuitenkaan sitä, miten vieraalla kielellä puhuminen muuttaa puhujan ääntä ja tapaa käyttää ääntä, ei ole juurikaan aiemmin tutkittu. Tämän tutkimuksen tarkoituksena on selvittää, miten äänen eri parametrit muuttuvat puhuttaessa vierasta kieltä verrattuna äidinkielellä puhumiseen. Tutkimus koostuu äänen akustisista ja kuulohavaintoon perustuvista analyyseistä sekä koehenkilöiden subjektiivisista näkemyksistä kielen vaihtamisen vaikutuksesta ääneen ja äänenkäyttöön sekä äänen väsymiseen. Subjektiiviset havainnot on kerätty kyselylomakkeella.

Tutkimus koostuu kolmesta osatutkimuksesta, joista ensimmäinen keskittyi perustaajuuden muutoksiin. Toisessa osatutkimuksessa on tutkittu äänen väsymistä ja kuormitusta äidinkielen ja vieraan kielen välillä subjektiivisten itsearviointien sekä keskiarvospektrin ja kuormituksen arvioinnissa käytettävien annosmittausten avulla. Kolmannessa osatutkimuksessa on tutkittu äänen laadun muutoksia äidinkielen ja vieraan kielen välillä sekä akustisesti käänteissuodattamalla että kuulohavaintoon perustuen.

Koehenkilöinä tutkimuksessa oli kaikkiaan 20 suomea sekä 23 englantia äidinkielenään puhuvaa henkilöä. Heidän käyttämänsä vieraat kielet olivat englanti ja suomi. Osatutkimuksessa I koehenkilöinä oli 16 suomea ja 14 englantia äidinkielenään puhuvaa ja osatutkimuksessa II 20 suomea ja 23 englantia puhuvaa, yhteensä 23 henkilöä. Osatutkimuksen I tulosten julkaisemisen jälkeen tutkimukseen saatiin vielä kolmetoista koehenkilöä lisää. Osatutkimukseen III valittiin osatutkimuksen II perusteella 24 koehenkilöä. Koehenkilöistä 12 vastasi kyselylomakkeessa, että eivät kokeneet äänensä väsyvän nopeammin puhuessaan vierasta kieltä ja 12 vastasi äänensä väsyvän nopeammin puhuessaan vieraalla kielellä kuin äidinkielellä.

Koehenkilöiden ääninäytteet äänitettiin äänieristetyssä äänitysstudiossa Tampereen yliopistossa. Koehenkilöt lukivat kaksi tekstiä, ensin äidinkielellään ja sen jälkeen vieraalla kielellä. Lisäksi heille annettiin spontaanipuheen äänitystä varten sarjakuva, jonka sisällön he selittivät ensin äidinkielellään ja sen jälkeen vieraalla kielellä. Näytteet kalibroitiin voimakkuuden analysointia varten.

Tutkimuksen tulosten perusteella puhujien ääni ja äänenkäyttö muuttuvat puhuttaessa vierasta kieltä verrattuna äidinkielellä puhumiseen. Perustaajuus nousi etenkin suomalaisilla koehenkilöillä, kun he puhuivat englantia verrattuna suomen puhumiseen (Osatutkimus I). Myös englantilaisilla koehenkilöillä oli nähtävissä muutoksia perustaajuudessa, mutta ne eivät olleet yhtä selkeitä kuin suomalaisilla. Lähes 80 % koehenkilöistä koki, että vieraalla kielellä puhuminen muuttaa ääntä, akustisesti mitattu perustaajuuden nousu ei kuitenkaan välttämättä ollut yhteydessä subjektiiviseen käsitykseen sävelkorkeuden muutoksessa. Suurin osa (70 %) koehenkilöistä koki, että ääni väsyy nopeammin vierasta kieltä puhuttaessa. Äänen annosmittaukset (Osatutkimus II) eivät kuitenkaan osoittaneet yksiselitteisesti, että vieraalla kielellä puhuminen olisi kuormittavampaa. Tämä johtuu siitä, että mittaukset käyttävät perustanaan perustaajuutta, voimakkuutta ja soinnillisen äännön määrää kokonaispuheessa. Vierasta kieltä puhuttaessa soinnillisuuden määrä laskee verrattuna äidinkieliseen puhumiseen, joten annosmittaukset eivät sovellu sellaisenaan kovin hyvin äänen kuormituksen analysointiin kahden kielen välillä. Normalisoituina, so. annos laskettuna sekunnissa, ne kuitenkin osoittivat, että oli löydettävissä trendi, jonka perusteella voidaan sanoa äänen kuormittumisen lisääntyvän vieraalla kielellä puhuttaessa. Myös akustiset spektriin liittyvät analyysit osoittivat samankaltaisen trendin. Äänen laadun tutkimuksessa (Osatutkimus III) havaittiin eroja äänen laadun kuulonvaraisissa arvioissa koehenkilöryhmien välillä. Henkilöillä, jotka kokivat äänensä väsyvän nopeammin vieraalla kielellä kuin äidinkielellä, äänen laatu arvioitiin vähemmän hyväksi, puristeisemmin tuotetuksi ja sävelkorkeudeltaan korkeammaksi kuin niillä henkilöillä, jotka eivät raportoineet äänen väsymisen lisääntymisestä vieraassa kielessä. Käänteissuodatuksella saadut äänen laadulliset muutokset osoittivat, että vierasta kieltä puhuttiin puristeisemmin kuin äidinkieltä, mikä voi aiheuttaa äänen kuormittumista.

Äänen laatu on syytä ottaa huomioon tutkittaessa äänen väsymistä ja kuormitusta. Lisäksi henkilöt, jotka joutuvat käyttämään runsaasti vierasta kieltä, hyötyisivät puhetekniikan opetuksesta myös vieraalla kielellä, koska optimaalinenkaan äänenkäyttö äidinkielessä ei välttämättä siirry vieraaseen kieleen.

CONTENTS

LIST OF ORIGINAL PUBLICATIONS

I Järvinen K, Laukkanen AM, Aaltonen O. Speaking a foreign language and its effect on F0. Log Phon Voc 2013;38:47–51.

II Järvinen K, Laukkanen AM. Vocal Loading in Speaking a Foreign Language. Folia Phoniatr Logop 2015;67:1–7.

III Järvinen K, Laukkanen AM, Geneid A. Voice quality in native and foreign languages investigated by perceptual analyses and inverse filtering. J Voice 2016, in press.

ABBREVIATIONS AND TERMS

Alpha ratio	The ratio of spectral energy below and above a certain frequency limit. The frequency limit is typically set to 1000 Hz but this study used 1500 Hz.
AQ	Amplitude quotient
ClQ	Closing quotient
D_c	Cycle dose
D_d	Distance dose
D_e	Energy dissipation dose
D_r	Radiated energy dose
D_t	Time dose
L1	Native language
L2	Foreign language
Formant	Vocal tract resonance
F0	Fundamental frequency
Hyperfunctional	Pressed phonation, too firm adduction of the vocal folds
Hypofunctional	Breathy phonation, too loose adduction of the vocal folds
IAIF	Iterative Adaptive Inverse Filtering
IQR	Interquartile range
Leq	Equivalent sound level
LTAS	Long-time average spectrum
L1-L0 level difference	The level difference of the F0 region (0-300Hz) in relation to F1 region (300-1200 Hz)
NAQ	Normalized amplitude quotient
OQ	Open quotient
Perceptual	Based on auditory analyses
SPL	Sound pressure level
SQ	Speed quotient

VLI	Vocal Loading Index
Vocal doses	Doses based on F0, SPL, and amount of voiced speech that can be used for quantifying exposure to loading of the vocal fold tissue
Vocal fatigue	In here, subjective sensation of tiredness of the voice
Vocal (over)loading	A combination of prolonged voice use and loading factors causing excessive mechanical stress on vocal fold tissue during phonation

LIST OF FIGURES AND TABLES

1 INTRODUCTION

In the modern and global world, it is necessary for people to be able to communicate with people from other cultural and language backgrounds. Especially people from a relatively small language environment, such as Finnish, need to learn other languages besides the native one in order to be able to interact in international relations. Also, in the multicultural societies it is important for immigrants to learn the language of the society. In modern work life, it has also become more usual than previously that workers need to have language skills, and skills in foreign languages are not considered as special skills anymore, but more as being a part of professional skills. (1.)

It is possible that speaking a language other than the native one affects the speaker´s voice production. For instance, people who need to use foreign languages to a large extent, e.g. professional interpreters and simultaneous translators, have been reported to have symptoms of vocal fatigue and mental tiredness after prolonged use of the foreign language (2,3), yet there is little, if any, voice training obligatory in their studies and/or in profession, see e.g. curricula guides in Multilingual Communication and Translation Studies (4,5). This notion set the idea that it would be important to investigate why people´s voices get problems when speaking a language other than the native one.

Knowing that interpreters have voice problems, and (at least in Finland) a certified business interpreter must master professional speech communication skills, which includes correct, clear and fluent speech, and correct, clear and flawless articulation as well as to be able to build trust and confidence through communication and presentation skills (6). Also, having to deal with multilingual and multicultural people, interpreters should have training in communication and cultures (7) which, then, should include training in voice and speech technique.

However, research on voice in a language other than the native one has not been carried out as vastly as one would expect. Majority of the research on the use of a foreign language has concentrated on the accuracy of articulation, or the influence of the native language on the production of the target language, and on the intelligibility of the speaker (8-13) or perception of the vowels in the target language (9,11,14-18). In most of these studies the target language has been English. Usually these studies are not made with the same person speaking two different languages, and, thus, the studies have not usually compared the production of the native language and second language and therefore vocal characteristics have not been investigated in both languages for the same speaker. However, some contrastive research on same speaker has been done, e.g. Morrison (19) has studied acoustic characteristics in L1 and L2 focusing on the production and perception of vowels /i/and /ɪ/ of Spanish and (Canadian) English, and Rauber et al. (20) have studied vowels in English and Brazilian Portuguese. According to the results language learners discriminate different English vowels with varying levels of accuracy, and a relationship between perception and production was found. Some contrastive research comparing Finnish and English has been done (e.g. 21-24).

This research aims to study what changes language shift from native language (L1) to foreign language (L2) causes in terms of voice production and voice use. The main focus is not to study accuracy of the spoken language as such, while it is touched on briefly. The research focuses mainly on vocal characteristics between L1 and L2 as well as the subjects′ subjective notions of the possible changes between the languages. Also, vocal loading and vocal fatigue in L2 compared to that of L1 are studied.

Interpreters was the starting point for the research, but it became obvious that it would be beneficial to study people with different levels of language skills. Therefore, the decision to study "ordinary" people was made, not only people with advanced skills in foreign languages, such as interpreters. Subjects were recruited to participate in the research mostly from University of Tampere, some were university lecturers, and some were students. Overall, a wonderful and diverse group of people participated in the research.

The text materials for both Finnish and English reading samples was chosen by the researcher, the traditional texts were not used, since it was essential to have two texts with the same content in both languages. The vowel-consonant ratio of the texts was satisfactory: 0.93 in the Finnish text and 0.61 in the English. According to Hakulinen (25), in Finnish the text based ratio is 0.92 and in English 0.63. Also, the same criteria applies for the spontaneous speech task: it was necessary to have a task where the subjects would use similar kind of language in both L1 and L2, and so a comic strip was chosen. The comic strip used here has been used previously broadly in voice therapy. Then a questionnaire for the subjects was planned. The main aim with the questionnaire was to investigate the subjects′ subjective views on their voices in L1 and L2 and whether they had noticed any changes in the language shift. Also, subjective evaluations on vocal endurance in L1 and L2 was included in the questionnaire.

The first recording was made by special laboratory technician Jarmo Helin, the researcher was in charge of the recordings of 42 subjects. The subjects filled in the questionnaire after the recordings at Tampere University.

The first study was done at the time when thirty subjects had been recorded, and it was on fundamental frequency and its standard deviation and range in native and foreign languages. This was a clear continuum after the article (26) where it was discovered that the fundamental frequency is affected by the language shift. The results of the first study were presented at PEVOC in Marseilles in 2011, and published in Logopedics Phoniatrics and Vocology in 2013.

Next a study of other parameters alongside with the fundamental frequency, such as intensity and spectral characteristics was carried out. The results were presented in Tallinn, Estonia at Phonetics Symposium in 2012, and the results have been published in the Proceedings of the symposium (27). This study is not included in the thesis, but nevertheless, it was an important part of the process. Also, vowels and formant shifts in language shift have been studied, and the preliminary results have been presented in Phonetics Symposium in Espoo in 2015 (28).

It is known that there are some factors which increase the loading of the voice, such as rise in fundamental frequency, increase in intensity, and hyperfunctional voice production (29,30). The results seemed to show independent of the study that some vocal characteristics are affected by the language shift, so a question arose whether it makes the voice production more loading. Also, it was interesting to study if the subjects considered their voices to getting more tired in L2 than in L1. The vocal doses (31,32) were included in the research for quantifying vocal loading. The results on vocal loading were presented in Folia Phoniatrica et Logopaedica in 2015. However, it seemed that the vocal doses were not all that suitable for quantifying vocal loading in two different languages, so another way of measuring it had to be studied and voice quality seemed to be the suitable parameter. And, at this point it was clear that the voice, going through the vocal tract, is affected by the language shift from native language to the foreign language. This lead to the final question, whether the language shift affects the voice source.

Since the subjective notions on vocal fatigue showed that some of the subjects reported more vocal fatigue than others, the change in vocal characteristics to possibly explain this was studied by inverse filtering, so the voice generated at the vocal folds could be analyzed. Also, as voice quality became an issue, perceptual analyses were conducted. The results of voice quality through inverse filtering and perceptual analyses have been accepted for publication in Journal of Voice in 2016.

The articles have been published in peer-reviewed publications. The researcher has been the first contributor in them, preparing the articles, and also making the corrections presented by the reviewers under supervision by Professor Anne-Maria Laukkanen and Professor Olli Aaltonen. The researcher has conducted the acoustical and statistical analyses, with help from Anne-Maria Laukkanen and MSc Jyrki Ollikainen.

2 LITERATURE

Many species communicate by voice, that is, by the airflow from the lungs through larynx and the vocal tract, but only humans have the ability to speak. (33.)

Speech and voice belong together, although they are not synonymous. Voice acts as a carrier for all voiced speech, and creates an independent source of expression with its broad variation possibilities. Speech is realization of a linguistic code and a medium for social interaction between people, but it is also a biological code as it gives out information about the speaker. The ability to speak makes the human acoustic communication unique, and speech enables the existence and survival of verbal languages. With the aid of vocal tract structures, e.g. lips, tongue, velum etc., it is possible to alter the size and shape of the vocal tract cavities and thereby vary the acoustic properties of it or to constrict the vocal tract to a varying degree to produce a second sound source in addition to vocal folds. This production of the speech sounds is called articulation.

Acoustically the speech sounds are formed by acoustical phenomena (such as formants, time-varying succession of sound and silence etc.) and their combinations. These are linked to the functions of the speech organs. The speech sounds can vary (depending on the speaker, the surrounding speech sounds, and speech tempo, for example) but speakers of a particular language can still recognize them as realizations of the same phonemes of the language. The special characteristic of speech is that a certain number of meaningless and constant phonemes can be modified into limitless number of combinations that have mutually agreed meanings in a given speech community, thereby creating the realization of a language. Language, then, can be defined as a social construct tightly connected to speech as a medium for communication. (33,34.)

2.1 Speech

Speech is primarily a dynamic action of the speech organs that results in speech signal. Speech consists of acoustical phenomena such as periodic and non-periodic signal (e.g. friction noise for example due to aspiration, fricatives and phonation type), formants, silent parts, and their combinations. We use our speech organs to speak, speech is received by our ears, and the perceived speech is processed in our brain. All of these organs have been shaped to serve one another, and finally our social behavior. (33.) It is likely that the human evolution has to a certain extent followed the evolution of speech, and the early elements of speech have been interjections and descriptive-onomatopoetic words which are based on speech sounds. Development of speech has been one major premise for human evolution. (35.)

Speech can be divided into communicative and informative speech, the first being significant for the sender, while the latter induces a change in the receiver being therefore significant to him/her. What the speaker says is communicative, the way the speaker speaks is informative. (36.)

For people to communicate visual and acoustic channels are used. The visual cues, also known as extra-linguistic features, are e.g. facial expressions, gestures and movements, and acoustical cues, which communicate paralinguistic information with intonation, accentuation and timbre. Some of the paralinguistic communication may be intentional and addressed to the listener in the context, such as variations in pitch, intensity and timbre, but occasionally these cues can be unintentionally signaled being a reaction to the speaker´s emotional state. (37.)

2.2 Language

The origin of language is possibly a result of biological processes, which have lead the necessary organs for speech to form. The upcoming of a language requires the ability of abstract thinking, and the actual breakthrough for the human language has occurred as voice as a communicative medium has become more essential than the visual messages. Also, the phoneme system has then evolved, and the vocal tract has changed. (33,35,38.)

The definition of language as having an acoustic channel for communication excludes sign language as a language, although it is the first and spontaneous native language for a large number of people. Sign language does not have an acoustical channel but it is a language altogether. (33,35). Then it would be reasonable to state that language is a medium for communication between people.

The phonology of languages is as much a result of evolution as is the ability to speech. The phonemes can vary depending on many reasons, such as differences between speakers, phoneme interaction, and varying speech tempo, but people who speak the same language still recognize them as one and the same phoneme in the native language. (33.)

Language has two basic functions: it enables thoughts to be symbolized out loud and thus communication (39), and it gives out and receives information (40). All languages that exist are the medium for communication and they fill social tasks in the society. If the needs in the society are changing, the language changes along with them. Languages, then, are constantly changing. (41.) According to Crystal (42), the functions of a language is to communicate ideas, and to exchange the facts and opinions to other people. It allows people to express emotions, and it works as the instrument of thought. Without language it is not possible to think rationally. However, this strong relation between language and thought has been criticized, since, while language is a powerful tool in forming thoughts, it does not entirely determine a person´s thinking (see e.g. 43,44).

Languages can be classified by typological or genealogical classifications. In typological classification languages are compared by semantics, phonology, morphology, and syntax. Genealogical classification explains the relationships between languages where the relationship is often manifested in similarities in vocabulary and structure. (35,45.) The two languages used in this research have no relationship as Finnish is a Uralic, Finnish-Ugric language, and English is an Indo-European, Germanic language (45), and they differ vastly from each other also in typology.

English was chosen to the study with Finnish, since the first foreign language that Finnish school children usually start studying is English: in 2009 ninety percent of children in grammar school chose English as the first foreign language (46). Also, in international companies in Finland the working language is mostly English, and there all official interaction is performed in English (1).

Languages may differ from each other in terms of vocal characteristics, such as fundamental frequency and voice quality. Also, languages have different formant frequencies of speech sounds, which affect the spectra, and that has an effect on vocal timbre (47-55). For example, in Finland a relatively low pitch is considered favourable whereas in Britain a higher pitch has traditionally been attributed to a high status in the society (56,57).

Some major differences between Finnish and English can be seen e.g. at the level of phonemes, articulation, and intonation. For example, Finnish /y/ does not exist in English, and Finnish does not have voiced affricate /dʒ/ such as English. Also, in articulation Finnish /r/ is a voiced tremulant, while in English, it is often realized as an approximant /ɹ/. Finnish also has short/long vowel opposition, while in English there is a quality opposition in vowels to distinguish meaning. Finnish does not have aspiration of consonants as a distinctive feature, such as English. (For differences in vowels, see vowel formant charts in chapter 2.4.3, page 32). In the Finnish language questions are marked mainly by interrogatives, while in English questions are marked by intonation. In the Finnish intonation, the F0 contour has the highest peak typically located at the start of a sentence and the end is low. Also, the intonation range is narrower in Finnish than in English. (see e.g. 24,58-60.)

2.2.1 L1 and L2

Foreign language is by definition a language which is taught at school and has no status as a routine medium of communication in the country, while second language can be defined as a non-native language which is widely used for purposes of communication (42). By these definitions it would be safe to say that the Finnish subjects´ English is a foreign language, and the English subjects´ Finnish is a second language, since the research has been made in Finland, and the English subjects had very little to none education in Finnish. However, such distinctions have not been made, the term foreign language is used for both Finnish and English subjects´ L2.

2.2.2 Foreign language learning

Foreign language learning has interested researchers for decades, and several models for L2 learning have been created. Crystal (42) (Fig. 1) presents two models for foreign language learning: behaviourist and cognitive. In the behaviourist view L2 is learned by formal teaching, while in the cognitive view language learning is obtained through authentic use of L2 as spoken by native speakers of the language. The behaviourist model sees the language learning as a process of imitation and reinforcement, that is, the language learners try to copy what they hear, and by practice and repetition they can establish acceptable habits in the target language. In the cognitive view, the language learners use their cognitive abilities to construct rules of the target language. The rules are tried out and altered if they turn out to be inadequate. Language learning is seen to proceed in series of transitional stages that are not equivalent to either L1 or L2, so called interlanguage.

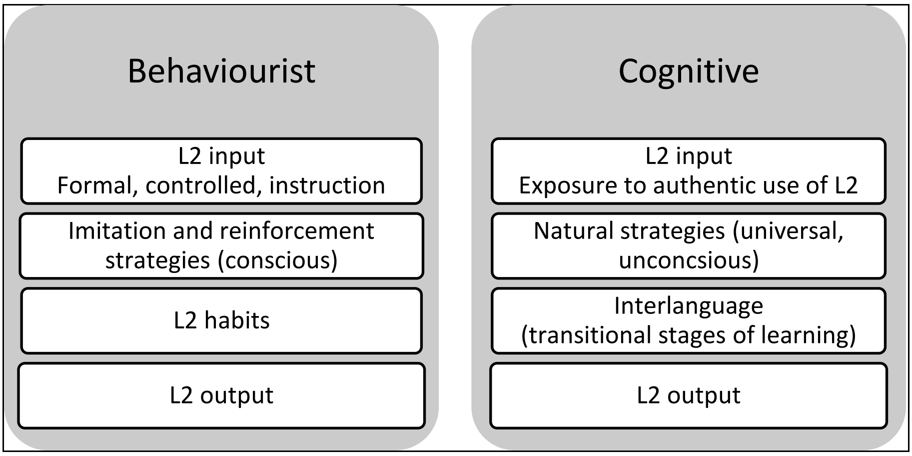


Figure 1. The foreign language learning models by Crystal (42).

It can be hypothesized that in this research, the Finnish subjects probably had mostly been using the behaviourist model, the English subjects mostly the cognitive one in their foreign language learning. This is, though, not that precise, since the Finnish subjects had probably been exposed to the English language as spoken by natives to a large extent through for example popular culture, and some of the English subjects had had some formal education in the Finnish language.

Most of the L2 learning models focus on language learner´s accuracy and accent in L2. For instance, Flege´s (61) Speech Learning Model (SLM) and Best´s (62) Perceptual Assimilation Model (PAM) aim to account for age-related changes in the ability to produce and perceive L2 sounds. In L2 learning the environment plays an important role across ages, such as the length of residence and relative usage of the language (16,63,64), as well as the relative quantity and quality of input from native L2 speakers (65,66).

Since accuracy is not the main focus in this research, Communication Accommodation Theory (CAT) (67) suits best for the purposes of this study. This theory has been developed in the 1970´s, and it is based on the hypothesis that, when people interact, they modify their speech, vocal patterns, and gestures to accommodate those of other people. The two main concepts in the theory are convergence and divergence. In convergence a person adapts a speech style to match the interlocutor´s speech, while in divergence the speaker starts using a strong regional or social variation of the language. The adaptation may be partial or total, for example, a speaker may speed up the speech tempo close to the interlocutor´s speech tempo, or to fully match it. Adaptation may also be verbal or nonverbal. In verbal adaptation, the speaker may use a different language, regional variation of language, or vocabulary. In nonverbal adaptation, the speaker may change e.g. the pitch or loudness closer to the interlocutor´s pitch or loudness, which can be seen already in infants, who, for instance, babble at a lower pitch in the presence of their father, compared to in the presence of their mother. Adaptation is subjective when the speaker makes changes that are based on stereotypes of the interlocutor, as the speaker believes some features to be present, and objective when the changes are in relation to the features that are, in fact, present in the interlocutor´s speech. Adaptation is usually unconscious, but to some extent, the speaker may be aware of it, more in the case of the divergence than of the convergence. (67-72.)

Adaptation can be seen, for instance, in voice of foreign language speakers even when speaking a language they do not have previous skills in. The foreign language speech is, then, imitation of a language and it bases primarily on the notions, stereotypes and images of the native speakers´ use of voice. (26.)

In this study, L2 is considered to be a learned language, and L2 learning seen as an ongoing process. This is in line with Flege´s terminology, when he states that learning a language does not have a clear endpoint and, thus, can be considered as a learning process rather than acquisition (73).

2.3 Voice

The voice of a speaker is the premise for speech. For person recognition, the voice is one of the most important cues, and it includes information on a person´s long-term traits, such as personality, sex and age, and speaker identification is mostly obtained by vocal tract characteristics, and speakers´ vocal identity may differ in acoustic features (72,74). The system for voice production consists of the lungs, the structures of the larynx, and the vocal tract (Fig. 2). The lungs create the energy source for voice, the vibration of the vocal folds in the larynx create the sound, and in the vocal tract it is modified to speech by resonances and articulation.

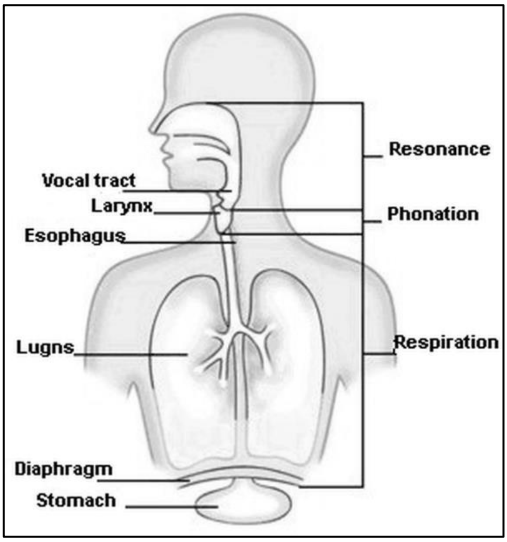


Figure 2. Voice production system. (Figure: http://www.gbmc.org/anatomyandphysiology, legends modified).

As air is inhaled, diaphragm lowers, the volume of the lungs expands, and air fills the space. To produce sound, adductor muscles are activated creating resistance to exhaled air. Air flows through the larynx, which is a tube shaped structure consisting of cartilages, muscles, and mucous membrane. The pressure between the vocal folds drops sucking them back together. This vibration is based on the elasticity of the vocal folds, aerodynamics of the larynx (Bernoulli Effect), inertia of the air column in the vocal tract, and the inner muscles of the larynx. The sound generated in the larynx is then modified by muscular changes in the pharynx and oral (and nasal) cavity, and by articulators (e.g. lips, tongue) to create speech. (75). (Fig. 3)

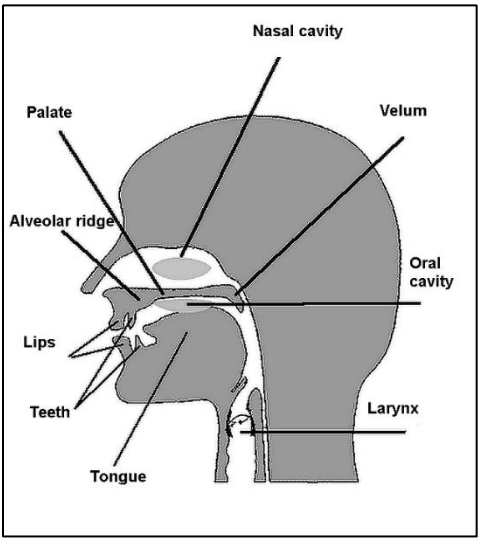


Figure 3. Vocal tract, articulators and places of articulation. (Figure: https://publicspeakingstrategies.wordpress.com/2015/06/04/a-ted-speakers-speech-impediment-and-ways-to-overcome-it/, legends modified).

The voice is, in addition to anatomy features, affected by the way it is produced (phonation), which may be influenced by geographic, social, and psychological factors (52). One way of investigating phonation is to study the firmness of the closing of the glottal closure i.e. the adduction (moving closer to each other) of the vocal folds: in hypofunctional voice the adduction of the vocal folds is too loose, and in hyperfunctional voice it is too hard in relation to subglottic air pressure (75). According to Laver (76) languages have articulatory settings that are typical for the language, and every person has his/her own personal settings in the use of the vocal organs which lead to individual articulatory settings.

2.4 Voice research

The acoustic research of voice gives an opportunity to study the vocal performance and the quality of the voice in an objective way. The characteristics of the voice to be studied here are fundamental frequency, sound pressure level, spectral qualities, quality of the voice generated at the voice source by inverse filtering (IAIF), and vocal loading. Also, voice quality by perceptual analyses will be studied, since acoustical parameters have been shown to have a relation to perceptual, auditive parameters (36).

2.4.1 Fundamental frequency

The fundamental frequency (F0) is the most important correlate to the perceptually observable pitch. It is based on the vocal fold vibration measuring its periodic duration in time. The unit of the fundamental frequency is a hertz, which is one vibration per second. (77.)

Pitch and its use are somewhat culture-dependent. According to Ohara (49,50) in Japanese society women´s high pitch has a great meaning in reflecting femininity, while in Western countries a lower pitch is more valued (56,57). Also in other languages pitch has a significance in perceptual observation (78-80). Pegoraro Krook (53) has studied differences in fundamental frequency in several languages, and it seems that differences in women´s mean F0 may be language dependent.

Previous studies have shown that F0 can be affected by the language shift from L1 to L2 (26,27,49,50). The reason for F0 to change between two languages has been explained by cultural and by physiological factors. Cultural factors may affect the use of pitch, as it is possible that speakers of L2, trying to achieve a native-like pitch, change the use of pitch in order to sound more native-like. (26,27,49,50,81-83). Also, another explanation has been offered by Järvinen et al. (26) who state that the rise in F0 may be a consequence of the fact that speaking a foreign language is a task more stressful than speaking the native one, which, then, increases the psycho-physiological stress of the speaker resulting in rising of F0 (84).

F0 may be affected also by linguistic features, such as intonation. Languages differ from each other in intonation, what kind and how wide it is (e.g. 13,85,86). In intonation research, some mutual tendencies between languages have been found, such as the rising intonation in questions and descending intonation in opposition phrases, also politeness is often expressed by high pitch (87,88). Even more common is that small things and submissiveness are expressed with a high pitch, and big things and threat with a low pitch. This is language and culture independent, and emerges even between species. (89.)

2.4.2 Intensity

The main physiological features correlating with intensity of speech are the glottal amplitude (amplitude of the vocal fold vibration, i.e. width of excursion of the vocal folds from the midline during vocal fold vibration), and the speed of glottal closing. Sound pressure level (SPL) and Equivalent sound level (Leq) are physical parameters that, to a large extent, correspond the perceptual observation of loudness, and the measuring unit is decibel (dB). SPL and Leq differ mainly by the fact that SPL measures the strongest peak of the sample, while Leq measures the total sound energy. (75,77.)

The normal conversational speech has an average SPL of 60-70 dB, when measured in a well damped space with the microphone 40 cm from the speaker´s mouth (77). Intensity and fundamental frequency are connected, a rise in the intensity raises the fundamental frequency (75,77).

Previous studies of SPL (or Leq) between L1 and L2 is scarce, if not non-existing, but, according to Järvinen et al., Leq is not affected by the language shift from L1 to L2 (27).

2.4.3 Spectral analyses

Long-time average spectrum (LTAS) gives out information about the sound energy and its distribution to different frequency areas. It is an average of many individual spectra in a long (usually one minute) speech sample. LTAS discloses the quality of voice, and it is used in analyses of speech and singing to describe the target voice, for example. (76.) LTAS can offer information on the individual characteristics of the speaker, despite the linguistic content (90). Voice quality can be analyzed acoustically from the long-time average spectrum (LTAS) as it provides an averaged spectrum of all voiced sound in the sample (91,92).

Contrastive research has concentrated on language and/or cultural factors and their influence on spectral structure and vowel formants (47,48,81,93-98). Previous studies have shown that the acoustically analyzed quality of voice can be language dependent, and that LTAS and its differences between languages can be due to the cultural use of voice (e.g. 27,47,51,55,99-102).

The alpha ratio is a method for describing the steepness of the spectrum. It is the ratio in RMS-energy in the frequency lanes of the sound signal, α = over 1000 Hz/ under 1000 Hz. (103.) The alpha ratio can be calculated also by reducing the sound energy in the high frequency area with the sound energy in the low frequency area, (SPL (1-5kHz)-SPL (50Hz-1 kHz). However, in this study the low frequency area was set at 50-1500 Hz, in order to avoid the effect of different formant frequencies in the two languages on LTAS.

The L1-L0 level difference is the difference in spectral amplitudes between the F0 region (0-300 Hz) and F1 region (300-1200 Hz), and it can give out information about the type of phonation. A higher alpha ratio and a lower sound level of F0 region in relation to F1 region may indicate a more pressed phonation type (Fig 4). (104-106.)

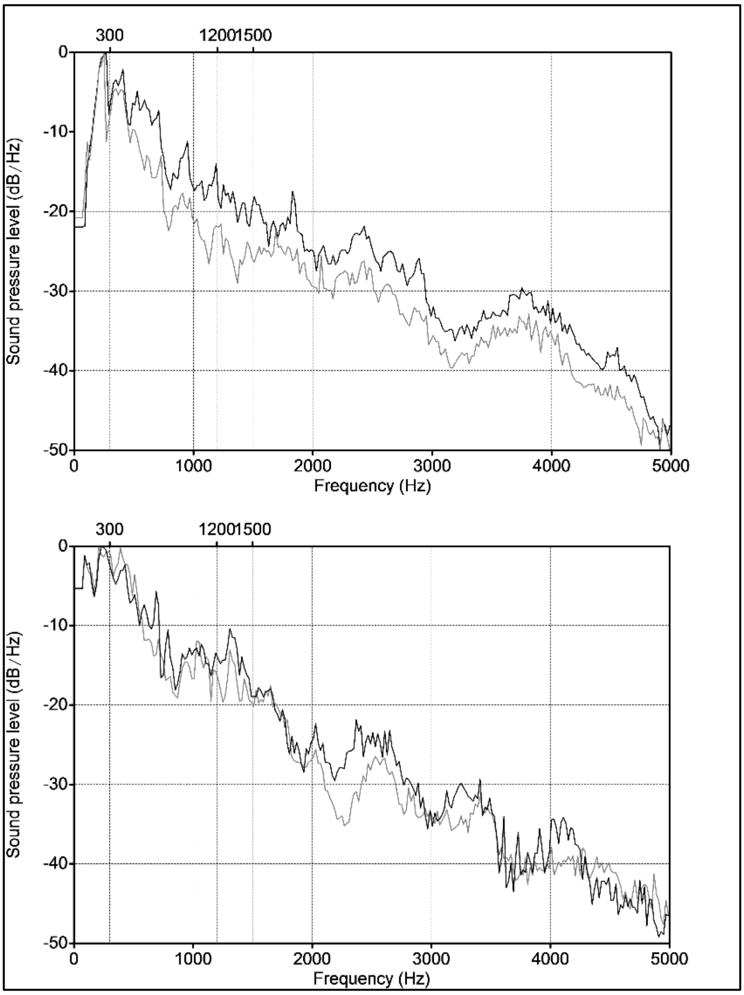


Figure 4. LTAS in L1 and L2. Above Finnish female, text reading, below English male, spontaneous speech. Black line Finnish, grey line English.

In different languages similar speech sounds may be actualized with different vocal tract settings, which may lead to differences in formant frequencies, as in the Finnish and English vowel systems (Fig. 5) (13,24,107). In addition to language dependent factors, anatomical features, such as length and shape of the vocal tract have a strong effect on formant frequencies. Formant frequencies are then somewhat speaker dependent (108). In addition to the formant frequencies, LTAS can also provide spectral information on changes in the articulatory settings (109).

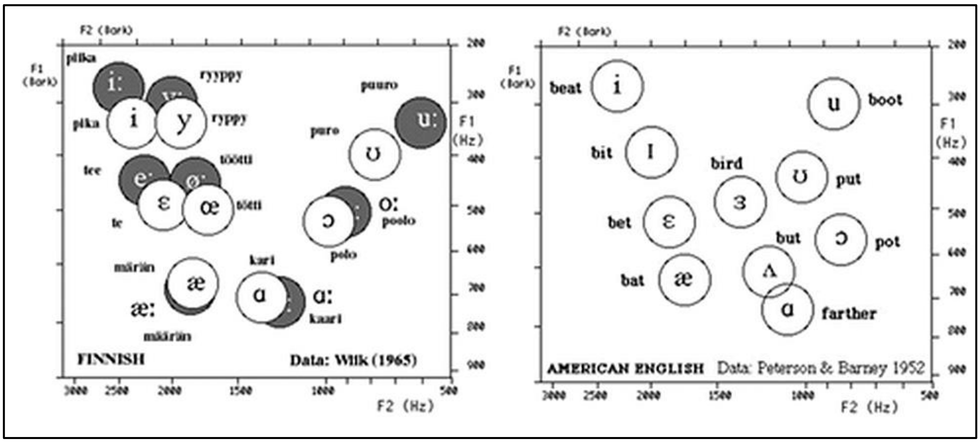


Figure 5. Vowel formant charts for Finnish and (American) English short vowels. (Figure for Finnish: http://www.helsinki.fi/puhetieteet/projektit/Finnish_Phonetics/vokaaliakustiikka_eng.htm, accessed March 29th 2016, modified, and for English: http://isip.piconepress.com/courses/msstate/ece_8463/lectures/current/lecture_17/index.html, accessed March 29th 2016).

2.4.4 Inverse filtering

Inverse filtering, developed in the 1950's by Miller (110), reveals the voice source, that is to say, the airflow pulses that are generated by the vibration of the vocal folds (111). The basis of the source-filter approach is the ideas that the source and the filter are independent of each other, and that speech consists of three separate and independent processes: glottal excitation, vocal tract filter, and lip radiation (112). Rothenberg (113) introduced an inverse filtering method that estimates the airflow out of the mouth through a mask avoiding the lip radiation effect. Iterative Adaptive

Inverse Filtering (IAIF), was developed by Alku (114), and it is one of the methods that estimates the glottal flow signal from a corresponding acoustic speech pressure signal.

The parameters investigated through inverse filtering can be time or amplitude based. The temporal (time based) parameters include open quotient (OQ), closing quotient (ClQ), and speed quotient (SQ). (110.) In previous studies OQ has been found to have a negative correlation with the intensity and the pressedness of speech, breathy voice implies increased open quotient, while an increase of the intensity increases SQ (115-117). ClQ decreases when pressedness and intensity increase (115). The closing phase affects most directly the voice quality, since the abrupt closure of the glottis at the end of the closing phase is to a large extent responsible for generation of the voicing energy (118). ClQ decreases when intensity and pressedness increase (115).

Time based parameters may be problematic, since the exact points of the glottal opening or closing can be hard to determine, so amplitude based parameters, indicating the features which are related to the time domain of the signal, have been shown to help avoid the problem (111). Amplitude quotient (AQ) is the ratio of the flow peak-to-peak amplitude of the flow (that is, the difference between maximum and minimum value within one period) and the minimum peak of the pulse derivative (111,119). In order to avoid the influence of F0 to the measuring of AQ, a parameter NAQ has been introduced, which is AQ with F0 normalization (111,118). AQ and NAQ have been found to correlate negatively with pressedness of voice (114,120).

The time and amplitude based parameters through IAIF to be studied are OQ1, OQ2, ClQ, SQ1, SQ2, AQ, and NAQ (see Fig. 6, in chapter 4.4, page 45).

2.4.5 Parameterization of vocal loading

Overloading of the vocal organs can be a result of excessive use of voice, and this overloading can lead to subjective sensation of vocal fatigue. Vocal fatigue can be manifested in poor quality of voice, unpleasant feelings in the throat and even in voice loss. (75.) Vocal overloading is regarded as being a result from a combination of loading factors such as excessively prolonged voice use, high pitch, intensity, and a pressed type of phonation (30,104).

Several parameters for quantifying vocal loading has been proposed, such as F0 and SPL (29,121), Vocal Loading Index (122), the phonation threshold pressure

(123), and the characteristics of vocal fold vibration (124). Also, vocal doses (31,32) can be used in quantifying vocal loading.

In studying the effects of vocal loading, the self-reported vocal loading symptoms (104) are included in the analyses in addition to the acoustic analyses. Also, the voice can be affected by individual factors such as gender, vocal endurance, general health, life habits, vocal skills, and experience and personality (29), and they may contribute to the effects of vocal loading.

This study concentrates on Vocal Loading Index and vocal doses in quantifying the loading of the voice and the subjective sensation of vocal fatigue.

2.4.5.1 Vocal Loading Index

Vocal loading index (VLI) has been introduced by Rantala and Vilkman (122). The index can be calculated by frequency of the vibration of the vocal folds multiplied by the time of the vocal fold vibration divided by thousand (VLI = F0 x F0 time / 1000, i.e. number of vocal fold periods in kilocycles). VLI indicates the amount of the vocal fold vibration, it is an index of the work the vocal folds do. According to Rantala and Vilkman, the subjects´ own observations of the problems in voice and VLI have a clear connection. (122.)

2.4.5.2 Vocal doses

Titze, Švec and Popolo have introduced five vocal doses for quantifying the amount of vocal loading which are based on F0, SPL and voicing time (31,32). The doses are:

- time dose (D_t) indicating the total voicing time
- cycle dose (D_c) which indicates the total number of vocal fold oscillatory cycles (i.e. the same as VLI times 1000)
- the distance dose (D_d) which indicates the total distance the vocal folds travel in an oscillatory path
- the energy dissipation dose (D_e) as total amount of dissipation of heat energy in the vocal folds
- radiated energy dose (D_r) indicating the total acoustic energy that radiates from the mouth.

According to Titze et al. (32), the speaking style (normal-monotone-exaggerated) may have an effect on the vocal doses, as the highest values were found in the exaggerated speaking style. Only the time dose (Dt) did not seem to be affected by the speaking style. When the doses were normalized to Dt they found a trend that the mean exposures per second were lowest in monotone and highest in exaggerated speech. The differences were substantially greater for the energy dose (De) and the radiation dose (Dr) than for the cycle dose (Dc) and distance dose (Dd). Also, some differences between men and women were found. The normalized vocal loading index (VLI/Dt) for men was 100-150 cycles/s and for women 200-250 cycles/s. The values correspond, as expected, to the fundamental frequency values. The normalized Dd was about 0.5m/s in the normal speech. The normalized De was 0.39–0.86 mJ/(cm3s) for men and 0.16–0.25 mJ/(cm3s) for women. The normalized Dr was 0.009–0.036 mJ/s for men and 0.02–0.057 mJ/s for women. The results suggest that women were more efficient in their voice production than men, since their Dr/Dt was about twice the amount for men but De/Dt only about half of that for men. (32.)

The dose measurements have been used in previous research in studying vocal loading among schoolteachers and/or kindergarten teachers (e.g. 125-128), also research has been done on vocal doses among singers (e.g. 129,130). The results of the studies have shown that vocal doses are suitable in quantifying vocal loading. Kindergarten teachers have higher vocal doses than elementary school teachers (125) and, also, that the vocal load is higher in professional voice use environment than in the nonprofessional one (125,126). Also, a large vocal dose may have an effect on inability to produce a soft voice, vocal effort, and the subjective ratings of voice (130).

Previous studies on vocal doses in two languages is to this author´s knowledge non existing, but a short mention on language effect has been made by Titze et al. (32) as they have found that the total amount of voiced speech was lower with a non-native speaker than with the native speakers. The dose measures, then, may be affected by language.

2.4.6 Perceptual analyses of voice quality

According to American National Standards Institute (ANSI), voice quality is a perceptual term, defined as “that attribute of auditory sensation in terms of which a listener can judge that two sounds similarly presented and having the same loudness and pitch are dissimilar” (131). In addition to the acoustical analyses of the voice signal, perceptual analysis of the voice quality is widely used as a research method of voice. Laver (76) defines voice quality as the characteristic auditory colouring of an individual speaker´s voice. Both laryngeal and supralaryngeal features can be seen as contributing to voice quality. The perceptual analyses can reveal the voice quality of the speaker (132), since it is possible to perceptually detect even small changes in the voice signal, although the ability to sense these changes are individual (133). Vocologists may be defined as expert listeners of voice quality, since training in perceptual evaluation of voice is included in their education (134).

Perceptual analyses can provide information on the speaker´s phonation type. Kankare et al. (135) have found that perceptual analyses of pressedness are in line with EGG parameters (glottal pulse parameters obtained by electroglottography), however, the perceptual evaluations did not correlate with the self-evaluated voice problems. They argue that voice quality is not necessarily a major factor in vocal loading in healthy voices, since pressed phonation type is only one factor in vocal loading. On the other hand, they discuss that the voice samples used in the perceptual analyses may have had an effect on the results, as they used sustained vowels and extracted vowels from continuous speech. In some previous studies continuous speech has been argued to be more favourable as describing the phonation type than sustained vowels (136,137).

In previous studies, different systems for analyzing voice quality perceptually have been developed, and most of the studies have been focused on evaluation of pathological voices (138). The GRBAS scale (139) is widely used as a tool for perceptual analysis of voice quality. The scale includes a scale from zero to three, 0 corresponding to healthy voice and 3 to severe disease, in five qualitative characteristics: Grade of dysphony (G), Roughness (R), Breathiness (B), Asthenia (A), and Strain (S). A visual analog scale (VAS) may, also, be used in evaluation of voice quality with GRBAS (140,141), but Wuyts et al. (141) recommend the use of the original 4-point scale. However, VAS offers more detailed information as the listeners tend to perceive vocal characteristics along a continuum, and not by intervals (138). GIRBAS scale is a GRBAS scale with an addition of instability (I) (142). The Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V) is based on perceptual vocal characteristics, which are vastly used and easily understood. The characteristics are overall severity, roughness, breathiness, strain, pitch, and loudness (143,144). GRBAS has been considered to be a fast and simple method and having good correlation with acoustic parameters (140), and the CAPE-V scale as being more sensitive, in especially for detecting small quality changes in voice (144). The Vocal Profile Analysis (VPA) is a system suitable for both normal and pathological voices. It involves both the articulatory and phonatory settings, and the general principles are that an individual´s characteristic voice quality is affected by the whole of the vocal apparatus, and that voice quality is a combination of more or less independent settings. Also, a normal baseline is not used, since VPA is based on "neutral" setting. (145.)

In perceptual analyses the reliability of the listeners´ judgements seem, to some extent, be affected by the method used. For example, Webb et al. (146) have found GRBAS to be more reliable than VPA. Also, to enhance the reliability of the perceptual judgements, it has been suggested that external referenced standards (anchors) should be added to the perceptual analyses (147) since every listener has his/her internal subjective standards which may affect the judgements (148).

The present study used sentences extracted from continuous speech in perceptual analyses of voice quality. Also, in perceptual analyses overall voice quality, strenuousity of voice production and suitability of pitch were used in addition to phonation type of the speakers. Anchors were not used in this study, since the comparison was made between same speaker´s L1 and L2 samples, not between speakers.

3 AIMS OF THE RESEARCH

3.1 Research questions and research hypothesis

This present research aimed to answer the question if speaking a foreign language affects the speaker´s voice and voice use compared to speaking the native language, and whether speaking L2 is more loading than speaking L1. The research hypothesis was that voice changes when speaking in a foreign language compared to speaking in the native language. Research questions in studies were as follows:

Study I:

1. Do mean F0 and the F0 range change when speaking a foreign language compared to speaking the native language?
2. Are the speakers aware of the changes?
3. Does experience, e.g. residence and education, in speaking the foreign language have an effect on the changes in F0?

Study II:

1. Does speaking in L2 affect the subjective notions of vocal fatigue?
2. Can increased vocal loading be seen in acoustical analyses?

Study III:

1. Is the perceived phonation quality more pressed when speaking L2 than when speaking L1?
2. Do the characteristics of the voice source (revealed through inverse filtering) differ between L1 and L2?
3. Do the subjects who report more vocal fatigue in L2 than in L1 have more pressed phonation in L2?

4 METHODS

4.1 Subjects

The first step was to recruit subjects for recordings. A letter (Appendix 1) was sent to Department of English at the University of Tampere, and to the students in courses “Finnish for foreigners” held at the University of Tampere. Also, the researcher used her own connections and recruited people she knew as subjects. The only criteria for the subject selection was that every subject should know the foreign language at a level that the text-reading and spontaneous speech samples would be possible to produce. However, three of the English subjects were not able to produce the spontaneous speech samples.

The subjects were native speakers of either Finnish or English, and the foreign languages they used were English or Finnish. The Finnish subject group consisted of twenty people, eight males and twelve females, and the English group consisted of 23 subjects, fourteen males and nine females.

The mean age of the Finnish subjects was 39.8 years (sd 15.9, range 19-69) and the English subjects´ 36.0 years (sd 16.9, range 21-79). Finnish subjects had a longer formal education in L2 (mean 12.0 years, sd 9.2, range 0-40) than the English subjects (mean 0.8 years, sd 1.1, range 0-6).

The English subjects had a longer residence in Finland than the Finnish subjects in an English speaking country, 22 percent of the English subjects had resided in Finland over 10 years, while only 5 percent of the Finnish had residence of 10 years or more.

The Finnish subjects considered themselves as more experienced in speaking English than the English subjects in speaking Finnish: seventy percent of the Finnish subjects considered to be at least experienced, while almost seventy percent of the English considered to have only some experience. The data was gathered by a questionnaire (Appendix 2).

4.2 Recordings

The subjects were recorded at the University of Tampere. The recordings took place in a well-damped recording studio with a combined level meter and microphone (Bruel et Kjaer Mediator, Type 2206), placed in front of the subject at a distance of 40 cm from the mouth. The signal was recorded with Sound Forge 7.0 software, using 44.1 kHz sampling frequency and 16-bit amplitude quantization. The signal was calibrated for calculation of SPL and Leq by recording a reference signal with a known intensity.

The subjects read a text (Appendix 3) first in their native language and then in the foreign language. The subjects read aloud from a text for one minute first in the native language (Finnish or English) and then in the foreign language (English or Finnish). The texts were the same in content in both languages. Also, spontaneous speech samples were recorded. The subjects were shown a comic strip (Appendix 4) which they described in their own words first in the native language and then in the foreign one. The duration of the spontaneous speech samples varied between 21.5 seconds and one minute (mean of duration was 42.9 seconds) in L1 and between 26 seconds and one minute (mean of duration 43.6 seconds) in L2.

4.3 Questionnaire

After the recordings, the subjects were asked to fill in a questionnaire (Appendix 2). The subjects were able to fill in the questionnaire at the same location the recordings took place. This helped both the subjects as well as the researcher. Every subject filled in the questionnaire and the researcher was able to clarify if some questions were not all so clear. Had the subjects been asked to fill in the questionnaire for example at home, it is possible that not all of them would have filled it in. Also, the questionnaire was planned to be filled in after the recordings so the subjects would not have an exact idea of the aim of the study.

4.4 Acoustical analyses

The acoustical analysis were carried out with Praat speech analysis system (149) and TKK Aparat (150). Praat was used in Studies I and II. TKK Aparat was used in Study III.

In Study I, mean, standard deviation, and range of F0 were analyzed. Individual pitch ranges and analysis method cross-correlation were used in the analyses.

In Study II mean fundamental frequency (F0), equivalent sound level (Leq, intensity calculated period by period), speech rate (syllables per second), total duration of voiced speech (unvoiced speech segments and pauses were excluded), and VLI were analyzed. Alpha ratio was calculated by subtracting the Leq within the range of 1.5-5 kHz from the Leq in range of 50-1500 Hz. 1500 Hz was chosen as a level value, because it may then ignore the different formant frequencies in Finnish and English (for an example of different formant frequencies in Finnish and in English see Figure 4, page 33 and Figure 5, page 34).

The level difference between peak in the first formant region (300-1200 Hz) and peak in the F0 region (0-300 Hz) was measured from LTAS. Analysis of LTAS was pitch-corrected, and unvoiced segments were excluded.

The dose measures were used to quantify vocal loading. For the dose measurements, Leq 40 cm was converted to Leq 50 cm. The formula Leq50 cm = Leq40 cm – 10 · log √ 40/50 was used. The time resolution for dose measurements was 0.01 s. Since the duration of the L1 and L2 spontaneous speech samples varied, the same length of speech was analyzed in dose measurements in both languages.

In calculation of doses, the equations introduced by Titze et al. (31) were used. Time dose (Dt) was calculated with the total time and the voicing unit step function. This equation is the basis of every dose calculation. In calculation of cycle dose (Dc), F0 was added in the equation. In the distance dose (Dd) the amplitude of the vocal folds was added in the equation, and also a factor 4, since, theoretically, the vocal folds travel a distance of four times the amplitude within a cycle. The distance dose accounts for both intensity and fundamental frequency in voicing. To overcome the problem with the distance dose (the vibration amplitude of the vocal folds being very difficult to measure), the amplitude can be approximated from the SPL and F0. This means that, a distance dose typical for an average person when speaking at the measured SPL and F0 are calculated. The amplitude can be approximated using the empirical rules which take into account the reference vocal fold length (0.016 m for males and 0.01 m for females), the lung pressure and the phonation threshold pressure. See Titze (151) for the empirical rule for the phonation threshold pressure. Also, a nominal fundamental frequency (120 Hz for males and 190 Hz for females) is added in the equation. In energy dissipation dose (De) the shear viscosity of the vocal fold tissue (see 152,153), including the vertical thickness of the vocal folds and the angular frequency of the vocal fold vibration are added to the equation.

Calculating the radiated energy dose (Dr) the distance from the mouth at which SPL is registered is taken into account. For the equations, see Titze et al. (31).

The doses are derived for a specified measurement time as three basic parameters of speech are extracted: the parameter for voiced and voiceless parts of the speech, F0 and SPL. D_t, D_c and D_r are the true doses for the person, whereas D_d and D_e are approximations, since they are based on typical data for vocal fold amplitudes, thicknesses, and viscosities for both male and female speakers (31).

In Study III, 24 subjects were divided into two groups, subjects who did not report more vocal fatigue in L2 than in L1 formed Group 1 (12 subjects in total), and subjects who reported more vocal fatigue in L2 than in L1 Group 2 (12 subjects in total). From the speech samples short stressed vowels (/a/ in Finnish and /ʌ/ in English) were extracted in three different phonetic contexts both in text reading and in spontaneous speech. The length of the analyzed vowels was 0.045 seconds. The order of the vocal tract model and the lip radiation coefficient were adjusted separately for every vowel. To remove low-frequency ambient noise, the cut-off frequency of the high-pass filter was set at 80 Hz, and changed when necessary. The mean of the three acoustical parameters for text reading and spontaneous samples was calculated for every subject.

The following time and amplitude based parameters were studied from the inverse filtered signal (Fig. 6): Open quotient (OQ), closing quotient (ClQ), speed quotient (SQ), amplitude quotient, AQ, and normalized amplitude quotient, NAQ. Calculation of these parameters (111,114,118,119):

- the glottal open phase in comparison to the cycle duration = open quotient (OQ)
- the ratio between the duration of the glottal closing phase to the period length = closing quotient (ClQ)
- the ratio between the duration of the opening phase and the duration of the closing phase = speed quotient (SQ)
- the ratio of the peak-to-peak amplitude of the flow (Aac) and the minimum peak of the pulse derivative (Admin) = amplitude quotient, AQ
- AQ divided by period length = normalized amplitude quotient, NAQ.

The opening phase has two opening points (to1 and to2 in Fig. 6): OQ1 is calculated from the primary opening point while OQ2 is calculated from the secondary opening point (154). Similarly, SQ1 is calculated from the primary opening point, and SQ2 from the secondary opening point.

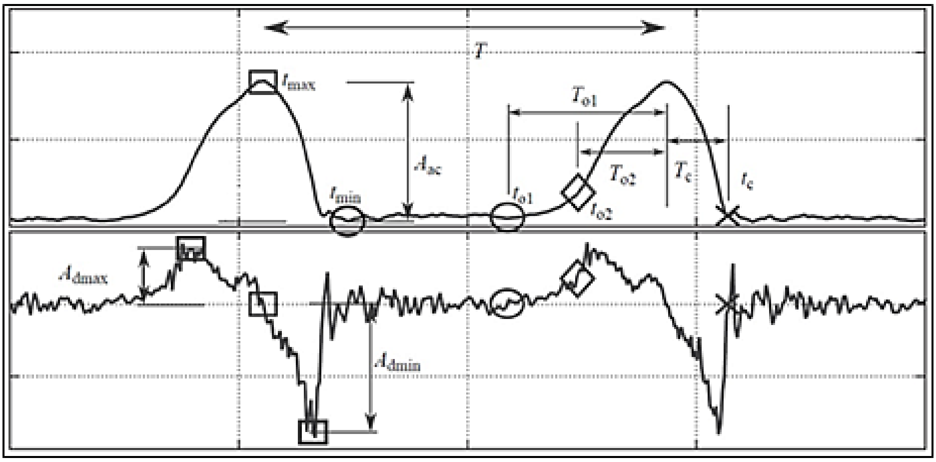


Figure 6. Time and amplitude points in calculations of the inverse filtering parameters. (Figure: Järvinen et al. Voice quality in native and foreign languages investigated by perceptual analyses and inverse filtering. J Voice 2016).

4.5 Perceptual analyses

Perceptual analyses were used in Study II and Study III. In Study II, two native listeners of the spoken language (1 native Finnish and 1 native English) evaluated the proficiency of the speakers in the foreign language. The listening samples were extracted from the speech samples, same sentences in text reading samples were used for both Finnish and English subjects, in spontaneous speech the listening samples were one sentence at the beginning of the speech sample. The listeners evaluated the speech on a scale from 0 to 4, 0 representing no skill/fluency/accuracy and 4 representing native-like speech. Proficiency was created by combining skill, fluency and accuracy into a sum of the variables. Questionnaire for these perceptual analyses is presented in Appendix 5.

In Study III three vocologists listened to the voice samples (one sentence from the same point in L1 and L2 with duration from five to eight seconds) and evaluated the speakers´ overall voice quality, strenuousity of voice production, firmness of phonation and pitch suitability. Questionnaire for perceptual analyses used in Study III is presented in Appendix 6.

4.6 Statistical analyses

Statistical analyses were carried out by SPSS software (versions 18 and 22). Related samples Wilcoxon signed ranks test was used for the significance of change in the parameters, Mann-Whitney U-test was used to test the significance of the difference in the change of the parameters between groups. Repeated measures analysis of variances (RM-ANOVA) was used for studying the differences in inverse filtering parameters between native and foreign languages and the effect of subjective vocal fatigue. Spearman correlation coefficient was used for investigating the relations between self-reported notions and the change in acoustical parameters, and perceptual evaluations. Also, Spearman correlation coefficient was used for investigating the relations between the acoustical and perceptual results, and Cronbach´s alpha for the inter-rater reliability. Nonparametric tests were used, since the results were not normally distributed.

4.7 Summary of methods and materials used in the studies

The materials, subjects and analyses in each study used are combined in Table 1.

Study	I Speaking a foreign language and its effect on F0	II Vocal loading in speaking a foreign language	III Voice quality characteristics in native and foreign languages
Subjects	30 16 Finnish 14 English	43 20 Finnish 23 English	24 12 Finnish 12 English
Material	Text reading Questionnaire	Text reading Spontaneous speech Questionnaire	Text reading Spontaneous speech Questionnaire
Acoustical analyses	F0 Standard deviation of F0 F0 range	F0 Leq Alpha ratio L1-L0 level difference Amount of voiced speech Speech rate Vocal doses	F0 OQ1 and OQ2 ClQ SQ1 and SQ2 NAQ AQ
Perceptual analyses		Skill Fluency Accuracy	Overall voice quality Strenuousity of voice production Firmness of phonation Suitability of pitch
Statistical analyses	Wilcoxon Signed Ranks Test Mann Whitney U-test Spearman correlation coefficient	Wilcoxon Signed Ranks Test Spearman correlation coefficient	Wilcoxon Signed Ranks Test RM-ANOVA Mann Whitney U-test Spearman correlation coefficient Cronbach´s alpha

Table 1. Summary of the subjects and methods used in studies.

5 RESULTS OF THE STUDIES

5.1 Mean, SD, and range of F0 in L1 and L2 (Study I)

The results showed that F0 was significantly higher in English than in Finnish for the Finnish subjects, but for the English subjects the change between L1 and L2 was not as clear. For the Finnish subjects F0 was in average four percentages higher in English than in Finnish (p=.001), for the English 0.84 percentage higher in Finnish than in English (statistically non-significant), (Figure 7). The change in F0 showed significant differences between Finnish and English subjects (p=.008). Standard deviation of F0 and the range of F0 showed no significant changes between native and foreign languages.

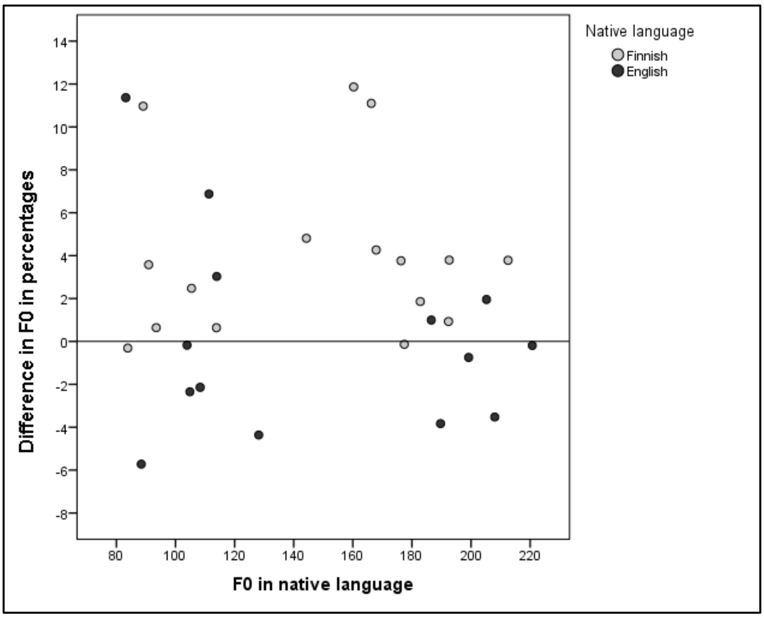


Figure 7. Difference in F0 in percentages in the foreign language compared to F0 in the native language in Finnish and English subjects.

Twelve (75 %) of the Finnish subjects considered that their pitch was higher in English than in Finnish, only two considered it to be lower. Eight (57 %) of the English subjects considered the pitch to be lower in Finnish than in English, two to be higher. Only two subjects in the Finnish and four in the English group reported that the pitch is not affected by the language change. The subjective notions of the change in pitch did not have significant correlation with the actual rise in F0.

The Finnish subjects had a longer formal education in the foreign language than the English subjects (14 and 0.7 years respectively), while the English subjects had a longer residence than the Finnish subjects. Fifty-seven percentages of the English subjects had resided in Finland over five years, while for the Finns all of the subjects had resided in an English speaking country for five years or less. The Finnish subjects considered themselves as more experienced in the foreign language use than the English subjects (Figure 8). No significant correlation between experience in speaking the foreign language and the change in F0 was found.

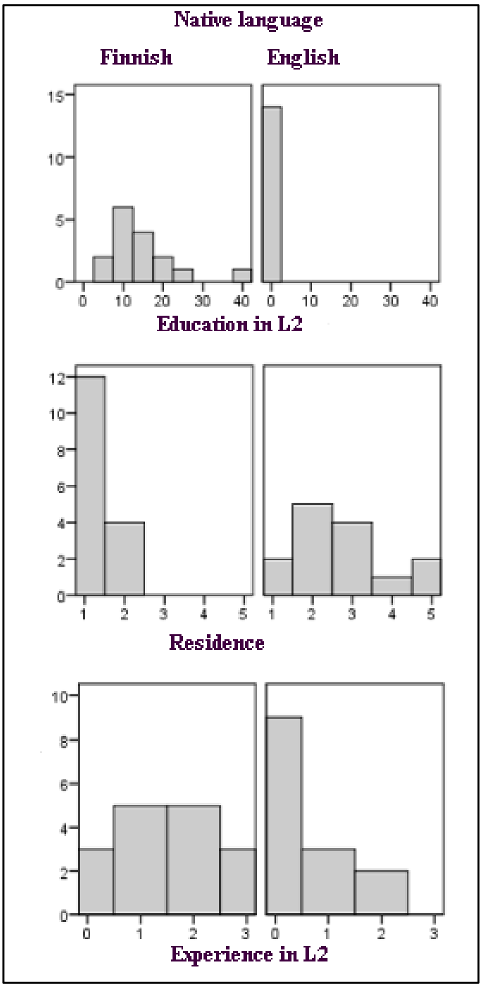


Figure 8. The subjects´ education in the foreign language (in years), the subjects´ residence in a country where L2 is spoken (1=less than 1 year, 2=1-5, 3=5-10, 4=10-20, 5=over 20 years), and the subjects´ own estimation of level of experience in speaking L2 (0=some experience, 1=experienced, 2=very experienced, 3=native-like). The vertical axes represents frequency of the responses.

5.2 Vocal loading in L1 and L2 (Study II)

The majority of the subjects (79%) considered that language shift from L1 to L2 has an effect on voice, as they reported that the voice changes when speaking L2 compared to speaking L1. Only two percent of the subjects had not noticed any changes in voice between L1 and L2, and nineteen percent only slight changes. Thirty percent of the subjects reported that language shift from L1 to L2 has an effect on vocal fatigue (Figure 9).

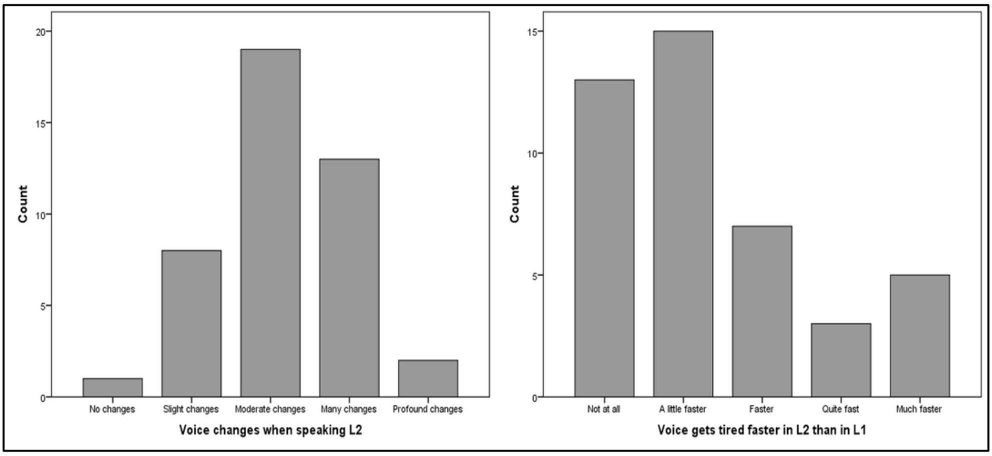


Figure 9. The subjects´ estimations of how much voice changes and whether voice gets tired faster when speaking L2 compared to speaking L1.

The subjects´ own estimation of experience in speaking L2 had a mild negative correlation with the estimation of vocal fatigue ($r = -0.48$, $p < .01$). The level of experience correlated for the Finnish subjects with education in L2 ($r = 0.55$, $p < .05$), and for the English subjects with the duration of residence ($r = 0.66$, $p < .01$).

According to perceptual evaluation by a native speaker of the target language, the Finnish subjects were evaluated as more proficient than the English subjects in speaking the target language. Seventy five percent of the Finnish subjects was evaluated as proficient, very proficient or even native-like in English while 56.5 percent of the English subjects was evaluated as proficient or very proficient. Males were evaluated as a little more proficient than females (72.7 % and 57.1 % respectively). No correlation between the perceived proficiency and the age or sex of the speaker was found. A moderate correlation between the self-evaluated experience in speaking L2 and perceptually evaluated proficiency was found ($r = 0.54$, $p < .01$). Also, a slight negative correlation between the self-reported notion of vocal fatigue and perceptually evaluated proficiency was found ($r = -0.34$, $p < .05$).

Proficiency correlated moderately with self-evaluated experience in the target language for the Finnish subjects ($r = 0.55$, $p < .05$), and slightly with residence with the English subjects ($r = 0.44$, $p < .05$).

Between L1 and L2 significant changes in the acoustical parameters were found in F0, and L1-L0 level difference. F0 was higher in L2 than in L1 ($p < .05$), and the L1-L0 level difference was higher in L2 than in L1 ($p < .01$). Changes in the alpha ratio and L1-L0 level difference did not have a correlation with the changes in Leq. The amount of voiced speech and the speech rate were significantly lower in L2 than in L1 ($p < .01$ in both).

The speech doses D_t, D_c, and D_d showed significant changes between L1 and L2 ($p < .01$ in all three dose measurement), the doses being lower in L2 than in L1. Age and sex of the speaker did not have a correlation with the changes in acoustical parameters.

Normalizing VLI and D_t and calculating doses per second showed significant changes between L1 and L2 in VLI/D_t ($p < .05$) for all subjects. For Finnish subjects a significant change was found in D_r/D_t ($p < .05$), and for females in VLI/D_t ($p < .05$). The normalized doses are presented in Figure 10.

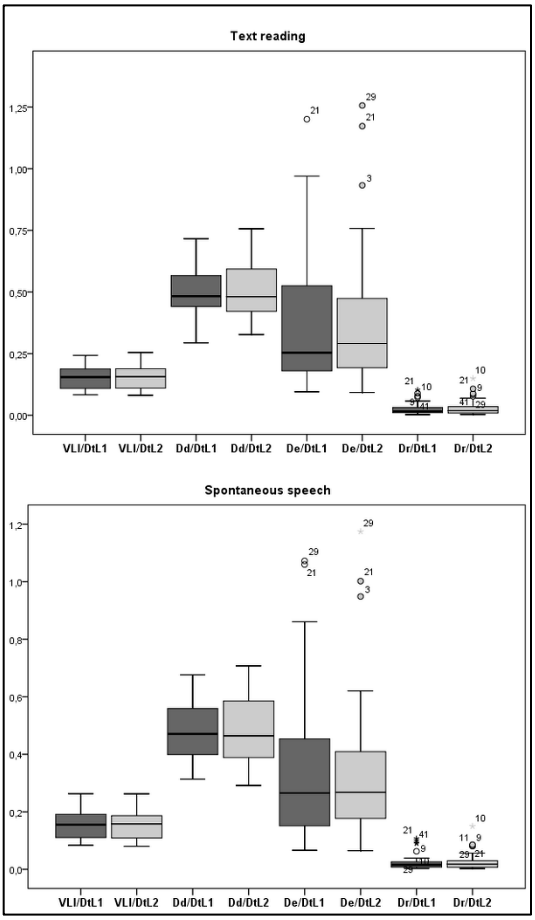


Figure 10. Distributions of dose measurements normalized to D_t in text reading and spontaneous speech samples.

5.3 Voice quality in L1 and L2 (Study III)

Results showed significant differences in the perceptual evaluation of voice in L1 and L2: in text reading the voice quality was evaluated as poorer in L2 than in L1 ($p < .05$), and in text reading and in spontaneous speech the firmness of phonation was lower (i.e. type of phonation more pressed) in L2 than in L1.

Significant differences between native and foreign languages were found in the acoustical analyses: in OQ2 ($p < .05$) in text reading, and in ClQ ($p < .05$), SQ2 ($p < .05$), and NAQ ($p < .01$) in spontaneous speech (Figure 11). Also, RM-ANOVA showed significant differences between L1 and L2 in ClQ (F = 6.20, df = 1, p < .05), NAQ (F = 5.32, df = 1, p < .05), and SQ2 (F = 13.81, df = 1, p < .01), but no significant differences were found when vocal fatigue was set as a between subject factor.

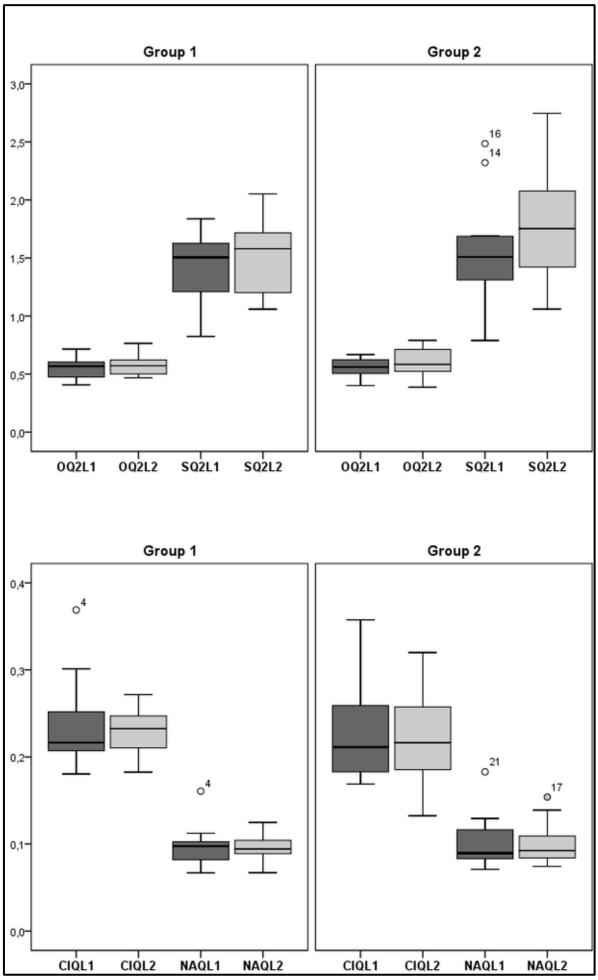


Figure 11. Distributions of parameters obtained through inverse filtering for the two groups of subjects: subjects who did not report more vocal fatigue in L2 than in L1 (Group 1), and subjects who reported more vocal fatigue in L2 than in L1 (Group 2).

Between the groups based on subjective sensations of vocal fatigue, significant differences were found. For subjects in Group 1 (subjects who did not report having more vocal fatigue in L2 than in L1) voice quality in L2 was evaluated as better (p =

.017), firmness of phonation lower (i.e. phonation type less pressed) (p = .001), and pitch more suitable (p = .01) than for subjects in Group 2 (subjects who reported more vocal fatigue in L2 than in L1). Significant differences between L1 and L2 in text reading were found in all the perceptual parameters for Group 2 as voice quality was evaluated poorer, phonation more pressed, voice production more strenuous, and pitch higher than suitable in L2 (Figure 12).

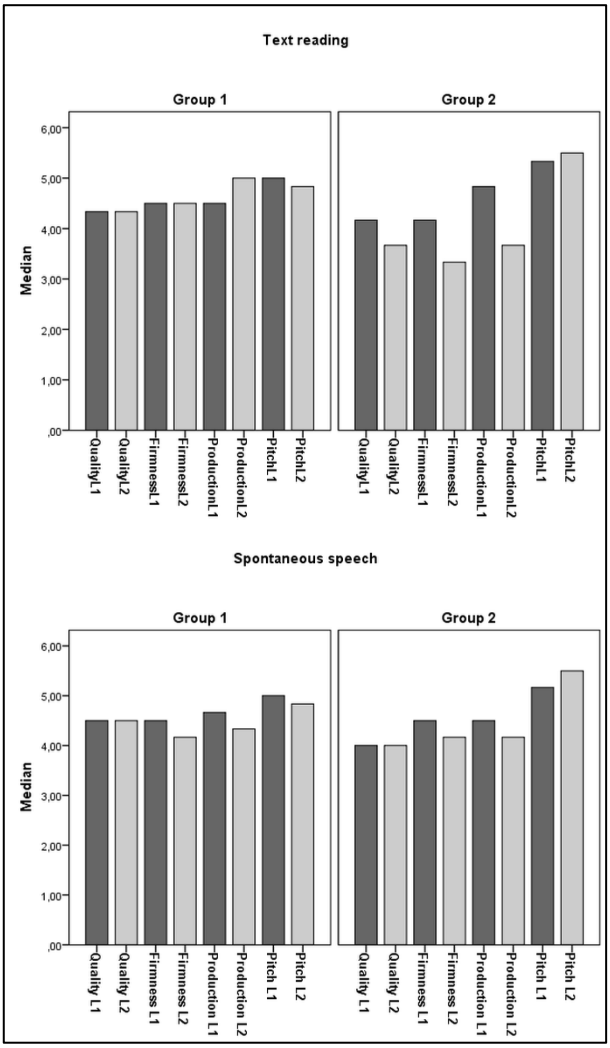


Figure 12. Medians of perceptual evaluations for subjects´ general voice quality (0=very poor – 10=very good), firmness of phonation (0=very pressed – 10=very breathy), strenuousity of voice production (0=very strenuous – 10=very easy), and pitch suitability (0=too low – 10=too high).

Both groups showed a significant increase in firmness in L2 in spontaneous speech ($p<.05$).

For some of the inverse filtering parameters significant differences were found within the same group. In spontaneous speech, Group 1 showed significant reductions of both NAQ and ClQ, and Group 2 showed significant increase in SQ2 in L2. Text reading showed the same trend but remained insignificant. Between Group 1 and Group 2 no significant differences in the inverse filtering was found.

The perceptual evaluation of strenuousity of voice production had a mild correlation with the acoustical inverse filtered parameters. The correlation varied from $r = .26$, $p < .01$ for OQ1 to $r = .47$, $p < .001$ for NAQ. The evaluated strenuousity of voice production correlated mildly negatively with SQ1 ($r = -.32$, $p < .01$), and SQ2 ($r = -.22$, $p < .05$). Also, the perceptual evaluation of pitch suitability had a mild correlation with F0 ($r = .41$, $p < .01$), and NAQ ($r = .30$, $p < .01$), and a mild negative correlation with AQ ($r = -.28$, $p < .01$).

6 DISCUSSION OF STUDIES

6.1 Speaking a foreign language and its effect on voice characteristics

The mean fundamental frequency changes when speaking a foreign language compared to speaking the native one, and the change is, to some point, related to the target language. According to Lewis (155) the use of pitch can be gender and culture dependent. This can lead the speakers of L2 to use a certain pitch level they think the native speaker of the target language would use.

The difference between Finnish and English groups indicates that the use of pitch in the foreign language may be a result of adaptation to a certain pitch level the speakers estimate the native speakers would use. In Speech Accommodation theory it is argued that people tend to modify their communication in interaction with other people. This adaptation includes both linguistic and non-verbal cues, and in vocal characteristics pitch is one characteristic which is modified in speech (67-71). Ohara (49,50) has found that Japanese women have lower pitch in English than in Japanese, but for men the difference between Japanese and English was not as clear. She argues that this may be due to the fact that in Japanese culture the women are expected to manifest their femininity by a high pitch, but it is not as important when they speak English. Ohara also argues that some of the women who learn Japanese as L2 try to change their pitch to meet the expectation of a higher pitch in Japanese. Boka (156) has found in his preliminary study that the fundamental frequency is, according to the results, affected by the language shift. In the study only one woman was studied (Japanese L1, English L2), so vast generalizations cannot be made. Surprisingly, the subject had a lower F0 in Japanese than in English, which Boka argues to be a result of the fact that the speaker was grown up in a Western society and did not, then, grow up with the Japanese gender expectations. He argues that, if changing the F0 is a conscious choice, after leaving the native speaking country (in this case Japan) and living in a country without the expectations and social constraints it will eventually stop maintaining the typical high pitch.

In this present study, the subjective notions of change in pitch in L2 and the actual change in F0 did not seem to be correlated, so one might ask whether the change in pitch is a conscious choice at all. Perhaps the rise in F0 is more related to the possibility that the speakers were somewhat more unconfident in speaking the foreign language than the native one, and it may have resulted in the rise in F0. Previous studies (87,88) have shown that lack of confidence, or increased psycho-physiological stress (84) may have an effect on F0 causing it to rise. The text reading as a task may also have an effect on F0 (157), so another kind of task, such as spontaneous speech samples, should be studied.

The range of F0 did not show changes between L1 and L2 which can be due to the fact that the range of F0 can be relatively constant in speech and therefore can be difficult to modify in the language shift. Also, it is possible that the range was affected by the native language, since the pattern for intonation in the native language is easily shifted in the foreign language speech (86). For example, the Finnish speakers of L2 have been shown to have difficulties in achieving a prosody or intonation of the target language (see e.g. 158,159).

The lower alpha ratio and the rise in L1-L0 level difference in L2 than in L1 may indicate a more pressed phonation (104,160), which is regarded increasing the risk of vocal loading and voice disorders (161). However, the L1-L0 level difference may be affected by the different formant frequencies in the two languages (47) (see Figure 4, page 33 and Figure 5, page 34) which make the acoustical voice quality analyses somewhat hard to make. Earlier studies suggest that language and/or culture may affect the spectral structure (47,48,81,93-98) and, therefore, the change in alpha ratio and the L1-L0 level difference between languages may also be resulted from the language shift itself, as, according to Majewski et al. (55), the first formant is mostly affected by the language. This has been taken into account in the analyses of alpha ratio by setting the border of low and high frequencies at 1500 Hz. By this, the effect of the first formant differences between Finnish and English should be eliminated from the analyses. Also, it has been previously stated that a speech sample of adequate length (at least one minute) should ensure that the individual sounds of the language do not affect the long time average spectrum, and with LTAS it is possible to study the laryngeal features of a speaker rather than the subglottal ones (90).

The effect different languages have on LTAS, alpha ratio, and L1-L0 level difference needs to be studied further with a larger number of subjects as well as with several languages. Also, the correlation with L1-L0 level difference changes and changes in F0 should be taken into account. Also, perceptual analyses and their correlation with the spectral based parameters should be added. In this study, the perceptual analyses were aimed to investigate the proficiency of the L2 speaker.

Significant changes were found only in the dose measures normalized to Time dose (Dt) in Vocal Loading Index (VLI) in text reading. These changes correspond to changes in F0 (32), as expected. Speaking a foreign language, though, may increase vocal loading, as a trend was found that the mean exposure per second tends to be higher in L2 than in L1. This trend towards increased vocal loading was clearer for Finnish and female speakers, and it may be resulted from the larger changes in F0. As previous studies have shown the change in F0 may be gender and language dependent as, according to the results, the change in F0 between languages is more distinctive with women than men (49,50), and more distinctive with Finnish than English subjects.

People may have differences in sensitivity to adaptation, as experience in speaking the target language did not show correlation with the actual changes in F0. It would have been expected that the more experienced the subjects felt in speaking the target language, the more clearly he/she would have adapted to the F0 characteristics of the target language, had the changes in F0 been clearly a result from adaptation. It may also be that quantifying the overall experience in the target language is not very easy, as in this study it was considered as the subjective notion of one´s own perspectives on experience level in the target language. Maybe native listeners´ evaluation on the experience of the speakers would have given different results. Also, the subjects varied extensively in their education and residence. These factors should be considered separately.

6.2 Vocal fatigue and vocal loading in speaking a foreign language

Some evidence was found in the acoustical parameters indicating vocal loading that speaking L2 may, in fact, be more loading than speaking L1. A specific cause of vocal fatigue is not always obvious, and acoustical changes do not always correlate with the subjective sensations of vocal fatigue (104). In comparison of L1 and L2 an increase in effort in phonation and/or articulation may cause an increase in vocal overloading and, thus, increase the speakers´ symptoms of vocal fatigue. The stressfulness of the task of speaking L2 may increase the mental effort itself which may, then, lead to an increased subjective sensation of vocal overloading (161). Also, it is possible that the subjects are not always able to separate the tiredness of the articulators from the tiredness of the voice. The subjects reported that their voices get tired faster in L2 than in L1, but the question did not specify how tiredness of the voice should be understood. In addition to voice getting tired faster, the subjects reported the quality of voice as poorer in L2 than in L1.

Excessive vibration, in terms of number and amplitude of vocal fold collision, may cause damage to the vocal fold tissue, and, therefore, F0, SPL and the duration of voiced speech are considered to be major factors in vocal loading (32). The rise in F0 was significant in L2, and that, then, may contribute to the vocal loading and to the subjective sensation of vocal fatigue. Previous studies have shown that the amount of voiced speech is approximately 50 percent of the total amount of speech (162), and it tends to be lower in the foreign language than in the native one (32). The results in the present study were in line with the previous findings. According to Spilková and van Dommelen (163), word duration is longer for L2 speakers than L1 speakers, which can have a relation to the fact that speech rate is slower in L2 than in L1, and that can result in a lower amount of voiced speech of the total amount of speech. This, in turn, may compensate the increase of vocal loading due to the rise of F0. The dose measures are based on F0, SPL, and total amount of voiced speech in the sample (31,32) and, therefore, not all that suitable for studying the vocal loading in two languages. Some other means for quantifying vocal loading should be considered. Voice quality investigated by inverse filtering, as well as perceptual analyses could give additional information on vocal loading.

The negative correlation between self-reported vocal fatigue and the level of experience in L2, and the native listeners´ evaluations of proficiency in the target language may express that the less experienced speakers tend to experience the task of speaking L2 more stressful or that they use more effort in the L2 production than the more experienced speakers and, thus, have more symptoms of vocal fatigue. Lack of experience, then, may increase their sensitivity of noticing symptoms of vocal and/or articulatory overloading.

The subjective sensation of vocal fatigue does not always correlate with the objective measurements of acoustical parameters (104), and a specific reason for vocal fatigue can sometimes be hard to find. Here, the acoustical parameters obtained through inverse filtering did not show differences between the subjects who reported more vocal fatigue in L2 (Group 2) from those subjects´ parameters who did not report more vocal fatigue (Group 1). The subjects in Group 2 had, though, less optimal voice use (i.e. poorer voice quality and more pressed phonation) also in L1. This may indicate that the sensation of vocal fatigue was affected by the voice use in both L1 and L2.

The subjects in Group 1 were more experienced in speaking L2 than the subjects in Group 2, which may cause increased psycho-physiological stress and mental effort with subjects in Group 2, which may, then, lead to increased muscle tension and therefore cause F0 to rise and pressedness in voice to increase. These factors have been reported to contribute to the increase of subjective sensation of vocal overloading (161,164). However, a clear answer that F0 rose or that NAQ lowered more with the subjects in Group 2 than with the subjects in Group 1 was not found. The fact that F0 was lower in L2 for subjects in Group 2 can be due to the composition of the groups, since nine of the twelve subjects were native English speakers in Group 2 while in Group 1 nine of the subjects were native Finnish speakers. Languages may differ in the mean F0 (53), and it is possible that the subjects tried to reach a level of pitch they thought the native speaker of the target language would use (26,49) which, then, affected the results in F0. The composition of groups was made by the subjective sensations of vocal fatigue, and, therefore, it was impossible to manage that both groups would have the same amount of native Finnish and native English speakers. To exclude the effect of different composition of groups, a further study with more subjects is warranted.

When the inverse filtering parameters of Group 1 and Group 2 were examined separately, a trend towards increased pressedness in L2 was found. It would have been expected that it had been significant for Group 2, but, surprisingly, it was significant only for Group 1. This may be an indication of that the subjects in Group 1 were not as familiar with the symptoms of vocal overloading as the subjects in Group 2, or possibly the acoustical parameters investigated here did not reveal the subjective sensations of vocal fatigue (104). The possibility of the lack of familiarity of the symptoms is in line with Rantala (165) who has found, that the rise of F0 and levelling of the spectrum appeared with subjects with less complaints of vocal fatigue symptoms rather than subjects with more complaints. According to Rantala this may reflect a normal adaptation of the human body, as the subjects with more complaints, in order to avoid excessive strain and exhaustion, tend to shift their voice use towards more hypofunctional (165).

The length of residence in a country speaking the foreign language (66) and, also, other factors such as the age when the foreign language studies have begun (9), or the amount of language use (16), may affect the performance in the foreign language. For the English subjects a correlation between residence and self-evaluated experience in L2, and for the Finnish subjects a correlation between self-evaluated experience and education was found. These findings may indicate that the subjective notion of experience in using the L2 (or perhaps the self-confidence in it) is gained differently with the native Finnish and native English speakers, which may, in fact, be a direct result from the fact that the Finnish and English subjects had learned the L2 in quite different ways. The Finnish had mostly been taught English in formal educational settings while the English had learned Finnish in more informal ways.

The differences in language learning between Finnish and English subjects, then, may result to the fact that the Finnish subjects´ changes in voice between L1 and L2 are somewhat more conscious than those of the English subjects´. (42.) On the other hand, in Finnish schools English is mostly taught be native Finnish speaking teachers which may lead to mispronunciation of English (166), and maybe this could be the case in non-verbal cues, too, but Finnish people tend to have quite a broad knowledge of native English speakers´ English, since it can be heard very extensively in the media. Therefore, it cannot be stated that the Finnish subjects had learned the English language solely by Crystal´s behaviouristic model. Also, some of the English subjects did have some formal education in the Finnish language and, then, their language learning model was perhaps not solely cognitive (42). Such distinctions between the two models by Crystal are hard to make in this kind of study. In order to find the possible distinctions the language learning model have on the changes in voice between L1 and L2, it would be favorable to recruit the L2 Finnish speakers from another country.

This study did not take the age when the L2 learning had begun or the amount of the L2 use into account. It is possible that these factors could have given more detailed information about the subjects´ actual experience in L2. In future research, it seems essential to include these factors to the study. Also, longer speech samples should be added. Maybe even conversational speech samples would be beneficial, since the type of the speaking task, i.e. text reading versus spontaneous speech, may have an effect on acoustical parameters, such as fundamental frequency (82,159). However, no indication that the subjects´ voice use was largely affected by the task was found here.

According to the results, a possibility that language shift from L1 to L2 may influence the speaker´s vocal use causing voice problems in the course of time has been found, and, therefore, voice coaching in L2 would be beneficial especially to people who use foreign languages to a large extent, e.g. professionally.

6.3 Voice quality in foreign language

The phonation type in L2 seemed to differ from that of L1. Between L2 and L1 differences were obvious in the perceptual analyses as the voices were perceptually evaluated as having poorer voice quality, more pressed and strenuous voice production, and higher pitch in L2 than in L1.

It may be speculated that the strong accent or mispronunciation of some of the L2 speakers may have affected the assessment of voice quality and strenuousity of voice production, as some listeners stated that the mispronunciation or strong accent was sometimes easily noticeable. However, the listeners were experienced in perceptual analyses of voice and should be able to ignore the influence of language and fluency of speech. The language background of the listeners may influence the perceptual evaluation of some vocal features, for example, asthenia and strain (167), and roughness (168) have been reported to be evaluated differently by listeners from different languages. Neither of the studies found that language background had any effect on the perceptual ratings of breathiness in GRBAS scale. Also, a recent study (169) on perceptual analyses between native speakers of Finnish and Brazilian Portuguese has found no significant differences between listeners´ evaluations on voice quality (169,170). In this study, the listeners were native Finnish speakers, and the cultural and language background was homogenous in a way that it avoided the possible influence of cultural and language background on the perception of voice quality (171). However, further study on perceptual analyses of voice quality with native English listeners is warranted to ensure that the language background has no effect on evaluations. The scale used in this study was developed for the purposes of the study, and the evaluated features were overall voice quality, voice production, firmness, and pitch suitability, since these features can be expected to change in the language shift (Study I, Study II), and, because the speakers had normal voices instead of pathological voices, a GRBAS scale would not have been suitable. The scale used in the study was from zero to ten. Previous studies have shown that a numerical scale can be as suitable as a visual analog scale (VAS) (141), and in this study the scale was even broader than the traditional GRBAS scale, which is from zero to three (140,141). By this, the subtle changes were perhaps more easily detectable. Also, the numerical scale was easier to fill, the perceptual analyses were made entirely on line.

Between L1 and L2 acoustical differences were found. A decrease in NAQ and ClQ in L2 compared to that of L1 point to increased pressedness of voice in L2 (114-117,120). Also, the acoustical analyses and the perceptual evaluations of voice quality correlated with each other, when acoustical parameters indicated a more pressed voice the voice production was evaluated also as more strenuous. Furthermore, voices with evaluated higher pitches had a correlation with a decrease in AQ, which indicate that a raised pitch and increased pressedness of voice were at least in some cases related. Pressed phonation may increase vocal loading (114) which may result in increased subjective sensations of vocal fatigue.

Important factors in vocal loading are also voicing time, F0, and SPL (29), a more pressed phonation does not necessarily singly indicate total vocal loading. Since the amount of voiced speech can vary between languages (Study II), the acoustical parameters based on voicing time, F0 and intensity, such as vocal doses (31,32) or Vocal Loading Index (122), are not sufficient in studying vocal loading between two languages, and, therefore, perceptual voice quality ratings and investigation of the voice source characteristics are important in studying differences in vocal loading between L1 and L2. It is, though, important to notice that when F0 rises it affects the length of the fundamental period, and for that reason, the normalized amplitude quotient (NAQ) may be more suitable for studying the vocal loading than AQ (119).

The subjects in Group 2 had perceptually evaluated poorer voice quality, more pressed phonation, more strenuous voice production, and higher pitch also in L1 than subjects in Group 1 (with the exception of strenuousity in text reading and of firmness in spontaneous speech which was the same for Group1 and Group 2). This may lead to the fact that the subjects with less optimal voice use in L1 possibly were selected to Group 2, and that these subjects also experience more symptoms of vocal fatigue even with smaller changes in L2 than the subjects in Group 1. This question should be addressed in further studies by studying a larger number of subjects in order to possibly avoid the differences in voice use L1.

Speakers have articulatory and phonatory settings, and the neutral phonatory settings are composed of modal phonation in voicing, that is, the average muscular tension throughout the whole vocal apparatus is moderate (172). Speaking L2 seems to set additional challenges to the vocal use that may lead the L2 speakers to use a less optimal voice and increase the sensitivity to vocal overloading as well as vocal fatigue, and it also causes changes in the voice that are perceptually perceivable. Therefore, people who must speak extensively in L2 could benefit from teaching in speech technique in L2, not only because the amount of voice use is large, but also because speaking L2 increases the vocal overloading compared to speaking L1. Due to the different structures and phonatory and articulatory settings in L2 compared to that of L1, the speech technique in L1 may be insufficient, since the technique from L1 does not always transfer to L2.

7 CONCLUSIONS OF STUDIES

7.1 Language shift and its effect on voice characteristics

The mean F0 changed, but the standard deviation or range of F0 did not change in L2 compared to L1, and the changes were to some extent related to the target language. The reasons for the changes in F0 may be due to adaptation to a certain level the speakers consider more native like, or it can be resulted from lack of confidence and uncertainty. People differ in their ability to observe their own voices and the changes in it, as the subjective evaluation of change in pitch between L1 and L2 did not always correspond to the actual changes in F0. F0 range of voice is probably very difficult for a language learner to modify which explains the results of the standard deviation and the range of F0.

The self-estimated experience in L2 was not related to the actual changes in the mean F0, which supports the idea of adding perceptual analyses of e.g. fluency to the study as an evaluation of speakers´ experience in the target language. The quantification of overall experience in L2 should be further studied.

Also, it would be beneficial to study larger number of subjects as well as multiple language pairs.

7.2 Language shift and its effect on vocal fatigue and vocal loading

The results showed that majority of the subjects experienced more vocal fatigue in L2 than in L1. However, the vocal doses, as such, did not show more loading in L2 than in L1, which is due to the lower amount of voiced speech in L2 than in L1. Vocal doses can be used in quantifying vocal loading in two languages but they need to be normalized with the amount of voiced speech, since a trend was found that the mean exposure per second was higher in L2 than in L1. Also, the trend of L1-L0 level difference indicated that the speakers´ voice use changed towards more pressed, which supported the idea of including analyses of voice quality measured by acoustical and perceptual analyses in the study of vocal loading. Also, longer speech samples should be studied, since 1 minute does not necessarily reveal long term changes in the voice use.

7.3 Language shift and its effect on voice quality

Language shift from L1 to L2 caused voice quality changes that are perceptually observable, and acoustical findings correlated with them. Perceptual analyses showed that speakers´ voices were evaluated as more pressed in L2 than in L1, and acoustical analyses revealed a decrease in NAQ and ClQ in L2 compared to L1 indicating increased pressedness of voice. Also, subjects who reported more vocal fatigue in L2 than in L1 were evaluated for the most part as having poorer voice quality, more pressed and strenuous voice quality, and higher pitch than suitable in both L1 and L2 than the speakers who did not report having more vocal fatigue in L2 than in L1.

7.4 Conclusions

Speaking a language other than the native one changes the speakers´ voice production towards a more suboptimal, in terms of higher pitch and more pressed phonation, causing symptoms of vocal fatigue and possible vocal overloading. In the study of changes in voice use between two languages perceptual analyses as well as acoustical analyses through inverse filtering seem to be the most efficient tool for revealing the changes. These methods are least influenced by the different characteristics of languages. Inverse filtering suits for research since it ignores the articulatory differences in the vocal tract between languages. Listening tests are supported as perceptual evaluations are in line with the acoustical findings through inverse filtering. Also, spectral analyses may be useful, but levels must be set so that the changes are not affected by e.g. different formant frequencies in the languages.

Speakers of foreign languages would benefit from having vocal training in the foreign language. The optimal voice use in L1 does not necessarily shift to voice use in L2. The group of people who could have this useful is extremely vast, since it has been estimated that over half of the world´s population uses two or more languages in their everyday lives (173). Communication between languages and cultures should include training in speech technique.

REFERENCES

1. Työelämässä tarvitaan yhä useampia kieliä. EK:n henkilöstö- ja koulutustiedustelu 2009. (Several languages are needed in the work life. Personnel and training inquiry by EK in 2009.) EK 2010. http://ek.fi/wp-content/uploads/Henko2009_ Tyoelamassa_ tarvitaan_yha_ useampia_ kielia.pdf, accessed March 29th 2016. (In Finnish.)
2. Izdebski K. Clinical voice assessment: The role & value of the Phonatory Function Studies. In Lalwani A. Current diagnosis & treatment. Otolaryngology head and neck surgery 2007. New York, McGraw Hill Lange, 416-429.
3. Gebhard S. 2000. Masters of the voice. http://aiic.net/page/130/masters-of-the-voice/lang/1, accessed March 30th 2016.
4. http://blogs.helsinki.fi/translationstudies/
5 https://www10.uta.fi/opas/koulutus.htm?opsId=145&uiLang=en&lang=en&lvv=2015&koulid=195
6. Opetushallitus 2006. Näyttötutkinnon perusteet. Asioimistulkin ammattitutkinto 2006. (Board of Education 2006. Basis for competence test. Professional degree for community interpreter.) http://www.oph.fi/download/110919_asioimistulkki_ammattitutkinto.pdf, accessed March 30th 2016. (In Finnish.)
7. Reflection Forum on Multilingualism and Interpreter Training. Final Report. 2009 European Commission, http://ec.europa.eu/dgs/scic/docs/finall_reflection_forum_ report_en.pdf, accessed April 5th 2016
8. Flege JE, Schirru C, MacKay IRA. Interaction between the native and second language phonetic subsystems. Speech Communication 2003; 40: 467–491.
9. Flege JE, MacKay, IRA, Meador D. Native Italian speakers´ perception and production of English vowels. J Acoust Soc Am 1999; 106, 5: 2973-2987.
10. Flege JE, Bohn O-S, Jang S. Effects of experience on non-native speakers´ production and perception of English vowels. J of Phonetics 1997; 25: 437-470.
11. Ingram JCL, See-Gyoon P. Cross-language vowel perception and production by Japanese and Korean learners of English. J of Phonetics 1997; 25: 343-370.
12. Munro MJ. Productions of English vowels by native speakers of Arabic: Acoustic measurements and accentedness ratings. Language and Speech 1993; 36, 1: 39-66.
13. Flege JE, Hillenbrand J. Limits on phonetic accuracy in foreign language speech production. J Acoust Soc Am 1984; 76: 708-721.

14. Levy ES, Strange W. Perception of French vowels by American English adults with and without French language experience. J of Phonetics 2008; 36: 141–157.
15. Peltola MS, Tuomainen O, Koskinen M, Aaltonen O. The Effect of Language Immersion Education on the Preattentive Perception of Native and Non-native Vowel Contrasts. J Psycholinguist Res, 2007; 36: 15–23.
16. Flege JE, MacKay IRA. Perceiving vowels in a second language. Studies in Second Language Acquisition 2004; 26, 1: 1-34.
17. Aaltonen O, Eerola O, Hellström Å, Uusipaikka E, Lang AH. Perceptual magnet effect in the light of behavioral and psychophysiological data. J. Acoust. Soc. Am. 1997; 101, 2: 1090-1105.
18. Aaltonen O. 1997. Vowel perception: behavioural and psychophysiological experiments. Academic dissertation. Turun yliopiston julkaisuja. Sarja D, Medica - Odontologica, 0355-9483; 261. University of Turku. Finland: Turku.
19. Morrison GS. 2006. L1 L2 Production and Perception of English and Spanish Vowels. A Statistical Modelling Approach. Doctoral thesis. Department of Linguistics. University of Alberta. Canada: Edmonton.
20. Rauber AS, Escudero P, Bion RAH, Baptista BO. The Interrelation between the Perception and Production of English Vowels by Native Speakers of Brazilian Portuguese. Proceedings of Interspeech 2005; 2913-2916, http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.422.8294&rep=rep1&type=pdf, accessed April 11th 2016.
21. Ylinen S, Uther M, Latvala A, Vepsäläinen S, Iverson P, Akahane-Yamada R, Näätänen R. Training the brain to weight speech cues differently: a study of Finnish second-language users of English. J Cogn Neurosci 2009; 22: 1319–1332.
22. Peltola MS, Kujala T, Tuomainen J, Ek M, Aaltonen O, Näätänen R. Native and foreign vowel discrimination as indexed by the mismatch negativity MMN response. Neuroscience Letters 2003; 352, 1: 25–28.
23. Toivanen J. 1999. Perspectives on intonation: English, Finnish and English spoken by Finns. Academic dissertation. Oulu: University of Oulu.
24. Wiik K. 1965. Finnish and English Vowels. Academic dissertation. Turun yliopiston julkaisuja B: 94. Turku: University of Turku.
25. Hakulinen, L. 1979. Suomen kielen rakenne ja kehitys. (Finnish language, structure and development.) (4.ed.) Helsinki: Otava. (In Finnish.)
26. Järvinen K, Laukkanen AM, Izdebski K. Voice Fundamental Frequency Changes as a Function of Foreign Languages Familiarity: An Emotional Effect? In Izdebski K. (Ed.) 2008. Emotions in the Human Voice 1. USA: Plural Publishing. 203-213.
27. Järvinen K, Laukkanen AM. Native Finnish and English Speakers' Fundamental Frequency, Equivalent Sound Level, and Long-Time Average Spectrum Characteristics in Text-reading in Finnish and English, http://digi.lib.ttu.ee/i/?802&, accessed April 11th 2016.
28. Järvinen K, O´Dell M, Aaltonen O. Vowel Characteristics in Native and Foreign Languages. XXVII Phonetics symposium 2012, Proceedings,

http://fp2015.aalto.fi/submission.html, accessed April 11th 2016.
29. Vilkman E. Occupational Safety and Health Aspects of Voice and Speech Professions. Folia Phoniatr Logop 2004; 56: 220–253
30. Jiang JJ, Titze IR: Measurement of vocal fold intraglottal pressure and impact stress. J of Voice 1994; 8, 2: 132-144.
31. Švec JG, Popolo PS, Titze IR. Measurement of vocal doses in speech: experimental procedure and signal processing. Logoped Phoniatr Vocol 2003; 28: 181-192.
32. Titze IR, Švec JG, Popolo PS. Vocal Dose Measures: Quantifying Accumulated Vibration Exposure in Vocal Fold Tissues. J of Speech, Language, and Hearing Research 2003; 46: 919–932.
33. Aaltonen O. Puhe kommunikaatiomuotona ja tutkimuskohteena, (Speech as a form of communication and an object of research.) Puhe ja kieli (Speech and Language) 2008; 28, 2: 85–94. (In Finnish, abstract in English.)
34. Aaltonen O. Puheen luonne ja kehitys. (The nature and development of speech.) In Aaltonen O, Aulanko R, Iivonen A, Klippi A, Vainio M. (Eds.). 2009. Puhuva ihminen. (The speaking human.) Keuruu: Otava. 10-18. (In Finnish.)
35. Korhonen M. 1993. Kielen synty. (Language and its origin.) Juva: WSOY. (In Finnish.)
36. Laver J, Trudgill P. Phonetic and linguistic markers in speech. In Scherer KR, Giles H. (eds.) 1979. Social markers in speech. GB: Cambridge University Press. 1-32.
37. Lehtonen J, Hurme P. Puheketju ja puheentutkimus. (Speech chain and speech research.) In Sajavaara K. (Ed.). 1980. Soveltava kielitiede. (Applied linguistics.) Huhmari: Gaudeamus. 78-96. (In Finnish.)
38. Itkonen E. 1966. Kieli ja sen tutkimus. (Language and its research.) Helsinki: WSOY. (In Finnish.)
39. Langacker RW. Conceptualization, Symbolization, and Grammar. In Tomasello M. (ed.) The New Psychology of Language. 1998. Cognitive and Functional Approaches to Language Structure. USA: Lawrence Erlbaum Associates, Inc. 1-40.
40. Givón T. The Functional Approach to Grammar. In Tomasello M. (ed.) 1998. The New Psychology of Language. Cognitive and Functional Approaches to Language Structure. USA: Lawrence Erlbaum Associates, Inc. 41-66.
41. Lyons J. 1968. Introduction to Theoretical Linguistics. Great Britain: Cambridge University Press.
42. Crystal D. 1987. The Cambridge Encyclopedia of Language. GB: Cambridge University Press.
43. Gleitman L, Papafragou A. Language and Thought. In Holyoak KJ, Morrison RG (Eds.) 2005. The Cambridge Handbook of Thinking and Reasoning. New York: Cambridge University Press.
44. Boroditsky L. Does Language Shape Thought?: Mandarin and English Speakers´ Conceptions of Time. Cognitive Psychology 2001; 43: 1–22.
45. Karlsson F. 1994. Yleinen kielitiede. (General Linguistics.) Helsinki: Yliopistopaino. (In Finnish.)

46. Kangasvieri T, Miettinen E, Kukkohovi P, Härmälä M. Kielten tarjonta ja kielivalintojen perusteet. Tilannekatsaus joulukuu 2011. Opetushallituksen muistio 201:3. (Supply of languages and basis for language choices in basic education. Status survey December 2011. Memo of Finnish National Board of Education 201:3.), http://www.oph.fi/download/138072_Kielten_tarjonta_ja_kielivalintojen_perusteet_perusopetuksessa.pdf, accessed March 30th 2016. (In Finnish.)
47. Wagner A, Braun A. Is voice quality language-dependent? Acoustic analyses based on speakers of three different languages. 15th ICPhS Barcelona 2003: 651-654, https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2003/papers/ p15_0651. pdf, accessed April 11th 2016.
48. Andrianopoulos MV, Darrow K, Chen J. Multimodal standardization of voice among four multicultural populations formant structures. J Voice 2001; 15, 1: 61-77.
49. Ohara Y. Performing gender through voice pitch: A Cross-Cultural analysis of Japanese and American English. In Pasero U, Braun F. (eds.) 1999. Perceiving and Performing Gender. Opladen/Wiesbaden: Westdeutscher Verlag. 105-116.
50. Ohara Y. Gender-dependent pitch levels: A Comparative study in Japanese and English. In Hall K, Bucholtz M, Moonwomon B. (Eds.) 1992. Locating power. Proceedings of the second Berkeley Women and Language Conference. Berkeley, CA: Berkeley Women and Language Group. 2: 468-477.
51. Bruyninckx M. Language-induced voice quality variability in bilinguals. J of Phonetics 1994; 22: 19-31.
52. Sundberg J. 1988. The Science of the Singing Voice. Illinois. USA: Northern Illinois University Press.
53. Pegoraro Krook MI. Speaking Fundamental Frequency Characteristics of Normal Swedish Subjects Obtained by Glottal Frequency Analysis. Folia Phoniatrica 1988; 40: 82-90.
54. Esling JH, Wong RF. Voice Quality Settings and the Teaching of Pronunciation. Tesol Quarterly 1983; 17, 1: 89-95.
55. Majewski W, Hollien H, Zalewski J. Acoustic comparisons of American English and Polish. J of Phonetics 1977; 5: 247-251.
56. Laukkanen AM, Leino T. 1999. Ihmeellinen ihmisääni. (The Amazing Human Voice.) Tampere: Gaudeamus. (In Finnish.)
57. Valo M. 1994. Käsitykset ja vaikutelmat äänestä. Kuuntelijoiden arviointia radiopuheen äänellisistä ominaisuuksista. (Perceptions and appearances of voice. Listeners´ evaluation on vocal characteristics in radio speech.) Academic dissertation. Studia Philologica Jyväkylaensia 33. Jyväskylä: University of Jyväskylä. (In Finnish, abstract in English.)
58. Ylinen S, Uther M, Latvala A, Vepsäläinen S, Iverson P, Akahane-Yamada R, Näätänen R. Training the brain to weight speech cues differently: a study of Finnish second-language users of English. J Cogn Neurosci 2009; 22: 1319–1332.

59. Suomi K, Toivanen J, Ylitalo R. 2008. Finnish Sound Structure. Phonetics, Phonology, Phonotactics and Prosody. Studia Humaniora Ouluensia. Oulu: University of Oulu.
60. Vihanta VV. Suomi vieraana kielenä foneettiselta kannalta. (Finnish as a foreign language from the phonetic point of view.). In Tommola J (Ed.). 1990. Vieraan kielen ymmärtäminen ja tuottaminen (Foreign Language Comprehension and Production). AFinLA Yearbook. 199–225, https://helda.helsinki.fi/bitstream/handle/10224/3668/vihanta199-225.pdf?sequence=2, accessed March 30th 2016. (In Finnish, abstract in English.)
61. Flege JE. Second language speech learning: Theory, findings, and problems. In Strange W. (Ed.) 1995. Speech perception and linguistic experience: Issues in cross-language research. Timonium, MD: York Press. 233-276.
62. Best CT. The Emergence of Native-Language Phonological Influences in Infants: A Perceptual Assimilation Model. Haskins Laboratories Status on Speech Research. SR 1991; 107, 108: 1-30.
63. Flege JE. Interactions between the native and second-language phonetic systems. In Burmeister P, Piske T, Rohde A. (Eds.) 2002. An integrated view of language development: Papers in honor of Henning Wode. Trier, Germany: Wissenschaftlicher Verlag Trier. 217-243.
64. Flege JE. Age of learning and second language speech. In Birdsong D. (Ed.) 1999. Second language acquisition and the Critical Period Hypothesis. Mahwah, NJ: Lawrence Erlbaum Associates. 101-131.
65. Jia G, Aaronson D. A longitudinal study of Chinese children and adolescents learning English in the United States. Applied Psycholinguistics 2003; 24: 131-161.
66. Flege JE, Liu S. The effect of experience on adults´ acquisition of a second language. Studies in Second Language Acquisition 2001; 23: 527-552.
67. Giles H. Accent mobility: A model and some data. Anthropological Linguistics 1973; 15: 87-105.
68. Littlejohn S.W. 1999. Theories of Human Communication. (6th ed.) USA: Wadsworth Publishing Company. 107-108.
69. Giles H, Coupland N, Coupland J. Accommodation Theory: Communication, Context, and Consequence. Contexts of Accommodation. Developments in Applied Sociolinguistics 1991; 1-68.
70. Burgoon J, Stern L, Dillman L. 1995. Interpersonal adaptation: dyadic interaction patterns. Cambridge, UK: Cambridge University Press.
71. Giles H, Mulac A, Bradac J, Johnson P. Speech accommodation theory: the first decade and beyond. In: McLaughlin ML. (Ed.) Communication yearbook, 10. 1987. London, UK: Sage.13-48.
72. Helfrich H. Age markers in speech. In Scherer KR, Giles H. (Eds.) 1979. Social markers in speech. GB: Cambridge University Press. 63-108.
73. http://jimflege.com/Research.html, accessed April 5th 2016

74. Lavner Y, Gath I, Rosenhouse J. The effects of acoustic modifications on the identification of familiar voices speaking isolated vowels. Speech Communication 2000; 30: 9-26.
75. Titze IR. 2000. Principles of Voice Production. 2nd ed. (National Center for Voice and Speech) Iowa City: Prentice Hall.
76. Laver J. 1980. The Phonetic Description of Voice Quality. Cambridge: Cambridge University Press.
77. Baken RJ. 1987. Clinical Measurement of Speech and Voice. Boston: College-Hill Press, A Division of Little, Brown and Company.
78. Mennen I, Schaeffer F, Docherty G. Pitching it differently: a comparison of the pitch ranges of German and English speakers. Proceedings 16th ICPhS, Saarbrücken 2007: 1769-72,
http://www.icphs2007.de/, accessed April 5th 2016.
79. Burkhardt F, Audibert N, Malatesta L, Türk O, Arslan L, Auberge V. Emotional Prosody - Does Culture Make A Difference? Speech Prosody 2006, http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.493.729&rep=rep1&type=pdf, accessed April 5th 2016.
80. Hirst D. Pitch parameters for prosodic typology. A preliminary comparison of English and French. Proceedings 15th ICPhS, Barcelona 2003,
https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2003/papers/ p15_1277.pdf, accessed April 5th 2016.
81. Awan SN, Mueller PB. Speaking fundamental frequency characteristics of white, African American and Hispanic kindergartners. J Speech and Hear Res 1996: 39: 573-577.
82. Hudson AI, Holbrook A. Fundamental Frequency Characteristics of Young Black Adults. Spontaneous Speaking and Oral Reading. J Speech and Hear Res 1982; 25: 25-28.
83. Laukkanen AM, Mäki E, Pukander J, Anttila I. Vertical laryngeal size and the lowest tone in the evaluation of the average fundamental frequency (F0) of Finnish speakers. Log Phon Voc 1999; 24: 170-177.
84. Orlikoff RB, Baken RJ. The Effect of the Heartbeat on Vocal Fundamental Frequency Perturbation. J Speech and Hear Res 1989; 32: 576-582.
85. Flege J. Factors Affecting degree of perceived foreign accent in English sentences. J Acoust Soc Am 1988; 84: 70-79.
86. Grover C, Jamieson DG, Dobrovolsky MB. Intonation in English, French and German: Perception and Production. Language and Speech 1987; 30, 3: 277-295.
87. Ohala J. Cross-Language Use of Pitch. An Ethological View. Phonetica 1983; 40: 1-18.
88. Ohala J. An Ethological Perspective on Common Cross-Language Utilization of F0 of Voice. Phonetica 1984; 41: 1-16.
89. Bauer HR. Frequency Code: Orofacial Correlates of Fundamental Frequency. Phonetica 1987; 44: 173-191.

90. Nolan F. 1983. The Phonetic Bases of Speaker Recognition. Great Britain: Cambridge University Press.
91. Cleveland TF, Sundberg J, Stone RE. Long-Term-Average Spectrum Characteristics of Country Singers During speaking and Singing. J Voice 2001; 15, 1: 54–60.
92. Mendoza E, Valencia N, Muños J, Trujillo H. Differences in Voice Quality Between Men and Women: Use of the Long-Term Average Spectrum LTAS. J Voice 1996; 10, 1: 59-66.
93. Rallo Fabra L, Romero J. Native Catalan learners´ perception and production of English vowels. J of Phonetics 2012; 40: 491–508.
94. Baker W, Trofimovic P, Flege JE, Mack M, Halter R. Child–Adult Differences in Second Language Phonological Learning: The Role of Cross-Language Similarity. Language and Speech 2008; 51, 4: 317-342.
95. Chen Y, Robb M, Gilbert H, Lerman J. Vowel production by Mandarin speakers of English. Clinical Linguistics & Phonetics 2001; 15, 6: 427-440.
96. Kuronen, M. 2000. Vokaluttalets akustik i sverigessvenska, finlandssvenska och finska. Vowel pronunciation in Swedish, Swedish in Finland and Finnish. Academic dissertation. University of Jyväskylä: Studia Philologica Jyväskyläensia 49. (In Swedish.)
97. Arslan LM, Hansen JHL. A Study of Temporal Features and Frequency Characteristics in American English Foreign Accent. J Acoust Soc Am 1997; 102, 28: 28-40.
98. Walton JH, Orlikoff RF. Speaker race identification from acoustic cues in the vocal signal. J Speech Hear Res 1994; 37: 738-745.
99. O´Neil EN, Jones GW, Nye C. Acoustic characteristics of children who speak Arabic. International Journal of Pediatric Otorhinolaryngology 1997; 42: 117-124.
100. Hanley TD. Snidecor JC, Ringel RL. Some Acoustic Differences among Languages. Phonetica 1966; 14: 97-107.
101. Byrne D. et al. An international comparison of long-term average speech spectra. J Acoust Soc Am 1994; 96, 4: 2018-2120.
102. Esling JH., Wong R.F. Voice Quality Settings and the Teaching of Pronunciation. Tesol Quarterly 1983; 17, 1: 89-95.
103. Frøkjær-Jensen B, Prytz S. Registration of Voice Quality. Brüel and Kjær Technical Review 1976; 3: 3-17.
104. Laukkanen AM, Ilomäki I, Leppänen K, Vilkman E. Acoustic Measures and Self-reports of Vocal Fatigue by Female Teachers. J Voice 2008; 22, 3: 283–289.
105. Laukkanen AM, Syrjä T, Laitala M, Leino T: Effects of two-month vocal exercising with and without spectral biofeedback on student actors´ speaking voice. Log Phon Vocol 2004; 29: 66-76.
106. Rantala L, Paavola L, Körkkö P, Vilkman E. Working-day effects on the spectral characteristics of teaching voice. Folia Phoniatr Logop 1998; 50: 205–211.

107. Wells JC. 1962. A study of the formants of the pure vowels of British English. http://www.phon.ucl.ac.uk/ home/wells/ formants/frequency.htm, accessed Aug 20th 2012.
108. Atal BS. Automatic Recognition of Speakers from their Voices. Proceedings of the IEEE 64 1976; 4: 460-475.
109. Linville SE, Rens J. Vocal Tract Resonance Analysis of Aging Voice Using Long-Term Average Spectra. J Voice 2001; 15, 3: 323-330.
110. Miller RL. Nature of the vocal cord wave. J Acoust Soc Am. 1959; 31: 667–677.
111. Airas M. 2008. Methods and Studies of Laryngeal Voice Quality Analysis in Speech Production. Academic dissertation. Espoo: Finland. Helsinki University of Technology Faculty of Electronics, Communications and Automation Department of Signal Processing and Acoustics.
112. Fant G. 1973. Speech Sounds and Features. Cambridge: MIT Press.
113. Rothenberg M. A new inverse-filtering technique for deriving the glottal air flow waveform during voicing. J Acoust Soc Am. 1973; 53: 1632-1645.
114. Alku P. Glottal wave analysis with pitch synchronous iterative adaptive inverse filtering. Speech Commun. 1992; 11: 109–118.
115. Klatt DH, Klatt LC. Analysis, synthesis, and perception of voice quality variations among female and male talkers. J Acoust Soc Am. 1990; 87; 2: 820–857.
116. Childers DG, Lee CK. Vocal quality factors: Analysis, synthesis, and perception. J Acoust Soc Am. 1991; 90; 5: 2394–2410.
117. Holmberg EB, Hillman RE, and Perkell JS. Glottal airflow and transglottal air pressure measurements for male and female speakers in soft, normal, and loud voice. J Acoust Soc Am. 1988; 84; 2: 511–1787.
118. Alku P, Vilkman E. Amplitude domain quotient of the glottal volume velocity waveform estimated by inverse filtering. Speech Commun. 1996; 18: 131-138.
119. Alku P, Bäckström T, Vilkman E. Normalized amplitude quotient for parametrization of the glottal flow. J Acoust Soc Am. 2002; 112: 701-710.
120. Gobl C, Ní Chasaide. A. Amplitude-Based Source Parameters for Measuring Voice Quality. VOQUAL'03. 2003, http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.385.540&rep=rep1&type=pdf, accessed April 18th 2015.
121. Buekers R, Bierens E, Kingma H, Marres EHMA. Vocal Load as Measured by the Voice Accumulator. Folia Phoniatr Logop 1995; 47; 5: 252-261.
122. Rantala L, Vilkman E. Relationship between subjective voice complaints and acoustic parameters in female teachers′ voices. J Voice 1999; 13; 4: 484–495.
123. Chang A, Karnell MP. Perceived phonatory effort and phonation threshold pressure across a prolonged voice loading task: a study of vocal fatigue. J Voice 2004; 18; 4: 454–466.
124. Lohscheller J, Doellinger M, McWhorter AJ, Kunduk M. Preliminary Study on the Quantitative Analysis of Vocal Loading Effects on Vocal Fold Dynamics Using

Phonovibrograms. Annals of Otology, Rhinology & Laryngology 2008; 117; 7: 484-493.
125. Remacle A, Morsomme D, Finck C. Comparison of Vocal Loading Parameters in Kindergarten and Elementary School Teachers. Journal of Speech, Language, and Hearing Research 2014; 57: 406–415.
126. Bottalico P, Astolfi A. Investigations into vocal doses and parameters pertaining to primary school teachers in classrooms. J Acoust Soc Am. 2012; 131; 4: 2817-2827.
127. Gaskill CS, O'Brien SG, Tinter SR. The Effect of Voice Amplification on Occupational Vocal Dose in Elementary School Teachers. J Voice 2012; 26; 5: 667.e19-667.e27.
128. Morrow SL, Connor NP. Comparison of Voice-Use Profiles between Elementary Classroom and Music Teachers. J Voice 2011; 25; 3: 367-372.
129. Schloneger MJ. Graduate Student Voice Use and Vocal Efficiency in an Opera Rehearsal Week: A Case Study. J Voice 2011; 25; 6: e265-e273.
130. Carroll T, Nix J, Hunter E, Emerich K, Titze I, Abaza M. Objective measurement of vocal fatigue in classical singers: A vocal dosimetry pilot study. Otolaryngology–Head and Neck Surgery 2006; 135: 595-602.
131. ANSI. 1960. USA Standard Acoustical Terminology (Including Mechanical Shock and Vibration). American National Standards Institute. New York: USA.
132. Kreiman J, Sidtis D. (Eds.) 2013. Foundations of Voice Studies. An Interdisciplinary Approach to Voice Production and Perception. Chichester: Wiley-Blackwell.
133. Dahmen JC, King AJ. Learning to hear: plasticity of auditory cortical processing. Curr Opin Neurobiol 2007; 17; 4: 456-464.
134. Kankare E, Laukkanen AM. Quasi-output-cost-ratio, perceived voice quality, and subjective evaluation in female kindergarten teachers. Log Phon Voc 2012; 37: 62-68.
135. Kankare E, Dong L, Laukkanen AM, Geneid A. EGG and acoustic analyses of different voice samples: comparison between perceptual evaluation and voice activity and participation profile. Folia Phoniatr Logop 2013; 65: 98–104.
136. Kreiman J, Gerratt BR. The perceptual structure of pathologic voice quality. J Acoust Soc Am 1996; 100: 1787-1795.
137. Parsa V, Jamieson DG. Acoustic discrimination of pathological voice: Sustained vowels versus continuous speech. J Speech Lang Hear Res 2001; 44: 327-339.
138. Bele IV. Reliability in Perceptual Analysis of Voice Quality. J Voice 2005; 19; 4: 555-573.
139. Hirano, M. 1981. Clinical Examination of Voice. New York: Springer Verlag.
140. Sáenz-Lechón N, Godino-Llorente JL, Osma-Ruiz V, Blanco-Velasco M, Cruz-Roldán F. Automatic Assessment of Voice Quality According to the GRBAS Scale, http://s3.amazonaws.com/academia.edu.documents/43172414/Automatic_assessment_of_voice_quality_ac20160228-29386spdjgc.pdf?AWSAccessKeyId=AKIAJ56TQJRTWSMTNPEA&Expires=1467913279&Signature=05gaKTfdlND9ykONkCoOXFLDuT8%3D&response-contentdisposition=inline%3B%20

filename%3DAutomatic_Assessment_of_Voice_Quality _Ac.pdf, accessed Jul 7th 2016.
141. Wuyts FL, De Bodt MS, Van de Heyning PH. Is the Reliability of a Visual Analog Scale Higher Than An Ordinal Scale? An Experiment with the GRBAS Scale for the Perceptual Evaluation of Dysphonia. J Voice 1999; 13; 4: 508-517.
142. Morsomme D, Jamart J, Wéry C, Giovanni A, Remacle M. Comparison between the GIRBAS Scale and the Acoustic and Aerodynamic Measures Provided by EVA for the Assessment of Dysphonia following Unilateral Vocal Fold Paralysis. Folia Phoniatr Logop 2001; 53: 317–325.
143. Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V) ASHA Special Interest Division 3, Voice and Voice Disorders, http://www.asha.org/uploadedFiles/ASHA/SIG/03/CAPE-V-Procedures.pdf, accessed Jul 7th 2016.
144. Nemr K, Simoes-Zenari M, Ferro Cordeiro G, Tsuji D, Ogawa AI, Ubrig MT, Moreira Menezes MH. GRBAS and Cape-V Scales: High Reliability and Consensus When Applied at Different Times. J Voice 2012; 26; 6: 812.e17-812.e22.
145. Mackenzie Beck J. Perceptual Analysis of Voice Quality: The Place of Vocal Profile Analysis. In Hardcastle WJ, Mackenzie Beck J (eds.). 2005. A Figure of Speech. A Festschrift for John Laver. New York: Routledge. 285-322.
146. Webb AL, Carding PN, Deary IJ, MacKenzie K, Steen N, Wilson JA. The reliability of three perceptual evaluation scales for dysphonia. Eur Arch Otorhinolaryngol 2004; 261: 429–434.
147. Kreiman J, Gerratt BR, Kempster GB, Erman A, Berke GS. Perceptual evaluation of voice quality: review, tutorial, and a framework for future research. J Speech and Hear Res 1993; 36: 21-40.
148. Chan MKK. 2000. The Effect of Anchors and Training on the Reliability of Perceptual Voice Evaluation. Academic dissertation. Speech and Hearing Sciences. The University of Hong Kong, https://www.researchgate.net/profile/Edwin_Yiu/publication/8897884_The_Effect_of_Anchors_and_Training_on_the_Reliability_of_Perceptual_Voice_Evaluation/links/00b49519cd63160ae3000000.pdf, accessed Jul 15th 2016.
149. Boersma P, Weenink D. Praat: Doing phonetics by computer. Version 5.3.39, http://www.fon.hum.uva.nl/praat/, accessed 16th Aug 2013.
150. Airas M. TKK Aparat: An Environment for Voice Inverse Filtering and Parameterization. Log Phon Voc 2008; 33; 1:49–64.
151. Titze IR. Phonation threshold pressure: A missing link in glottal aerodynamics. J Acoust Soc Am 1992; 91: 2926-2935.
152. Chan RW, Titze IR. Viscosities of implantable biomaterials in vocal fold augmentation surgery. Laryngoscope 1998; 108: 725-731.
153. Chan RW, Titze IR. Viscoelastic shear properties of human vocal fold mucosa: Measurement methology and empirical results. J Acoust Soc Am 1999; 106: 2008-2021.

154. Pulakka H. Analysis of Human Voice Production Using Inverse Filtering, High-Speed Imaging, and Electroglottography. Master's Thesis. Helsinki, Finland: Helsinki University of Technology. 37-38.
155. Lewis JA. 2002. Social influences on female speakers' pitch. Dissertation in Linguistics. Berkeley: University of Berkeley: California.
156. Boka Z. Investigating fundamental frequency shifts in Japanese female bilingual speakers. Issues in Intercultural Communication 2010; 3; 2: 185-192, http://s3.amazonaws.com/academia.edu.documents/36858674/Boka_Japanese_paper.pdf?AWSAccessKeyId=AKIAJ56TQJRTWSMTNPEA&Expires=1468842112&Signature=34rL98FDhHRlyuNKgW6em%2BO84UE%3D&response-content-disposition=inline%3B%20filename%3D.Investigating_Fundamental_Frequency_Shi.pdf, accessed Jul 18th 2016.
157. Johns-Lewis C. 1986. Prosodic differentiation of discourse modes. In: Johns-Lewis C. (Ed.) Intonation in discourse. London: Croom Helm. 199-219.
158. Salo-Lee L. Kulttuurienvälisten viestintätaitojen oppiminen: "Prosessipuhuminen" keinona tehokkaaseen vieraskieliseen suulliseen ilmaisuun. (Learning the communication skills in intercultural communication: "Process speaking as a means for effective foreign language oral expression. In Salo-Lee L. (Ed.). 1995. Kieli & Kulttuuri oppimisessa ja opettamisessa. (Language & Culture in learning and teaching.) Jyväskylän yliopisto viestintätieteiden laitoksen julkaisuja 12. Jyväskylä: Kopi-Jyvä Oy. 153-169. (In Finnish.)
159. Kuosmanen A, de Silva V. Why don´t Russians answer my questions? Finnish students´ problems in producing Russian interrogative intonation. 15th ICPhS Barcelona 2003: 523-526, https://www.internationalphoneticassociation.org/icphsproceedings/ICPhS2003/papers/p15_ 0523.pdf, accessed Jul 18th 2016.
160. Kitzing P. LTAS criteria pertinent to the measurement of voice quality. J Phonet 1986; 14:477–482.
161. Solomon NP. Vocal fatigue and its relation to vocal hyperfunction. Int J Speech Lang Pathol 2008; 10: 254–266.
162. Löfqvist A, Mandersson B. Long-time average spectrum of speech and voice analysis. Folia Phoniatr Logop 1987; 39: 221–229.
163. Spilková H, van Dommelen WA. English 'of' in L1 and L2 speakers' read and spontaneous speech, https://www.researchgate.net/publication/266047717_English_of_in_L1_and_L2_speakers'_read_and_spontaneous_speech, accessed April 28th 2016.
164. Johannes B, Wittels P, Enne R, Eisinger G, Castro CA, Thomas JL, Adler AB, Gerzer R. Non-linear function model of voice pitch dependency on physical and mental load. Eur J Appl Physiol 2007; 101: 267–276.
165. Rantala L. 2000. Ääni työssä. Naisopettajien äänenkäyttö ja äänen kuormittuminen. (Voice at work. Female teachers´ use and loading of voice.) Academic dissertation. Oulu: Finland. Department of Finnish, Saami and Logopedics and

Otorhinolaryngology/Phoniatrics. University of Oulu. (In Finnish, abstract in English.)
166. Sajavaara K, Dufva H. Finnish-English Phonetics and Phonology. IJES 2001; 1: 241-256.
167. Yamaguchi H, Shrivastav R, Andrews ML, Niimi S. A Comparison of Voice Quality Ratings Made by Japanese and American Listeners Using the GRBAS Scale. Folia Phoniatr Logop. 2003; 55: 147–157.
168. Ghio A, Weisz F, Baracca G, Canatrella G, Robert D, Woisard V, Fussi F, Giovanni A. Is the perception of voice quality language-dependant? A comparison of French and Italian listeners and dysphonic speakers. INTERSPEECH. 2011, http://despho-apady.univ-avignon.fr/documents/GHIO-2011-Interspeech.pdf, accessed April 28th 2016.
169. Nilsson T, Master S, Järvinen K, Syrjä T, Laukkanen AM. Comparison of voice quality evaluations conducted by Brazilian and Finnish listeners. J of Voice, in review.
170. Nilsson T, Master S, Järvinen K, Syrjä T, Laukkanen AM. Comparison of voice quality evaluations conducted by Brazilian and Finnish listeners. Pan European Voice Conference Abstract Book PEVOC 11. 2015 http://www.fupress.com/archivio/pdf/2999_8023.pdf, accessed April 28th 2016.
171. Yiu E ML, Murdoch B, Hird K, Lau P and Ho EM. Cultural and Language Differences in Voice Quality Perception: A Preliminary Investigation Using Synthesized Signals. Folia Phoniatr Logop 2008; 60: 107–119.
172. Laver J. Phonetic Evaluation of Voice Quality. In Kent RD, Ball MJ. (eds.). 2000. Voice Quality Measurement. San Diego. Singular Publishing Group.37-48.
173. Grosjean F. 2010. Bilingual: life and reality. Cambridge: Harvard University Press.

APPENDICES

APPENDIX 1 Research request

APPENDIX 2 Questionnaire for subjects

APPENDIX 3 Text reading texts

APPENDIX 4 Comic strip for spontaneous speech test

APPENDIX 5 Listening test for proficiency, fluency, and accuracy

APPENDIX 6 Listening test for perceptual evaluation of voice quality, strenuousity of production, firmness of phonation, and pitch suitability

APPENDIX 1

Research request

I´m a doctoral student in vocology, Voice Research Laboratory in School of Education at Tampere University making my dissertation on Foreign language and its effect on voice. My research needs test subjects who have English as native language and some skills in Finnish. The test includes 4 1-minute speech samples, two in English and two in Finnish. The samples will be recorded in the Voice laboratory at Tampere University. Additionally, a questionnaire will be filled. The test takes about 20 minutes in total.

Additional information and signing up in the recordings, please contact: kati.jarvinen(@)uta.fi

With best regards,

Kati Järvinen

APPENDIX 2

Questionnaire for test subjects

Questionnaire

Choose from the alternatives by clicking the circle or by writing in the box. Answer carefully all the questions.

After answering all the questions you can send the questionnaire to the researcher by clicking the button "Tallenna".

Background information

Name

Sex

- male
- female

Age

Education in Finnish language, in years

Residence in a country speaking Finnish

- 0-1
- 1-5
- 5-10
- 10-20
- over 20 years

Estimated level of experience in speaking Finnish

- some experience
- experienced
- very experienced
- native

Own perspective on speaking Finnish and its effects

Speaking Finnish changes the voice

- not at all
- slightly
- moderately
- many changes
- profound changes

Pitch changes when speaking Finnish compared to speaking English

- much lower
- somewhat lower
- no changes
- somewhat higher
- much higher

Pitch variation changes when speaking Finnish compared to speaking English

- much less
- slightly less
- no changes
- some increase
- major increase

Intensity changes when speaking Finnish compared to speaking English

- much lower
- somewhat lower
- no changes
- somewhat higher
- much higher

Variation of intensity changes when speaking Finnish compared to speaking English

- decreases a lot
- decreases somewhat
- no changes
- increases somewhat
- increases a lot

Voice tires faster speaking Finnish than speaking English

- not at all
- a little faster
- faster
- quite fast
- much faster

In your opinion, what are the reasons for the changes?

Finnish differs from English

In articulation

- not at all
- a little
- moderately
- a lot
- very much

How?

In accentuation

- not at all
- a little
- moderately
- a lot
- very much

How?

Intonation range

- much narrower
- a little narrower
- no difference
- a little wider
- much wider

In speech tempo

- much slower
- a little slower
- no difference
- a little faster
- much faster

Something else? What? How?

Vocal resources in English

Voice quality

- very poor
- poor
- not good or poor
- good
- very good

Voice resonance

- very poor
- poor
- not good or poor
- good
- very good

Vocal endurance

- very poor
- poor
- not good or poor
- good
- very good

Vocal expressiveness

- very poor
- poor
- not good or poor
- good
- very good

Vocal resources in Finnish

Voice quality

- very poor
- poor
- not good or poor
- good
- very good

Voice resonance

- very poor
- poor
- not good or poor
- good

- very good

Vocal endurance

- very poor
- poor
- not good or poor
- good
- very good

Vocal expressiveness

- very poor
- poor
- not good or poor
- good
- very good

Background information on voice

Voice training in speech

- none
- some
- moderately
- a lot
- very much indeed

If yes, what and how long

Voice therapy

- none
- some
- moderately
- a lot
- very much indeed

If yes, what and how long

Smoking

- never smoked
- have smoked but stopped
- smoking regularly

If you smoke, how many cigarettes per day

Tietojen lähetys

Thank you for answering this questionnaire!

APPENDIX 3

The English and the Finnish texts

The English text:

The island was moving all night. The fisherman´s point drifted imperceptibly a little farther out to sea. Shudder after shudder shook the whole island like chills running up and down its spine, and the black pool seemed to creep deeper and deeper into the rocks. It was sucked in and out and fresh waves broke in from the sea, but the pool never seemed to fill up. Its enormous mirror-like black eye sank lower and lower, surrounded by a fringe of sea-grass round the edges. On the beach on the leeward side, little field-mice ran backwards and forwards at the edge of the water, the sand slipping away from under their paws. Boulders turned over heavily, revealing the pale roots of the sea-pinks. At dawn the island slept. The trees had reached the lighthouse-rock; deep holes were left where great boulders had been before, now lying scattered among the heather. They were waiting for another night to come so that they could roll nearer and nearer the lighthouse. The great autumn gale continued to blow. At seven o´clock father went out to look at the boat. The water had risen again and the southwest wind was blowing the sea higher and higher. He found the fisherman lying rolled up at the bottom of the boat. The fisherman twisted his bent legs over the side of the boat and tumbled on to the beach. His eyes were just as kind and gentle as ever, and he said:"I haven´t done any harm..."

Jansson, T. 1965. Moominpappa at Sea, translated by Kingsley Hart (1993). USA: A Sunburst Book Farrar, Straus and Giroux. 199-200.

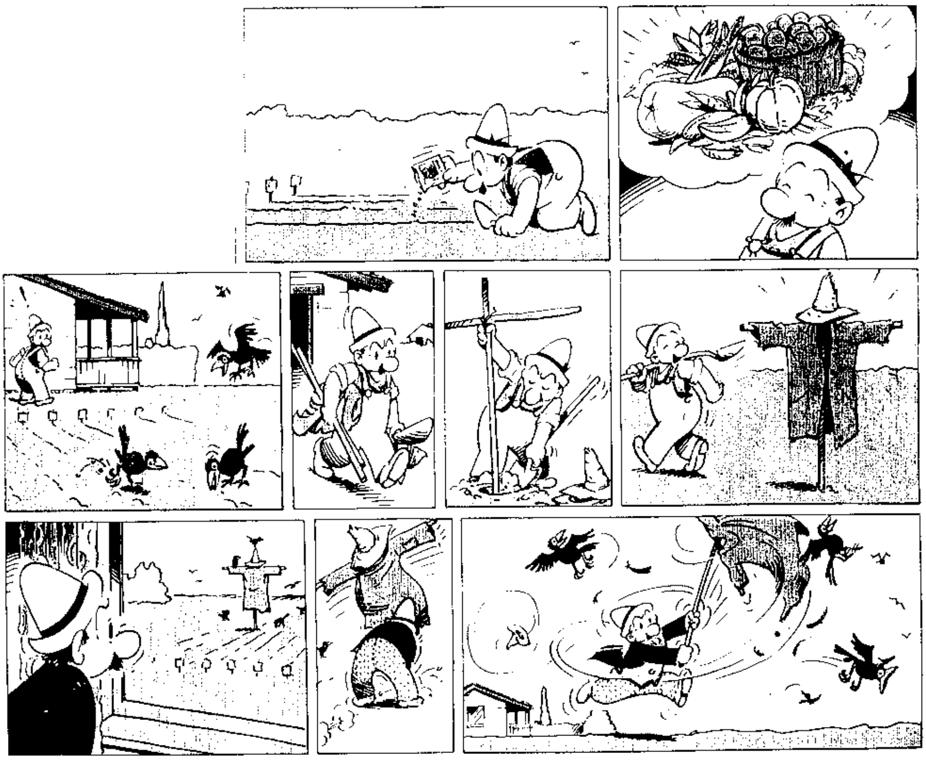
The Finnish text:

Saari vaelsi koko yön. Kalastajan niemi kulki huomaamatta kauemmas mereen. Ankarat väreet puistattivat kallioselänteitä aivan kuin tumma viri käy pitkin vedenpintaa, ja musta lampi ryömi syvemmälle alkukallioon. Se vetäytyi koristen alaspäin ja sisäänpäin, ja aallot syöksyivät merestä yli kannaksen kiiltävänä, vihreänä vesiputouksena. Mutta lampi ei täyttynyt. Se vetäytyi pakoon, nyt sen tumma peilisilmä hohti kaukaa saaren uumenista meriheinäripsien reunustamana. Tyynenpuoleisessa rannassa juoksentelivat myyrät ja metsähiiret edestakaisin vedenrajassa. Hiekka valui pois niiden tassujen alta. Kivet kääntyivät raskaasti niin että rantakauran valkoiset juuret jäivät paljaiksi. Aamun sarastaessa saari nukahti. Silloin puut olivat ehtineet majakkakallion eteen, kivipellon tilalla oli syvä kuoppa, kokonainen armeija pyöreitä, harmaita mukulakiviä oli hajallaan kanervikossa. Ne odottivat seuraavaa yötä vieriäkseen edelleen majakkaa kohti. Suuri syysmyrsky riehui lakkaamatta. Seitsemän tienoilla isä meni katsomaan venettä. Vesi oli taas kohonnut ja pitkä lounaistuuli nosti aallokon yhä korkeammaksi. Ja silloin isä löysi kalastajan veneen pohjalta. Kalastaja nosti ryppyiset säärensä veneen laidan yli ja muksahti maahan. Hänen silmänsä olivat lempeät ja rauhalliset hänen mumistessaan: ”En ole koskaan tehnyt mitään pahaa...”

Jansson, T. 1965 (7th ed, 1991). Muumipappa ja meri, translated by Laila Järvinen. Juva: WSOY. 161-162.

APPENDIX 4

The comic strip for spontaneous speech samples



APPENDIX 5

Listening test for proficiency, fluency, and accuracy

Listen to the samples and mark after every sample with a numeral evaluation of the speaker´s proficiency, fluency and accuracy in the English language.

subject	proficiency text reading	fluency text reading	accuracy text reading
	0=no skills 1=some skills 2=skilled 3=very skilled 4=native like	0=no fluency 1=some fluency 2=fluent 3=very fluent 4=native like	0=no accuracy 1=some accuracy 2=accurate 3=very accurate 4=native like
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			
13			
14			
15			
16			
17			
18			
19			
20			

subject	proficiency spontaneous	fluency spontaneous	accuracy spontaneous
	0=no skills 1=some skills 2=skilled 3=very skilled 4=native like	0=no fluency 1=some fluency 2=fluent 3=very fluent 4=native like	0=no accuracy 1=some accuracy 2=accurate 3=very accurate 4=native like
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			
13			
14			
15			
16			
17			
18			
19			
20			

APPENDIX 6

Listening test for perceptual evaluation of voice quality, strenuousity of production, firmness of phonation, and pitch suitability

Kuuntele näytteet pareina, esim. näytteet 1a ja 1b ovat pari. Klikkaa näytettä parhaiten kuvaavaa ympyrää. Jokaisessa ominaisuudessa on asteikko 0-10. (Listen to the samples as pairs, for example, samples 1a and 1b are a pair. Click the button that best represents the samples. Every characteristic has a scale from 0 to 10.)

Äänenlaatu 0=erittäin huono, 10=erittäin hyvä
(Voice quality 0=very poor, 10=very good)

Äänen tiiviys 0=erittäin puristeinen, 10=erittäin vuotoinen
(Firmness 0=very pressed, 10=very breathy)

Äänentuotto 0=työläs, 10=helppo
(Voice production 0=hard, 10=easy)

Korkeus 0=liian matala, 10=liian korkea
(Pitch 0=too low, 10=too high)

Jokaisen näyteparin perässä on myös mahdollisuus kommentoida vapaasti näytteitä, max. 250 merkkiä. (Every sample pair has a possibility to comment freely the samples, max. 250 signs.)

1a

ÄÄNENLAATU (VOICE QUALITY)

erittäin huono	0	1	2	3	4	5	6	7	8	9	10	erittäin hyvä
(very poor)	○	○	○	○	○	○	○	○	○	○	○	(very good)

ÄÄNEN TIIVIYS (FIRMNESS OF PHONATION)

erittäin puristeinen	0	1	2	3	4	5	6	7	8	9	10	erittäin vuotoinen
(very pressed)	○	○	○	○	○	○	○	○	○	○	○	(very breathy)

ÄÄNENTUOTTO (VOICE PRODUCTION)

työläs	0	1	2	3	4	5	6	7	8	9	10	helppo
(hard)	○	○	○	○	○	○	○	○	○	○	○	(easy)

KORKEUS (PITCH)

liian matala	0	1	2	3	4	5	6	7	8	9	10	liian korkea
(too low)	○	○	○	○	○	○	○	○	○	○	○	(too high)

1b

ÄÄNENLAATU (VOICE QUALITY)

erittäin huono	0	1	2	3	4	5	6	7	8	9	10	erittäin hyvä
(very poor)	○	○	○	○	○	○	○	○	○	○	○	(very good)

ÄÄNEN TIIVIYS (FIRMNESS OF PHONATION)

erittäin puristeinen	0	1	2	3	4	5	6	7	8	9	10	erittäin vuotoinen
(very pressed)	○	○	○	○	○	○	○	○	○	○	○	(very breathy)

ÄÄNENTUOTTO (VOICE PRODUCTION)

työläs	0	1	2	3	4	5	6	7	8	9	10	helppo
(hard)	○	○	○	○	○	○	○	○	○	○	○	(easy)

KORKEUS (PITCH)

liian matala	0	1	2	3	4	5	6	7	8	9	10	liian korkea
(too low)	○	○	○	○	○	○	○	○	○	○	○	(too high)

Muita huomioita

(Other remarks)

Kiitos vastauksistasi.

(Thank you for your answers.)

ORIGINAL ARTICLES

Logopedics Phoniatrics Vocology, 2013; 38: 47–51

informa
healthcare

ORIGINAL ARTICLE

Speaking a foreign language and its effect on F0

KATI JÄRVINEN[1], ANNE-MARIA LAUKKANEN[1] & OLLI AALTONEN[2]

[1]*University of Tampere, Voice Research laboratory, Tampere, Finland, and* [2]*Behavioural Sciences, Speech sciences, University of Helsinki, Helsinki, Finland*

Abstract
This study investigated whether speaking a foreign language affects the fundamental frequency (F0) of speech in 16 native Finnish and 14 native English subjects reading a text in Finnish and in English. The speech samples were analyzed for the mean and range of F0. Speaking a foreign language caused a change in F0 for the Finnish subjects, while the result was not as unambiguous for the English subjects. The change in F0 may be a result of adaptation to a certain pitch level in the foreign environment. Experience in using the foreign language did not show significant correlation to the change in F0, which suggests either individual differences in sensitivity to adaptation or difficulty in quantifying the amount of experience.

Key words: *Adaptation in speech, English, Finnish, fundamental frequency*

Introduction

Fundamental frequency (F0) is a widely studied parameter in speech, which includes F0 and its relation to speaker characteristics, the type of speech used, and text materials (1). F0 is determined especially at an individual level, depending on the physical characteristics of the voice production apparatus, the vibration of the vocal folds (2), and the habitual use of voice (3), but many linguistic features, such as intonation, may affect the mean fundamental frequency of speech (4–6).

Cultural factors can create ideals of the mean F0 of speech that are manifested in the speaker's voice. According to previous studies, differences between languages and cultures can be found (7–11). The differences in pitch can be influenced by the functions of expression in cultures (12–15), e.g. a relatively low pitch is considered favorable in Finland while a higher pitch is traditionally associated with a higher social status in Britain (16). Differences in perception of a speaker may thus be both language- and culture-dependent (17,18).

Not only linguistic but also non-linguistic factors can affect the F0. Such factors have in previous studies been mentioned as physiological and psychological factors, such as the condition of the vocal cords (19), physiological stress (20,21), and emotional factors (22–26). Psychological factors, such as stress, may be connected to the speaker's personality traits (27).

Speaking a foreign language may have an effect on voice and on F0. Previous studies suggest that when speaking a foreign language F0 can increase (12,13,28,29) which can be due to the speakers' attempt to adjust their voice. This is predominant in the cultural environment of the target language or a result of the fact that speaking a foreign language may be a task more demanding than speaking the native language, and may thus increase the level of stress and therefore cause the fundamental frequency to rise. Even with bilinguals differences between languages have been found (8), which may indicate that adaptation influences the pitch rather than mere stress. Other vocal characteristics, such as voice quality, may also be affected by the language change (30,31).

The majority of previous studies in language change and F0 have focused on women. This current study aimed to resolve the change in F0 with men and women when speaking a foreign language compared to speaking the native language, with the

Correspondence: Käti Jarvinen, University of Tampere, Voice Research laboratory, Kalevantie 4, Tampere, 33014 Finland. E-mail: kati.jarvinen@uta.fi

(Received 21 December 2011; accepted 27 March 2012)

ISSN 1401-5439 print/ISSN 1651-2022 online © 2013 Informa UK, Ltd.
DOI: 10.3109/14015439.2012.687764

focus on mean F0 and the F0 range, and whether the subjects are aware such changes occur. Also the experience in speaking the target language was taken into account, for e.g. residence in a country speaking the language may have a vast significance in the speech production (32).

Methods

Subjects

Sixteen native speakers of Finnish (7 males and 9 females, mean ages 36 and 32 years, standard deviations 13.7 and 8.8, respectively) and 14 native speakers of English (8 males and 6 females, mean ages 42 and 28, standard deviations 20.6 and 8.4, respectively) read aloud a 1-minute text passage in Finnish and in English; both groups read first the text in their native language. The texts were the same in content. Samples with duration of 1 minute have been considered to be sufficient to establish fundamental frequency and be an indication of the person's habitual pitch in a neutral speech situation (33).

Recordings

The recordings took place in a well-damped studio with Bruel & Kjaer Mediator; the microphone was placed in front of the mouth, and the distance from the mouth was 40 cm. The signal was recorded with Sound Forge 7.0 software, frequency rate 44.1 kHz, amplitude range 16 bits.

Acoustical analysis

The samples were acoustically analyzed for the mean, standard deviation, and range of F0 with Praat 5.1.15 signal analysis system. Individual pitch ranges and analysis method cross-correlation were used (www.fon.hum.uva.nl/praat/).

Statistical analysis

PASW Statistics SPSS 18 software was used for the statistical analysis. Related-samples Wilcoxon signed ranks test was used for the change in F0 in groups, Mann–Whitney *U* test for the difference in the change in F0 between groups, and Spearman's correlation coefficient for the correlation between experience in speaking the foreign language and the change in F0, and subjective notion on the change in F0 and the actual change in F0. The significance level was 5%.

Questionnaire

The subjective notions of the possible differences between speaking the two languages as well as the education and experience in the foreign language and residence in a country speaking the foreign language were asked by a questionnaire Appendix 1 Overall experience was calculated from education, residence, and own estimated level of experience by arbitrarily converting the answer into one measure, the minimum score being then 2 and maximum 12 in overall experience.

Results

When the change in F0 between the native and the foreign language was calculated in percentages, a significant difference between groups was found, $P = 0.008$.

As shown in Table I, mean F0 was significantly higher in the target language for the Finnish subjects, $P = 0.001$, but for the English subjects the change was not statistically significant.

F0 changed for the Finnish subjects significantly in English compared to Finnish, males $P = 0.028$ and females $P = 0.011$. For the English males or females the change was not statistically significant. Change of language did not show significant changes in F0 standard deviation or F0 range in either groups or between groups (Table II).

Twelve (75%) of the Finnish subjects thought that they used a higher pitch when speaking English than in Finnish speech; only two persons considered that they had a lower pitch in English than in Finnish. Two of the Finnish subjects had not noticed a difference in pitch between Finnish and English. Eight (57%) of the English subjects considered themselves to have a lower pitch in Finnish than in

Table I. Means and standard deviations for change in F0 (%), change in standard deviation of F0 (%), and change in F0 range (%) for Finnish and English subjects.

	F0 change % $\bar{X}$ (SD)	*P*	F0 standard deviation change (%) $\bar{X}$ (SD)	*P*	F0 range change % $\bar{X}$ (SD)	*P*
Finns ($n = 16$)	4.00 (3.97)	0.001	1.94 (9.99)	ns	0.82 (11.20)	ns
English ($n = 14$)	0.84 (4.64)	ns	–3.74 (12.52)	ns	–0.26 (7.48)	ns

ns = not significant.

Table II. Means and standard deviations for F0, change in F0 (%), standard deviation of F0, change in standard deviation of F0 (%), F0 range, and change in F0 range in native and foreign language.

	F0 (Hz) in native $\bar{X}$ (SD)	F0 (Hz) in foreign $\bar{X}$ (SD)	*P*	F0 (Hz) standard deviation in native $\bar{X}$ (SD)	F0 (Hz) standard deviation in foreign $\bar{X}$ (SD)	*P*	F0 range (Hz) in native $\bar{X}$ (SD)	F0 range (Hz) in foreign $\bar{X}$ (SD)	*P*
Finnish males ($n = 7$)	102.97 (20.92)	106.35 (22.61)	0.028	14.54 (4.03)	14.98 (4.66)	ns	99.83 (20.64)	98.51 (21.11)	ns
Finnish females ($n = 9$)	180.94 (16.26)	188.89 (14.14)	0.011	24.04 (4.54)	24.18 (4.63)	ns	171.83 (35.14)	173.36 (32.43)	ns
English males ($n = 8$)	105.26 (14.24)	105.87 (13.53)	ns	16.64 (6.06)	16.12 (5.81)	ns	117.25 (30.31)	115.84 (32.13)	ns
English females ($n = 6$)	201.56 (12.61)	199.79 (13.76)	ns	27.35 (7.27)	25.65 (7.03)	ns	192.83 (40.66)	195.19 (46.51)	ns

ns = not significant.

English, two to have a higher pitch in Finnish than in English, and four had not noticed a change in pitch between languages (Figure 1).

The subjective notion of how the pitch changes when speaking the target language did not show significant correlation with the actual change in F0.

Experience in the target language consisted of education, residence, and own estimation of the level of experience in speaking the language. The Finnish subjects had a longer formal education in English than the English subjects had in Finnish (mean 14 and 0.7 years, respectively). The residence in a country speaking the target language was much longer for the English subjects (57% more than 5 years) than the Finns (100% 0–5 years). The Finnish subjects considered themselves to be more experienced in speaking English than the English subjects in speaking Finnish (Table III).

Experience in speaking the target language did not show significant correlation with the change in F0.

Discussion and conclusions

The mean F0 of speech changed when speaking a foreign language compared to speaking the native language, and the change seemed to be somewhat

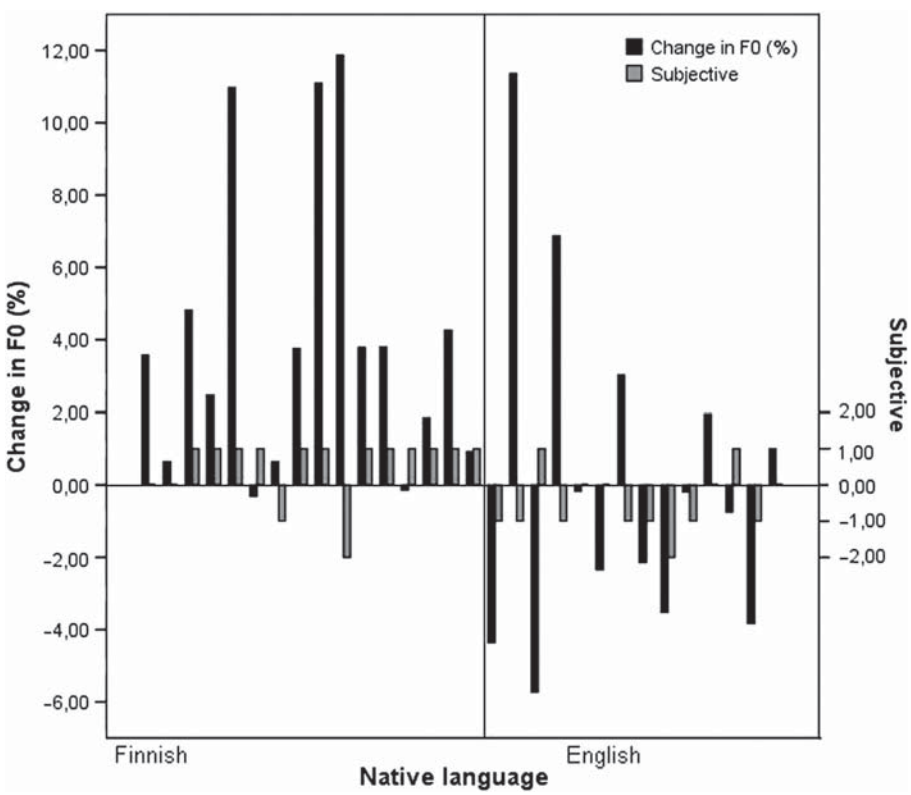


Figure 1. Distributions of change in F0 (%) and subjective notions of how pitch changes when speaking the target language compared to speaking the native language (–2 = much lower, –1 = somewhat lower, 0 = no changes, 1 = somewhat higher, 2 = much higher).

Table III. Means in experience in speaking the target language.

	Education (years)	Residence	Own estimation	Total experience
Finns	14.0	1.25	1.63	6.5
English	0.7	2.93	0.57	4.3

Total experience is calculated from education (1 = 0–1 year, 2 = 1–5 years, 3 = 5–10 years, 4 = over 10 years), residence (1 = 0–1 year, 2 = 1–5 years, 3 = 5–10 years, 4 = 10–20 years, 5 = over 20 years), and own estimation of level of experience (0 = some experience, 1 = experienced, 2 = very experienced, 3 = native-like); total experience: minimum score 2 and maximum score 12.

related to the target language. The use of pitch and its range can be gender- and culture-related (34), which can guide the non-native speakers to a certain pitch level. On the other hand the change in F0 was significantly different between the two groups. This might also be evidence of a certain kind of adaptation, since the English-speaking subjects' change in F0 was towards a lower F0 in Finnish than in English. Speech Accommodation Theory argues that people often modify their speech characteristics in interaction. Adaptation can occur as linguistic and non-verbal behavior between communicators. Both children and adults tend to adapt features, such as pitch, in their speech to match the interlocutor's speech (35–37). The rise of F0 in speaking the target language can also be a result of uncertainty, for lack of confidence can be signaled by a higher F0 (38), and it is possible that speaking a language other than the native one is a more demanding task and therefore can cause uncertainty in the speaker. The range of F0 did not change between languages, which can be caused by the fact that the range of F0 is somewhat constant in speech, and therefore a difficult characteristic to change consciously or unconsciously. In this study no correlation with experience and the changes in F0 was found, which may indicate that people have individual differences in sensitivity to adaptation, or it may reflect difficulty in quantifying experience in the use of the target language. The quantifying overall experience needs to be studied further. The subjects varied considerably in their education in the foreign language as well as in their residence, so it might be reasonable to view these factors separately. Also the task in this study (text-reading) may have influenced the F0 and the range of F0, because the task can affect the fundamental frequency (15,39). A further study with a larger number of subjects is warranted. It might be appropriate in the future to include speech samples based on something other than reading a text and to investigate the effects of age and fluency of the speakers.

Declaration of interest: This study was supported by the University of Tampere. The authors report no conflicts of interest.

References

1. Hollien H, Hollien PA, de Jong G. Effects of three parameters on speaking fundamental frequency. J Acoust Soc Am. 1997;102:2984–92.
2. Baken RJ, Orlikoff RF. Clinical measurement of speech and voice. London: Taylor & Francis Ltd; 2000.
3. Coleman RF, Markham IW. Normal variations in habitual pitch. J Voice. 1991;5:173–7.
4. Flege J. Factors affecting degree of perceived foreign accent in English sentences. J Acoust Soc Am. 1988;84:70–9.
5. Flege J, Hillenbrand J. Limits on phonetic accuracy in foreign language speech production. J Acoust Soc Am. 1984;76: 708–21.
6. Grover C, Jamieson D, Doborovolsky M. Intonation in English, French and German: perception and production. Lang Speech. 1987;30:277–95.
7. Leino T, Laukkanen AM, Ilomäki I, Mäki E. Assessment of vocal capacity of Finnish students. Folia Phoniatr Logop. 2008;60:199–209.
8. Altenberg EP, Ferrand CT. Fundamental frequency in monolingual English, bilingual English/Russian, and bilingual English/Cantonese young adult women. J Voice. 2006;20: 89–96.
9. Lindh J. Preliminary descriptive F0-statistics for young male speakers. Lund University. Working Papers. 2006;52:89–92.
10. Rantala L. Ääni työssä. Naisopettajien äänenkäyttö ja äänen kuormittuminen [Voice at work. Female teachers' voice use and loading] [Academic dissertation]. Acta Universitatis Ouluensis, B 37. University of Oulu; 2000. (In Finnish, abstract in English)
11. Pegoraro Krook MI. Speaking fundamental frequency characteristics of normal Swedish subjects obtained by glottal frequency analysis. Folia Phoniatrica. 1988;40:82–90.
12. Ohara Y. Performing gender through voice pitch: A cross-cultural analysis of Japanese and American English. In: Pasero U, Braun F, editors. Perceiving and performing gender. Opladen/Wiesbanden: Westdeutscher Verlag; 1999. p. 105–16.
13. Ohara Y. Gender-dependent pitch levels: a comparative study in Japanese and English. In: Hall K, Bucholtz M, Moonwomon B, editors. Locating power. Proceedings of the second Berkeley Women and Language Conference. Berkeley, CA, 4–5 May 1992. Berkeley, CA: Berkeley Women and Language Group; 1992. p. 468–77.
14. Awan SN, Mueller PB. Speaking fundamental frequency characteristics of white, African American, and Hispanic kindergartners. J Speech Hear Res. 1996;39:573–77.
15. Hudson AI, Holbrook A. Fundamental frequency characteristics of young black adults. Spontaneous speaking and oral reading. J Speech Hear Res. 1982;25:25–8.
16. Valo M. Käsitykset ja vaikutelmat äänestä. Kuuntelijoiden arviointia radiopuheen äänellisistä ominaisuuksista [Perceptions and appearances of voice. Listeners' evaluation on vocal characteristics in radio speech] [Academic dissertation]. Studia Philologica Jyväskyläensia 33. University of Jyväskylä; 1994. (In Finnish, abstract in English.)
17. Mennen I, Schaeffer F, Docherty G. Pitching it differently: a comparison of the pitch ranges of German and English speakers. 16th ICPhS, Saarbrücken, conference: 6–10 Aug 2007, p. 1769–72. www.icphs2007.de.

18. van Bezooijen R. Sociocultural aspects of pitch differences between Japanese and Dutch women. Lang Speech. 1995;38:253–65.
19. Guimarães I, Abberton E. Health and voice quality in smokers: an exploratory investigation. Logoped Phoniatr Vocol 2005;30:185–91.
20. Orlikoff RB, Baken RJ. Fundamental frequency modulation of the human voice by the heartbeat: preliminary results and possible mechanism. J Acoust Soc Am. 1988;85: 888–93.
21. Orlikoff RB, Baken RJ. The effect of the heartbeat on vocal fundamental frequency in speech. J Speech Hear Res. 1989;32:576–82.
22. Ruiz R, Legros C, Guell A. Voice analysis to predict the psychological or physical state of a speaker. Aviat Space Environ Med. 1990;61:266–71.
23. Vilkman E, Manninen O. Changes in prosodic features of speech due to environmental factors. Speech Commun. 1986; 5:331–45.
24. Shipp T, Izdebski K. Current evidence for the existence of laryngeal macrotremor and microtremor. J Forensic Sci. 1981;26:501–5.
25. Williams CE, Stevens KN. Vocal correlates of emotional states. In: Darby JK, editor. Speech evaluation in psychiatry. New York: Grune and Stratton; 1981. p. 221–40.
26. Hollien H. Vocal indicators of psychological stress. Ann N Y Acad Sci. 1980;347:47–72.
27. Johannes B, Salnitski VP, Gunga HC, Kirsch K. Voice stress monitoring in space—possibilities and limits. Aviat Space Environ Med. 2000;71 suppl 9:A58–65.
28. Järvinen K, Laukkanen AM, Izdebski K. Voice fundamental frequency changes as a function of foreign languages familiarity: an emotional effect? In: Izdebski K, editor. Emotions in the human voice. San Diego, CA, USA: Plural Publishing; 2008. p. 203–13.
29. Syrjä T. Vieras kieli suussa. Vieraalla kielellä näyttelemisen ulottuvuuksia näyttelijäopiskelijan äänessä, puheessa ja kehossa [A strange tongue in the mouth. The dimensions of acting in a foreign language in the students' voice, speech and body] [Academic dissertation]. Tampere: Tampere University; 2007. (In Finnish, abstract in English)
30. Acton W. Changing fossilized pronunciation. TESOL Quarterly. 1984;18:71–85.
31. Bruyninckx M, Harmegnies B, Llisterri J, Poch-Olivé, D. Language-included voice quality variability in bilinguals. J Phon. 1994;22:19–31.
32. Ullakonoja R. Da. Eto Vopros! Prosodic development of Finnish students' read-aloud Russian during study in Russia [Academic dissertation]. Jyväskylä Studies in Humanities 151. Jyväskylä: University of Jyväskylä; 2011.
33. Nolan, F. The phonetic bases of speaker recognition. Cambridge: Cambridge University Press; 1983.
34. Lewis JA. Social influences on female speakers' pitch [Dissertation in Linguistics]. Berkeley: University of Berkeley California; 2002.
35. Littlejohn SW. Theories of human communication. 6th ed. Belmont, CA, USA: Wadsworth Publishing Company; 1999.
36. Burgoon J, Stern L, Dillman L. Interpersonal adaptation: dyadic interaction patterns. Cambridge, UK: Cambridge University Press; 1995.
37. Giles H, Mulac A, Bradac J, Johnson P. Speech accommodation theory: the first decade and beyond. In: McLaughlin ML, editor. Communication yearbook, 10. London, UK: Sage; 1987. p. 13–48.
38. Ohala JJ. An ethological perspective on common cross-language utilization of F0 of voice. Phonetica. 1984;41:1–16.
39. Johns-Lewis C. Prosodic differentiation of discourse modes. In: Johns-Lewis C, editor. Intonation in discourse. London: Croom Helm; 1986. p. 199–219.

Appendix 1

Questionnaire

Choose from the alternatives by clicking the circle or by writing in the box. Answer carefully all the questions.
After answering all the questions you can send the questionnaire to the researcher by clicking the button "Tallenna".

Background information

Name

Sex

☐ male ☐ female

Age

Education in Finnish/English language, in years

Residence in a country speaking Finnish/English

○ 0-1 ○ 1-5 ○ 5-10 ○ 10-20 ○ over 20 years

Estimated level of experience in speaking Finnish/English

○ some experience ○ experienced ○ very experienced ○ native-like

Pitch changes when speaking Finnish/English compared to speaking English/Finnish

○ much lower ○ somewhat lower ○ no changes ○ somewhat higher ○ much higher

Original Paper

Folia Phoniatrica et Logopaedica

Folia Phoniatr Logop 2015;67:1–7
DOI: 10.1159/000381183

Received: October 3, 2014
Accepted: February 23, 2015
Published online: April 24, 2015

Vocal Loading in Speaking a Foreign Language

Kati Järvinen Anne-Maria Laukkanen

Speech and Voice Research Laboratory, School of Education, University of Tampere, Tampere, Finland

Key Words
Mean fundamental frequency · Equivalent sound level · Voicing amount · Vocal doses · Voice quality · Vocal fatigue · Bilingualism

Abstract
Aims: This study investigated whether speaking a foreign language affects the subjective notions of vocal fatigue, and whether acoustic measurements reveal a higher vocal loading. ***Methods:*** The speech samples of 20 native Finnish-speaking and 23 native English-speaking subjects were recorded in Finnish and in English. From the speech samples, fundamental frequency, equivalent sound level, total duration of voiced speech, speech rate, alpha ratio and L1-L0 level difference were analyzed. Vocal doses were calculated. ***Results:*** According to subjective notions, the voice gets tired more quickly when speaking a foreign language. The mean fundamental frequency increased but the speech rate and total duration of voiced speech decreased significantly when speaking a foreign language. Thus, the vocal doses decreased. ***Conclusions:*** The subjective sensations of increased vocal fatigue may be due to increased mental stress rather than to higher vocal loading. However, a trend that speaking a foreign language may involve more loading was found in L1-L0 level difference and in the doses normalized to time dose. Longer speech samples should be studied. Voice quality-based indicators of vocal loading are worth testing in addition to the measures based on the amount of voicing in speech.

Introduction

Skills in foreign languages are needed in multicultural societies and globally. Speaking a language other than the native one requires the speaker to adapt to different speech sounds and prosodic systems and to the vocal ideals of other cultures [1]. Contrastive research in language learners' production and perception of a foreign language has to a large extent focused on the accuracy of performance in the target language or the influence of the native language on the target language [for examples, see 2–5; for examples of differences between Finnish and English, see 6–8]. Voice parameters are prone to be affected by a language change between native and foreign languages. Parameters previously considered include mean fundamental frequency (F0) [9–11], F0 range [12] and sound energy distribution along frequencies [13, 14]. People who use foreign languages to a large extent, e.g. professional interpreters and simultaneous translators, have complained about symptoms of vocal fatigue and mental tiredness after prolonged use of the foreign language [15]. This leads to the question whether speaking a foreign language is vocally more loading than speaking the mother tongue.

Vocal overloading is regarded as resulting from a combination of loading factors such as excessively prolonged voice use, high pitch and intensity, and a pressed type of phonation [16, 17]. Loading factors imply a higher impact stress on the vocal fold tissue. A higher alpha ratio (the

KARGER 125

© 2015 S. Karger AG, Basel
1021–7762/15/0671–0001$39.50/0

E-Mail karger@karger.com
www.karger.com/fpl

Kati Järvinen
Speech and Voice Research Laboratory, School of Education
University of Tampere
FI–33014 Tampere (Finland)
E-Mail jarvinen.kati.h@student.uta.fi

ratio of spectral energy below and above 1,500 Hz) [18] and a lower sound level of the F0 region in relation to that of F1 [19] may indicate a more pressed phonation type [18, 20].

Previous studies have proposed different parameters for quantifying vocal loading, such as F0 and sound pressure level (SPL) [21], vocal loading index (VLI, F0 × F0 time/1,000, i.e. the number of vocal fold periods in kilocycles) [22], phonation threshold pressure [23] and characteristics of vocal fold vibration [24]. Five vocal doses based on F0, SPL and voicing detection have been introduced for quantifying the amount of vocal loading [25, 26]. The doses are: the time dose (D_t) as an indication of total voicing time, the cycle dose (D_c) indicating the total number of vocal fold oscillatory cycles (i.e. the same as VLI times 1,000), the distance dose (D_d) which indicates the total distance that vocal folds travel in an oscillatory path, the energy dissipation dose (D_e) indicating the total amount of dissipation of heat energy in the vocal folds, and the radiated energy dose (D_r) which indicates the total acoustic energy that radiates from the mouth.

In studying the effects of vocal loading, self-reported vocal loading symptoms [18] and psychological stress [27] should be considered in addition to acoustical analysis. Individual factors, such as gender, endurance, general health, life habits, vocal skills, and experience and personality, affect the voice [28] and may contribute to the effects of vocal loading.

This study investigates whether speaking a foreign language affects subjective notions of vocal fatigue and whether signs of increased vocal loading can be seen in acoustical analysis. The first foreign language that Finnish school children nowadays start studying is English; in 2009, 90% of children in school (attending the third class at the age of 8–10 years) chose English as their first foreign language [29], so it seemed logical to consider Finnish and English in the present study.

Some characteristic differences between Finnish and English include differences in phonemes (English does not have /y/, Finnish does not have voiced affricate /dʒ/), differences in articulation (e.g. Finnish /r/ is a voiced tremulant, while in English, it is often realized as an approximant /ɹ/, and vowels have different formant frequencies), and the fact that Finnish has short/long vowel opposition while English has quality opposition in vowels to distinguish meaning. Furthermore, Finnish has no aspiration of consonants as a distinctive feature. The Finnish language mainly uses interrogatives to mark questions, while English uses intonation type. In Finnish intonation, the highest peak in the F0 contour is typically located at the start of a sentence and the end of the sentence is low. Also, the intonation range is narrower in Finnish than in English [6, 30–32].

Methods

Subjects and Tasks

Forty-three subjects (mean age 37.8 years, SD 16.4, range 19–73; Finnish subjects' mean age 39.8 years, SD 15.9, range 19–69; English subjects' mean age 36.0 years, SD 16.9, range 21–73) participated in this study. Twenty of the subjects were native Finnish speakers (8 males and 12 females), and 23 were native English speakers (14 males and 9 females); the subjects were not simultaneous bilinguals but had learned the foreign language later. The English-speaking subjects came from different regions, with the majority of them being native speakers of American or British English. All subjects were recruited in Finland, and the English subjects were living in Finland at the time of the study. The subjects read aloud from a text for 1 min first in the native language (Finnish or English) and then in the foreign language (English or Finnish). The texts were the same in content in both languages. After reading the texts, the subjects recorded a sample of spontaneous speech in both languages, the task being a description of a comic strip. The duration of the spontaneous speech samples varied between 21.5 s and 1 min (mean 42.9 s) in the native language and between 26 s and 1 min (mean 43.6 s) in the foreign language. The spontaneous speech samples were gathered from 40 subjects; 3 of the English-speaking subjects were not proficient enough to produce spontaneous speech in the foreign language. For dose measurements, samples of the same duration were analyzed in both languages.

After the recordings, the subjects filled in a questionnaire collecting background information and their opinion whether speaking a foreign language affects their voice use and whether the voice gets tired faster in foreign language than in native language use (vocal fatigue). According to the answers to the questionnaire, the subjects' mean duration of formal education in the foreign language was 6 years (SD 8.4, range 0–40). The wide range in the duration of education is explained by the fact that some subjects were teachers of English at the university and, thus, most likely considered language learning as an ongoing process. Half of the subjects had resided in a country of the foreign language for ≤1 year (51%), 35% from 1 to 10 years and 14% for >10 years. Half of the subjects (50%) considered themselves as having some experience in speaking the foreign language, 43% as being experienced or very experienced and 7% as being native-like. Eighty-five percent of the Finnish subjects had >5 years of education in the foreign language and 60% of the English subjects <5 years. Seventy-five percent of the Finnish subjects had resided in an English-speaking country for ≤1 year, while 70% of the English subjects had resided in Finland for ≥5 years. Forty percent of the Finnish subjects and 9% of the English subjects considered themselves as very experienced or native-like in speaking the foreign language.

Recordings and Analyses

The recordings took place in a well-damped recording studio using a combined level meter and microphone (Bruel and Kjaer Mediator, type 2206), placed in front of the subject at a distance of 40 cm from the mouth. The signal was recorded with Sound Forge 7.0 software, using 44.1-kHz sampling frequency and 16-bit amplitude quantization. The signal was calibrated for calculation of equivalent sound level (Leq) by recording a reference tone.

From the speech samples, the mean F0, Leq, speech rate (syllables per second), total duration of voiced speech (unvoiced segments and pauses excluded) and VLI were analyzed. The alpha ratio was calculated by subtracting the Leq within the range of 1.5–5 kHz from that of the range from 50 Hz to 1.5 kHz, and the level difference between L1 (peak in the first formant region between 300 and 1,200 Hz) and L0 (peak in the F0 region between 0 and 300 Hz) was measured from long-term average spectra (LTAS, pitch-corrected, unvoiced segments excluded). For the acoustical analyses, the Praat speech analysis system was used [33]. The dose measures were used to quantify vocal loading. For the dose measurements, $Leq_{40\,cm}$ was converted to $Leq_{50\,cm}$ by using the formula $Leq_{50\,cm} = Leq_{40\,cm} - 10 \cdot \log\sqrt{40/50}$. The time resolution for dose measurements was 0.01 s.

Two native listeners (1 native Finnish and 1 native English speaker) listened to each subject's samples in the foreign language and evaluated the speaker's skill, fluency and accuracy in the language on a scale from 0 to 4, with 0 representing no skill/fluency/accuracy and 4 representing native-like speech. A variable reflecting proficiency was created by combining the three variables into a sum of the variable.

The difference between the native language and the foreign language in acoustical parameters was calculated by subtracting the measure in the native language from the corresponding measure in the foreign language, which results in a negative number if the measure is higher in the native language. For statistical analyses, the Wilcoxon signed-rank test was used for testing the significance of difference in parameters between the languages, and Spearman's correlation coefficients for studying the relations between self-reported notions and perceptual evaluations and acoustic parameters. The analyses were carried out with SPSS 18 software.

Results

The majority of subjects (79%) considered that the voice is affected by the language shift. Nineteen percent of those noticed only slight changes. Only 2% of the subjects did not notice any voice changes when speaking the foreign language compared to speaking their mother tongue (table 1).

Only 30% of the subjects reported no effect on voice fatigue while speaking the foreign language (table 1).

The subjects' estimation of the level of experience in speaking the foreign language had a mild negative correlation with the estimation of vocal fatigue ($r = -0.48$, $p < 0.01$). The level of experience correlated with education for the Finnish subjects ($r = 0.55$, $p < 0.05$) and with the duration of residence for the English subjects ($r = 0.66$, $p < 0.01$).

Table 1. Subjects' notions (%) on voice differences, vocal fatigue and quality of voice between native and foreign languages

	The voice changes when speaking the foreign language compared to the native language				
	not at all	slightly	moderately	many changes	profound changes
n = 43	2.3	18.6	44.2	30.2	4.7
	The voice tires faster in the foreign language than in the native language				
	not at all	a little faster	faster	quite fast	much faster
n = 43	30.2	34.9	16.3	7.0	11.6
	Voice quality in the native language				
	very poor	poor	neither good nor poor	good	very good
n = 42	0.0	2.4	33.3	52.4	11.9
	Voice quality in the foreign language				
	very poor	poor	neither good nor poor	good	very good
n = 42	2.3	18.6	44.2	34.9	0.0

Table 2. Native listeners' evaluation (%) of subjects' proficiency in the target language

	No proficiency	Some proficiency	Proficient	Very proficient	Native-like
Finnish (n = 20)	5.0	20.0	30.0	35.0	10.0
English (n = 23)	13.04	30.44	43.48	13.04	0.0
Male (n = 22)	13.64	13.64	40.90	27.27	4.55
Female (n = 21)	4.76	38.10	33.33	19.05	4.76

The majority of the subjects were evaluated as proficient in the foreign language: the Finnish subjects were evaluated more proficient than the English subjects, and males a little more proficient than females (table 2).

No correlation was found between perceived proficiency and the age or sex of the speaker. However, there was a moderate correlation between self-evaluated experience and proficiency ratings by native listeners ($r = 0.54$, $p < 0.01$) and a slight negative correlation between the self-reported notion of vocal fatigue and proficiency ($r = -0.34$, $p < 0.05$). Proficiency correlated moderately with

Table 3. Medians and statistical significance (p values) of differences in acoustical parameters from text reading and spontaneous speech between the native and the foreign language

	F0, Hz		p	$Leq_{50\ cm}$, dB		p	Alpha ratio, dB		p	L1-L0, dB		p
	native	foreign		native	foreign		native	foreign		native	foreign	
Text (n = 43)	154.78	156.50	*	68.10	68.67		–20.36	–20.47		–3.08	–3.75	
Spontaneous (n = 40)	154.96	158.12		67.86	68.11		–19.71	–18.86		–4.32	–3.40	**

	Voiced, %		p	Speech rate, syllables/s		p	D_t, s		p	VLI		p
	native	foreign		native	foreign		native	foreign		native	foreign	
Text (n = 43)	51.39	47.55	**	4.25	2.88	**	30.65	28.48	**	4.54	3.88	
Spontaneous (n = 40)	44.85	37.96	**	3.40	2.50	**	18.91	15.49	**	2.78	2.46	

	D_c, cycles		p	D_d, m		p	D_e, mJ/cm^3		p	D_r, mJ	
	native	foreign		native	foreign		native	foreign		native	foreign
Text (n = 43)	4,535.35	3,877.63	**	15.77	13.75	**	7.52	8.47		0.54	0.48
Spontaneous (n = 40)	2,780.83	2,460.77	**	9.23	7.13	**	4.46	4.01	**	0.26	0.24

* p < 0.5, ** p < 0.01, Wilcoxon signed-rank test.

Table 4. Medians of vocal doses per second and statistical significance (p values) of differences between the native and the foreign language for all subjects, for Finnish and English subjects, and for male and female subjects

	VLI/D_t, kcycles/s		p	D_d/D_t, m/s		p	D_e/D_t, mJ/cm^3/s		p	D_r/D_t, mJ/s		p
	native	foreign		native	foreign		native	foreign		native	foreign	
All												
Text (n = 43)	0.155	0.156	*	0.483	0.481		0.254	0.291		0.016	0.018	
Spontaneous (n = 40)	0.155	0.158		0.471	0.464		0.265	0.268		0.016	0.018	
Finnish												
Text (n = 20)	0.167	0.177	*	0.466	0.474		0.222	0.248		0.017	0.019	
Spontaneous (n = 20)	0.162	0.163		0.463	0.464		0.250	0.259		0.016	0.021	*
English												
Text (n = 23)	0.137	0.123		0.510	0.511		0.324	0.331		0.016	0.015	
Spontaneous (n = 20)	0.133	0.132		0.477	0.463		0.288	0.280		0.015	0.013	
Males												
Text (n = 22)	0.109	0.110		0.466	0.432		0.525	0.474		0.015	0.012	
Spontaneous (n = 20)	0.110	0.109		0.425	0.397		0.453	0.409		0.012	0.010	
Females												
Text (n = 21)	0.1888	0.1889	*	0.539	0.547		0.176	0.180		0.024	0.021	
Spontaneous (n = 20)	0.191	0.187		0.498	0.536		0.152	0.177		0.018	0.028	

* p < 0.05, Wilcoxon signed-rank test.

experience for the Finnish subjects ($r = 0.55$, $p < 0.05$) and slightly with residence for the English subjects ($r = 0.44$, $p < 0.05$).

The acoustical parameters showed some significant differences between the native language and the foreign language. F0 was significantly higher in the foreign language than in the native language in text reading. The L1-L0 level difference showed significant differences between the native language and the foreign language in spontaneous speech only, but there seemed to be a trend that the alpha ratio was lower and the L1-L0 level difference higher in the foreign language compared to the native language. Changes in the alpha ratio and the L1-L0 level difference did not correlate with changes in Leq. The amount of voiced speech and the speech rate were significantly lower in the foreign language than in the native language. The D_t, D_c and D_d dose measures showed significant differences being lower in the foreign language than in the native language (table 3). Age and sex did not correlate with the changes in acoustical parameters.

The dose measurements were gathered from continuous speech with different voicing times, so it was useful to normalize the VLI and D_t. Significant differences were found only in VLI/D_t for all subjects and in D_r/D_t for Finnish subjects. English and male subjects showed no significant changes. Females showed significant changes in VLI/D_t (table 4).

Discussion

Acoustical changes do not always correlate with the subjective sensations of vocal fatigue [18], and a specific cause of subjective sensations is not always apparent. Speaking a foreign language is probably a more stressful task than speaking the native one, which can lead to an increase in phonatory and articulatory effort, causing vocal overloading and, thus, symptoms of vocal fatigue. On the other hand, the mental effort itself may lead to a subjective sensation of vocal overloading [34]. It is also possible that the subjects are not always capable of differentiating between tiredness of the articulators and that of the voice. The subjects of the present study reported that their voice was getting tired faster in the foreign language than in the native language. The quality of voice was also reported as poorer in the foreign language than in the native language. The acoustical parameters which indicate vocal loading showed some evidence that speaking the foreign language showed more loading than speaking the native one. The normalized dose measures showed significant changes only in VLI/D_t in text reading which corresponds to changes in F0 [26]. However, the differences between the native language and the foreign language show a trend that the mean exposure per second tends to be higher in the foreign language than in the native language, which may indicate that speaking a foreign language may, in fact, increase vocal loading. Finnish and female speakers showed a clearer trend towards increased vocal loading due to a larger F0 change. Earlier results by Järvinen et al. [9] suggest that an F0 change in language shift is somewhat gender and language dependent: women seemed to have more distinctive changes in F0 between languages than men, and Finnish speakers more distinctive than English speakers. The trend of a higher L1-L0 level difference in the foreign language may be related to different formant frequencies in the languages [35]. However, it may also indicate a more pressed phonation [36], which has been regarded to increase the risk of vocal loading and voice disorders [34]. It seems that a language shift from the native to the foreign language may influence the speaker's vocal use in such a way that it may cause voice problems in the long run, and, therefore, people who use foreign languages to a large extent, e.g. professionally, may benefit from voice coaching in the target language.

F0, SPL and the duration of voiced speech are considered to be major factors in vocal loading [25] since excessive vibration (in terms of number and amplitude of vocal fold collisions) may cause damage to the vocal fold tissue [26]. In this study, F0 was significantly higher in the foreign language, which may contribute to the vocal loading and to the subjects' notion on it, too. An increase in F0 may also indicate a higher mental load [37]. The results obtained for the amount of voiced speech were in line with previous studies that have shown it to be approximately 50% of the total amount of speech [38] and to be lower in the foreign language than in the native one [26]. The change in the amount of voiced speech can be related to the fact that the speech rate is lower in the foreign language than in the native language, e.g., according to Spilková and van Dommelen [39], the word duration is longer for foreign language than for native language speakers.

A negative correlation between self-reported vocal fatigue and the level of experience in speaking the foreign language was found. This may indicate that the subjects with less experience in speaking the foreign language consider the task somewhat more stressful, and, therefore, their sensitivity to observe symptoms of vocal (and/or articulatory) overloading increases.

It seems natural that the length of residence in a country would improve the spoken language performance, and Flege and Liu [40] have provided such results. Not only the length of residence, but other factors, such as the age when the foreign language studies have begun [5], or the amount of language use [3], may affect the performance in the foreign language. In the present study, a correlation between residence and self-evaluated experience in language use was found for the English subjects, and, for the Finnish subjects, between experience and education. This may indicate that the Finnish subjects considered that experience in a foreign language is acquired by education or that the amount of formal education increases the Finnish speakers' self-confidence as language users. The English subjects had little to no education in Finnish so their experience had been gained through residence in Finland.

The speech doses that are based on voicing time are not necessarily sufficient for a comparative investigation of vocal loading in speaking different languages. Some other parameters, like voice quality, should also be taken into account in the investigation of vocal loading. Thus, perceptual analyses should be included to provide information on the speakers' phonation type [41] and on the possible vocal impairment due to vocal overloading after a prolonged voice use in the foreign language compared to that in the native language. More detailed acoustic analysis, e.g. through inverse filtering, is also warranted.

Previous studies have shown that the type of the speaking task, i.e. text reading versus spontaneous speech, may have an effect on acoustical parameters, such as fundamental frequency [42, 43], but, according to the results of the present study, there was no indication that the subjects' voice use was largely affected by the task. Longer speech samples would be worth studying.

Conclusions

The majority of subjects in the present study experienced that speaking a foreign language causes their voices to get tired faster than speaking the native language. The acoustical vocal doses, however, were lower in the foreign language due to a lower amount of voiced speech. The results suggest that either the reason for increased vocal fatigue in speaking a foreign language is mental stress and not changes in the speaking style or that vocal loading in two languages cannot be sufficiently determined through acoustical dose estimation which is based on the amount of voiced speech.

Acknowledgement

This study was supported by the Emil Aaltonen Foundation.

References

1 Mennen I, Scobbie JM, de Leeuw E, Schaeffler S, Schaeffler F: Measuring language-specific phonetic settings. Second Lang Res 2010;26: 13–41.
2 Levy ES, Strange W: Perception of French vowels by American English adults with and without French language experience. J Phonet 2008;36:141–157.
3 Flege JE, MacKay IRA: Perceiving vowels in a second language. Stud Second Lang Acquis 2004;26:1–34.
4 Flege JE, Schirru C, MacKay IRA: Interaction between the native and second language phonetic subsystems. Speech Commun 2003;40: 467–491.
5 Flege JE, MacKay IRA, Meador D: Native Italian speakers' perception and production of English vowels. J Acoust Soc Am 1999;106: 2973–2987.
6 Ylinen S, Uther M, Latvala A, Vepsäläinen S, Iverson P, Akahane-Yamada R, Näätänen R: Training the brain to weight speech cues differently: a study of Finnish second-language users of English. J Cogn Neurosci 2009;22: 1319–1332.
7 Peltola MS, Kujala T, Tuomainen J, Ek M, Aaltonen O, Näätänen R: Native and foreign vowel discrimination as indexed by the mismatch negativity (MMN) response. Neurosci Lett 2003;352:25–28.
8 Toivanen J: Perspectives on intonation: English, Finnish and English Spoken by Finns; dissertation, University of Oulu, Oulu, 1999.
9 Järvinen K, Laukkanen AM, Aaltonen O: Speaking a foreign language and its effect on F0. Logoped Phoniatr Vocol 2013;38:47–51.
10 Järvinen K, Laukkanen AM, Izdebski I: Voice fundamental frequency changes as a function of foreign language familiarity: an emotional effect? in Izdebski I (ed): Emotions in the Human Voice. San Diego, Plural Publishing, 2008, pp 203–214.
11 Ohara Y: Performing gender through voice pitch: a Cross-Cultural Analysis of Japanese and American English; in Pasero U, Braun F (eds): Perceiving and Performing Gender. Opladen/Wiesbaden, Westdeutscher Verlag, 1999, pp 105–116.
12 Ullakonoja R: Comparison of pitch range in Finnish (L1) fluency and Russian (L2); in Trouvain J, Barry WJ (eds): The Proceedings of the 16th International Congress of Phonetic Sciences, 6–10 August 2007, Saarbrücken. Saarbrücken, Universität des Saarlandes, 2007, pp 1701–1704.
13 Bruyninckx M: Language-induced voice quality variability in bilinguals. J Phonet 1994; 22:19–31.
14 Harmegnies B, Bruyninckx M, Llisterri J, Poch D: Effects of language change on voice quality in bilingual speakers. Corpus content effect. Eurospeech 1991. Proc 2nd Eur Conf Speech Commun Technol, Genova, 1991, pp 165–168.
15 Izdebski K: Clinical voice assessment: the role and value of the phonatory function studies; in Lalwani A: Current Diagnosis and Treatment. Otolaryngology Head and Neck Surgery. New York, McGraw-Hill/Lange, 2007, pp 416–429.
16 Vilkman E: Occupational safety and health aspects of voice and speech professions. Folia Phoniatr Logop 2004;56:220–253.
17 Jiang JJ, Titze IR: Measurement of vocal fold intraglottal pressure and impact stress. J Voice 1994;8:132–144.

18 Laukkanen AM, Ilomäki I, Leppänen K, Vilkman E: Acoustic measures and self-reports of vocal fatigue by female teachers. J Voice 2008;22:283–289.
19 Rantala L, Paavola L, Körkkö P, Vilkman E: Working-day effects on the spectral characteristics of teaching voice. Folia Phoniatr Logop 1998;50:205–211.
20 Laukkanen AM, Syrjä T, Laitala M, Leino T: Effects of two-month vocal exercising with and without spectral biofeedback on student actors' speaking voice. Logped Phoniatr Vocol 2004;29:66–76.
21 Buekers R, Bierens E, Kingma H, Marres EHMA: Vocal load as measured by the voice accumulator. Folia Phoniatr Logop 1995;47:252–261.
22 Rantala L, Vilkman E: Relationship between subjective voice complaints and acoustic parameters in female teachers' voices. J Voice 1999;13:484–495.
23 Chang A, Karnell MP: Perceived phonatory effort and phonation threshold pressure across a prolonged voice loading task: a study of vocal fatigue. J Voice 2004;18:454–466.
24 Lohscheller J, Doellinger M, McWhorter AJ, Kunduk M: Preliminary study on the quantitative analysis of vocal loading effects on vocal fold dynamics using phonovibrograms. Ann Otol Rhinol Laryngol 2008;117:484–493.
25 Švec JG, Popolo PS, Titze IR: Measurement of vocal doses in speech: experimental procedure and signal processing. Logoped Phoniatr Vocol 2003;28:181–192.
26 Titze IR, Švec JG, Popolo PS: Vocal dose measures: quantifying accumulated vibration exposure in vocal fold tissues. J Speech Lang Hear Res 2003;46:919–932.
27 Welham NV, Maclagan MA: Vocal fatigue: current knowledge and future directions. J Voice 2003;17:21–30.
28 Vilkman E: Occupational safety and health aspects of voice and speech professions. Folia Phoniatr Logop 2004;56:220–253.
29 Kangasvieri T, Miettinen E, Kukkohovi P, Härmälä M: Kielten tarjonta ja kielivalintojen perusteet. Tilannekatsaus joulukuu 2011. Opetushallituksen muistio 2013. (Supply of languages and basis for language choices in basic education. Status survey December 2011. Memo of the Finnish National Board of Education 2013). 2013. http://www.oph.fi/download/138072_Kielten_tarjonta_ja_kielivalintojen_perusteet_perusopetuksessa.pdf (accessed July 14, 2014).
30 Wiik K: Finnish and English Vowels; dissertation, Turun yliopisto, Turku, 1965.
31 Vihanta VV: Suomi vieraana kielenä foneettiselta kannalta (Finnish as a foreign language from the phonetic point of view; in Finnish, abstract in English); in Tommola J (ed): Vieraan kielen ymmärtäminen ja tuottaminen (Foreign Language Comprehension and Production). Turku, AFinLA Yearbook, 1990, pp 199–225. https://helda.helsinki.fi/bitstream/handle/10224/3668/vihanta199–225.pdf?sequence=2 (accessed January 14, 2015).
32 Suomi K, Toivanen J, Ylitalo R: Finnish Sound Structure. Phonetics, Phonology, Phonotactics and Prosody. Studia Humaniora Ouluensia. Oulu, University of Oulu, 2008.
33 Boersma P, Weenink D: Praat: doing phonetics by computer (version 5.3.39). http://www.fon.hum.uva.nl/praat/ (accessed August 16, 2013).
34 Solomon NP: Vocal fatigue and its relation to vocal hyperfunction. Int J Speech Lang Pathol 2008;10:254–266.
35 Andrianopoulos MV, Darroe K, Chen J: Multimodal standardization of voice among four multicultural populations formant structures. J Voice 2001;15:61–77.
36 Kitzing P: LTAS criteria pertinent to the measurement of voice quality. J Phonet 1986;14:477–482.
37 Johannes B, Wittels P, Enne R, Eisinger G, Castro CA, Thomas JL, Adler AB, Gerzer R: Non-linear function model of voice pitch dependency on physical and mental load. Eur J Appl Physiol 2007;101:267–276.
38 Löfqvist A, Mandersson B: Long-time average spectrum of speech and voice analysis. Folia Phoniatr Logop 1987;39:221–229.
39 Spilková H, van Dommelen WA: English 'of' in L1 and L2 speakers' read and spontaneous speech. http://131.111.166.124/images/uploads/SpilkovavanDommelen2010_English_of.pdf (accessed August 1, 2014).
40 Flege JE, Liu S: The effect of experience on adults' acquisition of a second language. Stud Second Lang Acquis 2001;23:527–552.
41 Kankare E, Dong L, Laukkanen AM, Geneid A: EGG and acoustic analyses of different voice samples: comparison between perceptual evaluation and voice activity and participation profile. Folia Phoniatr Logop 2013;65:98–104.
42 Johns-Lewis C: Prosodic differentiation of discourse modes; in Johns-Lewis C (ed): Intonation in Discourse. London, Croom Helm, 1986, pp 199–219.
43 Hudson AI, Holbrook A: Fundamental frequency characteristics of young black adults. Spontaneous speaking and oral reading. J Speech Hear Res 1982;25:25–28.

Voice Quality in Native and Foreign Languages Investigated by Inverse Filtering and Perceptual Analyses

*Kati Järvinen, *Anne-Maria Laukkanen, and †Ahmed Geneid, *Tampere and †Helsinki, Finland

Summary: Objectives. Language shift from native (L1) to foreign language (L2) may affect speaker's voice production and induce vocal fatigue. This study investigates the effects of language shift on voice source and perceptual voice quality.

Study Design. This is a comparative experimental study.

Subjects and Methods. Twenty-four subjects were recorded in L1 and L2. Twelve of the subjects were native Finnish speakers and 12 were native English speakers, and the foreign languages were English and Finnish. Two groups were created based on reports of fatigability. Group 1 had the subjects who did not report more vocal fatigue in L2 than in L1, and in group 2 those who reported more vocal fatigue in L2 than in L1. Acoustic analyses by inverse filtering were conducted in L1 and L2. Also, the subjects' voices were perceptually evaluated in both languages.

Results. Results show that language shift from L1 to L2 increased perceived pressedness of voice. Acoustic analyses correlated with the perceptual evaluations. Also, the subjects who reported more vocal loading had poorer voice quality, more strenuous voice production, more pressed phonation, and a higher pitch.

Conclusions. Voice production was less optimal in L2 than in L1. Speech training given in L2 could be beneficial for people who need to use L2 extensively.

Key Words: L1–L2–voice source–vocal fatigue.

INTRODUCTION

Speaking a foreign language (L2) can impose a number of challenges for the non-experienced speakers. Such challenges are not only limited to the mental stress associated with speaking a foreign language but also include possible effects on voice that are not encountered while speaking the native language (L1).

According to earlier studies[1,2] and clinical observations, people often report experiencing more symptoms of vocal fatigue when they speak a foreign language than when they speak their native language. It has been hypothesized that the experience may be related to the fact that more mental stress is associated with speaking a foreign language.[3] Therefore, either all sensations of fatigue may be intensified, or mental stress itself results in a more taxing speech production (increased phonatory and articulatory effort), which causes overloading of the vocal organ.

Whether or not mental stress is involved, a shift from one language to another is prone to cause changes in voice and speech parameters. In earlier studies, for instance, a change in mean fundamental frequency and pitch range has been reported.[4–8] Changes in voice spectrum have also been observed, as measured by the shape and slope of long-term average spectrum (LTAS), and the level difference between the first harmonics and noise-to-harmonics ratio.[4,9,10]

Mechanical loading posed on the vocal folds increases with fundamental frequency (F0), intensity, and degree of adduction of the vocal folds.[11] The degree of adduction corresponds to phonation type (breathy having a low adduction and pressed showing a high adduction). The amount of vocal loading has been quantified by calculating vocal doses.[12,13] Vocal doses are measures based on fundamental frequency, sound pressure level, and voicing time. A previous study[1] applied vocal doses on the study of vocal loading in L1 and L2. Contrary to the expectations, the vocal doses decreased in L2. This was due to the fact that the total duration of voiced speech decreased significantly when speaking a foreign language. The time of voicing may be related to speech tempo and also to the ratio between voiced and voiceless sounds in a particular language. Therefore, calculation of doses as such does not seem to be sufficient to reflect vocal loading in speaking a foreign language. However, it was found in Järvinen et al[1] that the mean exposure *per second* tended to be somewhat higher in L2 than in L1, which suggests that speaking a foreign language may be more loading than speaking L1. LTAS results also seemed to point toward the same direction. A higher level difference between the peak in the first formant region between 300 and 1200 Hz and the peak in the F0 region between 0 and 300 Hz in L2 was found,[1] possibly indicating a more pressed phonation. That, in turn, is considered to be one factor in vocal loading, in addition to high pitch and intensity, and excessively prolonged voice use.[14,15]

The present study continues the previous one. Here the phonation type is studied by applying inverse filtering, which reveals voice source, that is, the airflow pulses that are generated by the vibration of the vocal folds.[16] Inverse filtering was developed in the late 1950s by Miller.[17] The source-filter approach is based on the ideas that the source and the filter are independent of each other, and speech consists of three separate and independent processes: glottal excitation, vocal tract filter, and lip radiation.[18] Rothenberg[19] introduced an inverse filtering method that estimates the airflow out of the mouth through a mask that avoids

Accepted for publication May 5, 2016.

From the *Speech and Voice Laboratory, School of Education, University of Tampere, Tampere, Finland; and the †Department of Otolaryngology and Phoniatrics, Head and Neck Surgery, Helsinki University Hospital and University of Helsinki, Helsinki, Finland.

Address correspondence and reprint requests to Kati Järvinen, Speech and Voice Laboratory, School of Education, University of Tampere, 33014 Tampere, Finland. E-mail: jarvinen.kati.h@student.uta.fi

Journal of Voice, Vol. ■■, No. ■■, pp. ■■-■■
0892-1997

http://dx.doi.org/10.1016/j.jvoice.2016.05.003

the lip radiation effect. The iterative adaptive inverse filtering (IAIF), developed by Alku,[20] is one of the methods where the glottal flow signal is estimated from a corresponding acoustic speech pressure signal.

The following are the research questions of the present study: (1) Is the perceived phonation quality more pressed when speaking a foreign language (L2) than when speaking the native one (L1). (2) Do the characteristics of the voice source (revealed through inverse filtering) differ between L1 and L2? (3) Do the subjects who report more vocal fatigue in L2 than in L1 have more pressed phonation in L2?

METHODS

Participants and recordings

Originally, 43 subjects in total participated in the study. Twenty of the subjects were native Finnish speakers and 23 were native English speakers, and the foreign languages used were either English or Finnish. A questionnaire about vocal fatigue experienced when speaking L1 and L2 was filled in by all the subjects. Based on the answers to the questionnaire, we picked out 12 subjects (six men and six women) who reported that language shift from L1 to L2 does not increase vocal fatigue (group 1, G1), and another 12 subjects (also 6 men and 6 women) who instead reported more vocal fatigue in L2 than in L1 (group 2, G2). In G1, three of the subjects were native English speakers and nine were native Finnish speakers, and in G2 nine of the subjects were native English speakers and three were native Finnish speakers. The mean age for G1 was 31.6 (standard deviation [SD]: 9.1), and for G2 it was 35.6 years (SD: 15.4). The subjects in G1 considered themselves as more experienced in speaking L2 than the subjects in G2: 50% of the subjects in G1 considered themselves as very experienced or native like, whereas 83% of the subjects in G2 considered themselves to have only some experience.

The recordings took place in a well-damped recording studio using a combined level meter and microphone (Bruel et Kjær Mediator, Type 2206, Copenhagen, Denmark). The microphone was placed 40 cm from the mouth. *Sound Forge* software (Sony Creative Software Inc, Middleton WI, USA) was used. The input frequency was 44,100 Hz, and the amplitude quantization was 16 bits. One-minute text reading and spontaneous speech samples in L1 and L2 were recorded. The texts were the same in content in both languages, and the spontaneous speech samples were recorded from a description of a cartoon using the same picture for both languages.

Perceptual analyses

From the speech samples, one sentence (from 5 to 8 seconds in duration) was extracted for the listening analysis. The text extracts for L1 and L2 were taken from the same place in the text. Three experienced vocologists listened to each subject's samples in L1 and L2. The samples were replayed in random order, but the same speaker's L1 and L2 samples were always presented in pairs. Half of the samples had L1 sample first, and half L2 sample first. Two persons' samples in text reading and in spontaneous speech were repeated in order to study the intra-rater reliability of the perceptual analysis. The listeners first listened to the text reading samples and then to the spontaneous speech samples. After every sample pair, the listeners set a mark on a scale from 0 to 10 reflecting four voice characteristics: general voice quality (0 = very poor – 10 = very good), strenuousity of voice production (0 = very strenuous – 10 = very easy), firmness of phonation (0 = very pressed – 10 = very breathy), and suitability of the pitch for the speaker (0 = too low – 10 = too high).

Acoustic analyses

Short stressed vowels (/a/ in Finnish and /ʌ/ in English) were extracted in three different phonetic contexts both in text reading and in spontaneous speech. The length of the extracted vowels was 0.045 seconds. The extracted vowels were inverse-filtered by *TKK Aparat* software,[21] which has been developed for IAIF inverse filtering of speech pressure signal.[20] The formant and lip radiation effects were manually set for every vowel in the inverse filtering. The low frequency noise cutoff was set at 80 Hz, and changed if necessary. In the acoustic analyses, the mean of each acoustic parameter of the three vowels per text type was calculated for every subject.

The following time and amplitude-based parameters were studied from the inverse-filtered signal (Figure 1): (1) open quotient (OQ), which measures the glottal open phase in comparison to cycle duration; (2) closing quotient (ClQ), which is the ratio between the duration of the glottal closing phase to the period length; (3) speed quotient (SQ), which is the ratio between the duration of the opening phase and the duration of the closing phase, and two amplitude-based parameters; (4) amplitude quotient (AQ), which is the ratio of the peak-to-peak amplitude of the flow (Aac) and the minimum peak of the pulse derivative (Admin); and (5) normalized amplitude quotient (NAQ), which is the AQ divided by period length.[16,20,22,23]

It is common to find that the opening phase has rather two opening instants (t_{o1} and t_{o2} in Figure 1) after the closed phase with a possible (knee) shaped waveform in the beginning of the opening phase. Accordingly, OQ1 is calculated from the primary opening instant, whereas OQ2 is calculated from the secondary opening instant.[24] Similarly, SQ1 is calculated from the primary opening instant, whereas SQ2 is calculated from the secondary opening instant. If a flow pulse did not show two opening instants, then the two opening instants were actually one rather than two separate ones. In such a case, OQ1 = OQ2 and SQ1 = SQ2.

According to earlier studies, phonation quality is signaled by various voice source parameters as follows. Breathy voice implies increased open quotient.[25–27] The properties of the closing phase most directly affect the voice quality as the abrupt closure of the glottis at the end of the closing phase involves the majority of the voicing energy.[17] When intensity and pressedness increase, ClQ decreases.[25] Amplitude-based AQ and NAQ have been found to correlate negatively with pressedness of voice.[20,28]

Statistical analyses

Median and interquartile range were calculated for each parameter because parameters were not normally distributed. Wilcoxon signed-rank test was used for testing the statistical significance of differences between L1 and L2 in perceptual analyses.

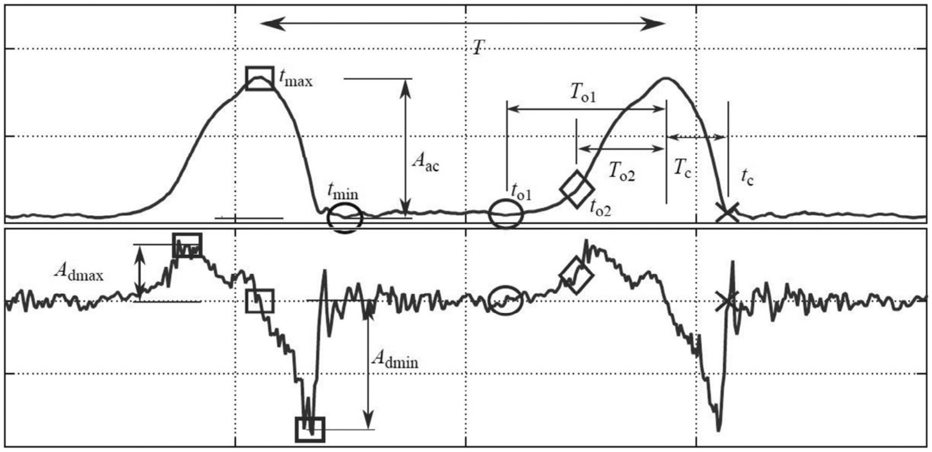


FIGURE 1. Time and amplitude instants used in calculating the parameters from inverse-filtered signal. The upper part shows the glottal flow estimate, whereas the lower part shows the respective derivative. The parameters used in this study are calculated as follows: $OQ1 = \frac{T_{o1}+T_c}{T}$, $OQ2 = \frac{T_{o2}+T_c}{T}$, $ClQ = \frac{T_c}{T}$, $SQ1 = \frac{T_{o1}}{T_c}$, $SQ2 = \frac{T_{o2}}{T_c}$, $AQ = \frac{A_{ac}}{Admin}$, $NAQ = \frac{AQ}{T}$.

Mann-Whitney *U* test was used for statistical significance of the differences between G1 and G2. Repeated measures analysis of variances (RM-ANOVA) was used for analyzing the differences in inverse filtering parameters between L1 and L2 and the effect of subjective vocal fatigue. For RM-ANOVA, the acoustic results from text reading and spontaneous speech were combined, and the mean of each parameter was calculated for L1 and L2. The inter-rater reliability was studied with Cronbach's alpha reliability coefficient, and intra-rater reliability was studied with Spearman correlation coefficient. Spearman correlation coefficient was also used for studying the relations between the acoustic and perceptual results. Significance level was set to $P < 0.05$ in all statistical analyses. The statistical analyses were carried out with *SPSS 22* software (IBM SPSS Statistics v. 22 for Windows, Armonk, NY).

RESULTS

Perceptual results

The reliability of the perceptual analysis was found satisfactory. The inter-rater reliability for listeners was quite good, Cronbach's alpha = 0.70. The intra-rater reliability for voice quality was r = 0.83, $P < 0.001$, for strenuousity of voice production r = 0.87, $P < 0.001$, for firmness r = 0.53, $P < 0.001$, and for suitability of pitch r = 0.47, $P < 0.05$.

Evaluation of voice in L1 and L2 showed significant differences; in text reading, the voice quality was poorer in L2 ($P < 0.05$), and in text reading and in spontaneous speech firmness of phonation was lower (ie, type of phonation more pressed) in L2 than in L1 (Table 1).

Acoustic results

The results of the analysis were first statistically investigated gender-wise, but because no significant differences between the gender were found except for F0, the male and female subjects' other parameter values are presented here together.

Acoustic analyses showed significant differences between L1 and L2 in OQ2 in text reading, and in ClQ, SQ2, and NAQ in spontaneous speech (Table 2).

RM-ANOVA showed significant differences between L1 and L2 in ClQ ($F = 6.20$, df = 1, $P < 0.05$), NAQ ($F = 5.32$, df = 1, $P < 0.05$), and SQ2 ($F = 13.81$, df = 1, $P < 0.01$), but vocal fatigue set as a between-subject factor showed no significant differences.

Differences between group 1 and group 2

For subjects in G1, voice quality in L2 was evaluated as better ($P = 0.017$), firmness of phonation lower (ie, phonation type less pressed) ($P = 0.001$), and pitch more suitable ($P = 0.01$) than for subjects in G2 (Figure 2). Significant differences between L1 and L2 in text reading were found in all the perceptual

TABLE 1.
Median and Interquartile Range (IQR) for Perceptual Analyses (*N* = 24), Wilcoxon Signed-Rank Test, Significance Level (0.05)

		Text Reading		Spontaneous Speech	
		L1	L2	L1	L2
Quality	Median	4.33	4.00*	4.33	4.33
	IQR	0.33	0.75	1.08	1.33
Strenuousity	Median	4.67	4.67	4.67	4.33
	IQR	1.08	2.00	1.33	1.42
Firmness	Median	4.33	3.83**	4.50	4.17**
	IQR	1.00	1.08	0.67	0.67
Pitch	Median	5.00	5.00*	5.00	5.00
	IQR	0.67	1.00	0.67	1.08

Notes: Voice quality (very poor = 0 – very good = 10), voice production (very strenuous = 0 – very easy = 10), firmness of phonation (very pressed = 0 – very breathy = 10), and pitch suitability (too low = 0 – too high = 10).
* $P < 0.05$, ** $P < 0.01$.

TABLE 2.
Median and IQR for Acoustic Analyses in Text Reading (*N* = 24) and Spontaneous Speech (*N* = 24), Wilcoxon Signed-Rank Test

		Text Reading		Spontaneous Speech	
		L1	L2	L1	L2
F0, men (*N* = 12)	Median	116.67	114.23	124.29	117.41
	IQR	40.09	23.99	64.10	34.14
F0, women (*N* = 12)	Median	208.60	218.04	195.12	191.91
	IQR	25.95	30.35	58.48	24.14
OQ1	Median	0.78	0.76	0.82	0.78
	IQR	0.14	0.19	0.11	0.14
OQ2	Median	0.57	0.58*	0.53	0.55
	IQR	0.12	0.12	0.14	0.11
ClQ	Median	0.22	0.23	0.24	0.21*
	IQR	0.06	0.05	0.08	0.04
SQ1	Median	2.49	2.55	2.54	2.66
	IQR	0.63	0.77	0.99	0.79
SQ2	Median	1.51	1.62	1.43	1.70*
	IQR	0.39	0.59	0.45	0.69
NAQ	Median	0.09	0.09	0.09	0.08**
	IQR	0.02	0.02	0.02	0.02
AQ	Median	0.0006	0.0006	0.0006	0.0006
	IQR	0.0004	0.0004	0.0003	0.0003

* $P < 0.05$, ** $P < 0.01$.
Abbreviations: OQ, open quotient; ClQ, closing quotient; SQ, speed quotient; NAQ, normalized amplitude quotient; AQ, amplitude quotient; IQR, interquartile range.

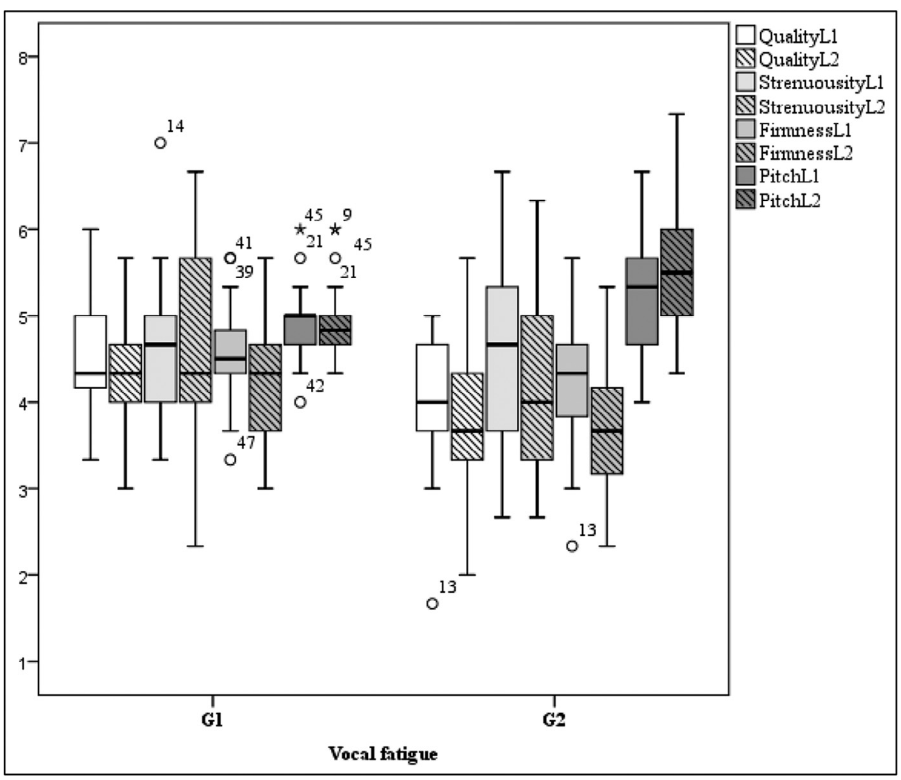


FIGURE 2. Perceptual evaluation of subjects' voices in L1 and L2. G1: subjects who did not report more vocal fatigue in L2 than in L1, G2: subjects who reported more vocal fatigue in L2 than in L1. Evaluated characteristics: general voice quality (0 = very poor – 10 = very good), strenuousity of voice production (0 = very strenuous – 10 = very easy), firmness of phonation (0 = very pressed – 10 = very breathy), and suitability of the pitch for the speaker (0 = too low – 10 = too high).

TABLE 3.
Medians and Interquartile Range (IQR) for Perceptual Analyses for G1 (N = 12) and G2 (N = 12), Wilcoxon Signed-Rank Test, Significance Level (0.05)

		Text Reading				Spontaneous Speech			
		G1		G2		G1		G2	
		L1	L2	L1	L2	L1	L2	L1	L2
Quality	Median	4.33	4.33	4.17	3.67*	4.50	4.50	4.00	4.00
	IQR	0.25	0.92	0.92	0.92	1.50	0.92	1.33	1.58
Strenuousity	Median	4.50	5.00	4.83	3.67**	4.67	4.33	4.50	4.17
	IQR	1.25	1.92	1.67	1.92	1.17	1.58	1.58	1.50
Firmness	Median	4.50	4.50	4.17	3.33*	4.50	4.17*	4.50	4.17*
	IQR	0.67	1.67	0.92	1.17	0.58	0.67	1.17	1.25
Pitch	Median	5.00	4.83	5.33	5.50*	5.00	4.83	5.17	5.50
	IQR	0.33	0.58	0.92	1.17	0.33	0.58	1.00	1.50

Notes: Voice quality (0 = very poor – 10 = very good), strenuousity of voice production (0 = very strenuous – 10 = very easy), firmness of phonation (0 = very pressed – 10 = very breathy), and pitch suitability (0 = too low – 10 = too high).
* $P < 0.05$, ** $P < 0.01$.

parameters for G2: Voice quality was evaluated poorer, phonation more pressed, voice production more strenuous, and pitch higher than suitable in L2. In spontaneous speech, a significant increase in firmness in L2 was found for both G1 and G2. Medians and interquartile ranges are presented in Table 3.

Significant differences were found within the same group for some of the inverse filtering parameters. G1 showed significant reductions in both NAQ and ClQ, and G2 showed significant increase in SQ2 for L2. These changes for both groups were found in spontaneous speech (Table 4). Text reading showed the same trend, but it remained insignificant. Medians and interquartile ranges are presented in Table 4.

The inverse filtering parameters did not differ significantly between G1 and G2 either in L1 or L2.

The perceptual evaluation of strenuousity of voice production correlated mildly with the acoustic inverse-filtered parameters; the correlation varied from r = 0.26, $P < 0.01$ for OQ1 to r = 0.47, $P < 0.001$ for NAQ. The evaluated strenuousity of voice pro-

TABLE 4.
Median and IQR for Acoustic Analyses for G1 (N = 12) and G2 (N = 12) in Text Reading and Spontaneous Speech, Wilcoxon Signed-Rank Test

		Text Reading				Spontaneous Speech			
		Group 1		Group 2		Group 1		Group 2	
		L1	L2	L1	L2	L1	L2	L1	L2
F0, men (N = 6)	Median	122.87	114.23	115.72	115.23	124.29	120.29	120.03	112.10
	IQR	35.33	24.84	66.13	39.35	60.67	46.14	39.92	33.51
F0, women (N = 6)	Median	202.31	218.94	209.32	218.04	195.12	191.34	189.05	199.22
	IQR	39.34	38.12	24.25	42.94	60.03	12.26	64.75	53.26
OQ1	Median	0.78	0.74	0.78	0.78	0.81	0.75	0.82	0.78
	IQR	0.16	0.19	0.13	0.17	0.11	0.15	0.12	0.12
OQ2	Median	0.57	0.57	0.56	0.58	0.58	0.54	0.52	0.56
	IQR	0.12	0.11	0.11	0.18	0.14	0.08	0.15	0.12
CIQ	Median	0.22	0.23	0.21	0.22	0.24	0.21*	0.23	0.21
	IQR	0.04	0.03	0.07	0.07	0.07	0.04	0.08	0.04
SQ1	Median	2.46	2.45	2.66	2.65	2.43	3.02	2.68	2.50
	IQR	0.37	0.49	0.97	1.03	0.51	0.71	1.06	0.39
SQ2	Median	1.51	1.58	1.51	1.75	1.39	1.57	1.47	1.70*
	IQR	0.39	0.49	0.34	0.62	0.40	0.46	0.63	0.79
NAQ	Median	0.10	0.09	0.09	0.09	0.09	0.08**	0.09	0.08
	IQR	0.02	0.01	0.03	0.02	0.02	0.02	0.02	0.02
AQ	Median	0.0007	0.0006	0.0006	0.0007	0.0007	0.0006	0.0006	0.0007
	IQR	0.0003	0.0003	0.0004	0.0005	0.0002	0.0002	0.0002	0.0004

* $P < 0.05$, ** $P < 0.01$.
Abbreviations: OQ, open quotient; CIQ, closing quotient; SQ, speed quotient; NAQ, normalized amplitude quotient; AQ, amplitude quotient; IQR, interquartile range.

duction had a mild negative correlation with SQ1 ($r = -0.32$, $P < 0.01$) and SQ2 ($r = -0.22$, $P < 0.05$). Also, the evaluation of pitch suitability correlated mildly with F0 ($r = 0.41$, $P < 0.01$) and NAQ ($r = 0.30$, $P < 0.01$), and negatively with AQ ($r = -0.28$, $P < 0.01$).

DISCUSSION

Languages as such may differ in the type of phonation.[4] The results of the present study, however, did not seem to reflect differences between languages as such, but L2 seemed to differ from L1 independent of the language. The perceptual differences between speaking L2 and L1 were obvious. Changes from L1 to L2 resulted in poorer voice quality, more pressed and strenuous voice production, and higher pitch.

Although the perceptual assessment was randomized, some listeners still stated that it was easy to notice mispronunciation or strong accent in some L2 samples. It is possible to speculate whether this may have had some effect on the assessment of voice quality and strenuousity of voice production. However, the listeners were experienced in perceptual analysis of voice regardless of language and fluency of speech. The perceptual evaluation of some vocal features may be dependent on the language background of the listeners. Yamaguchi et al[29] found asthenia and strain to be evaluated differently by listeners from different languages, and Ghio et al[30] reported the same finding for roughness, while neither of the studies found any influence of language background on the perceptual ratings of breathiness in the GRBAS scale, which is a common tool for subjective evaluation of voice, the GRBAS scale describes voice in five parameters: overall voice performance (G—grade); roughness (R), breathiness (B), asthenia (A), and strain (S). In the present study, all the listeners were native Finnish speakers, and accordingly the cultural and language background was homogenous among them in a way that the possible influence of cultural and language background on the perception of voice quality was avoided.[31] Further study with native English listeners is warranted.

The acoustic differences between L1 and L2 included a decrease in NAQ and ClQ in L2, which points to increased pressedness of voice in L2.[20,25–28] Acoustic analyses and perceptual evaluations also correlated with each other. Voice production was evaluated as more strenuous when acoustic parameters indicated a more pressed voice. Furthermore, voices that were evaluated as having higher pitch correlated with a decrease in AQ, suggesting that a raised pitch was at least in some cases related to increased pressedness of voice. More pressed phonation increases vocal loading,[20] and thus may result in subjective sensations of vocal fatigue. However, more pressed phonation does not necessarily singly indicate total vocal loading, as also voicing time, F0, and intensity are important factors in vocal loading.[14] Acoustic parameters based on voicing time, F0, and intensity are not sufficient in studying vocal loading between two languages due to the fact that the amount of voiced speech can vary vastly between languages.[1] According to the results in the present study, perceptual voice quality ratings and investigation of the voice source characteristics may be an important addition to studying differences in vocal loading between L1 and L2.

Subjects who reported more sensation of vocal fatigue in L2 did not differ from the others in terms of acoustic parameters. According to Laukkanen et al,[32] subjective sensation of vocal fatigue does not always correlate with the objective measurements of acoustic parameters, and it can sometimes be hard to find a specific reason for vocal fatigue. The results show that subjects with more subjective sensations of vocal fatigue in L2 had less optimal voice use also in L1, which may contribute to the sensation of vocal fatigue.

By examining the inverse filtering parameters of G1 and G2 separately, we found that both groups showed a similar trend like the previously reported, which points to increased pressedness in L2. However, surprisingly, this trend was only significant for G1 rather than G2. This may indicate that the subjects in G1 were not as familiar with the symptoms of vocal loading. As Rantala[33] has found, the rise of F0 and leveling of the spectrum during a vocally loading working day occurred in the group of subjects with less complaints of vocal fatigue symptoms rather than in the group with more complaints. She discussed that the result may reflect a normal adaptation of the human body, where the subjects with more complaints have a tendency to shift toward a more hypofunctional voice use to avoid excessive strain and exhaustion. According to the results from the perceptual evaluations, subjects in G2 had poorer voice quality, more pressed phonation, more strenuous voice production, and higher pitch in L1 than subjects in G1. It is possible that subjects with less optimal voice use in L1 may, in fact, experience more symptoms of vocal fatigue even with small changes in voice use in L2.

In this study, the two groups differed from each other as subjects in G1 were more experienced in speaking L2 than the subjects in G2. Lack of experience in speaking L2 may cause increase in psycho-physiological stress and mental effort, which may lead to more muscle tension and therefore to higher F0 and increased pressedness in voice. These factors may contribute to the increase of subjective sensation of vocal overloading.[1,3] However, F0 was lower in L2 for subjects in G2. This can be due to the composition of the groups. In G2, 9 of the 12 subjects were native English speakers, whereas in G1 9 of the subjects were native Finnish speakers. It is possible that the subjects tried to reach a certain level of pitch they thought the native speaker of the target language would use,[5–7] as languages have been found to differ in the mean F0.[34] Further study with more subjects is necessary to exclude this effect.

Laver[35] suggests that speakers have articulatory and phonatory settings, and neutral phonatory settings consist of modal phonation in voicing where the average muscular tension throughout the vocal apparatus should be moderate. Speaking L2 seems to impose additional vocal challenges that make them use less optimal voice production and increase the susceptibility to vocal overloading and vocal fatigue, and also causes perceivable changes in acoustic characteristics of the voice.

It is probable that people who must speak extensively in L2 (professional interpreters for example) may benefit from teaching in speech technique in L2, not only because they need to use voice to great extent, but also because the vocal loading seems to increase when speaking L2 compared with speaking L1. Although speech technique in L1 may be sufficient, it does not always transfer to L2, which has a different structure and phonatory and articulatory settings.

CONCLUSIONS

Language shift caused voice quality changes between L1 and L2 that are markedly perceptually observable, as speakers' voices were evaluated as more pressed in L2 than in L1. Acoustic findings correlated with the perceptual results. Decrease in NAQ and ClQ in L2 compared with L1 indicated increased pressedness of voice. The speakers who report more vocal fatigue in L2 than in L1 were also evaluated as having poorer, more pressed and strenuous voice quality, and higher pitch in both L1 and L2 than the speakers who did not report having more vocal fatigue in L2 than in L1.

Acknowledgment

This study was supported by Emil Aaltosen Säätiö (Emil Aaltonen Foundation).

REFERENCES

1. Järvinen K, Laukkanen AM. Vocal loading in speaking a foreign language. *Folia Phoniatr Logop*. 2015;67:1–7.
2. Izdebski K. Clinical voice assessment: the role and value of the phonatory function studies. In: Lalwani A, ed. *Current Diagnosis and Treatment. Otolaryngology Head and Neck Surgery*. New York: McGraw-Hill/Lange; 2007:416–429.
3. Solomon NP. Vocal fatigue and its relation to vocal hyperfunction. *Int J Speech Lang Pathol*. 2008;10:254–266.
4. Ferreira E. APP. Cross-linguistic effects on voice quality: a study on Brazilians' production of Portuguese and English. Proceedings of the International Symposium on the Acquisition of Second Language Speech Concordia Working Papers in Applied Linguistics. 2014. Available at: http://doe.concordia.ca/copal/documents/13_Engelbert_Vol5.pdf. Accessed July 21, 2015.
5. Järvinen K, Laukkanen AM, Aaltonen O. Speaking a foreign language and its effect on F0. *Logopen Phoniatr Vocol*. 2013;38:47–51.
6. Järvinen K, Laukkanen AM, Izdebski I. Voice fundamental frequency changes as a function of foreign language familiarity: an emotional effect? In: Izdebski I, ed. *Emotions in the Human Voice*. San Diego, CA: Plural Publishing; 2008:203–214.
7. Ohara Y. Performing gender through voice pitch: a cross-cultural analysis of Japanese and American English. In: Pasero U, Braun F, eds. *Perceiving and Performing Gender*. Opladen/Wiesbaden: Westdeutscher Verlag; 1999:105–116.
8. Ullakonoja R. Comparison of pitch range in Finnish (L1) fluency and Russian (L2). In: Trouvain J, Barry WJ, eds. *Proceedings of the 16th International Congress of Phonetic Sciences*. Saarbrücken, Germany: ICPhS; 2007:1701–1704.
9. Bruyninckx M, Harmegnies B, Llisterri J, et al. Language-induced voice quality variability in bilinguals. *J Phon*. 1994;22:19–31.
10. Harmegnies B, Bruyninckx M, Llisterri J, et al. Effects of language change in voice quality in bilingual speakers. Corpus content effects. In: *Eurospeech 1991. Proceedings of the 2nd European Conference on Speech Communication and Technology*, Vol. 1. Genova, Italy; ISCA; 1991:165–168.
11. Titze IR, Jiang JJ, Lin E. The dynamics of length change in canine vocal folds. *J Voice*. 1997;11:267–276.
12. Švec JG, Popolo PS, Titze IR. Measurement of vocal doses in speech: experimental procedure and signal processing. *Logoped Phoniatr Vocol*. 2003;28:181–192.
13. Titze IR, Švec JG, Popolo PS. Vocal dose measures: quantifying accumulated vibration exposure in vocal fold tissues. *J Speech Lang Hear Res*. 2003;46:919–932.
14. Vilkman E. Occupational safety and health aspects of voice and speech professions. *Folia Phoniatr Logop*. 2004;56:220–253.
15. Jiang JJ, Titze IR. Measurement of vocal fold intraglottal pressure and impact stress. *J Voice*. 1994;8:132–144.
16. Airas M. Methods and studies of laryngeal voice quality analysis in speech production [academic dissertation]. Espoo, Finland: Helsinki University of Technology Faculty of Electronics, Communications and Automation Department of Signal Processing and Acoustics; 2008.
17. Miller RL. Nature of the vocal cord wave. *J Acoust Soc Am*. 1959;31:667–677.
18. Fant G. *Speech Sounds and Features*. Cambridge: MIT Press; 1973.
19. Rothenberg M. A new inverse-filtering technique for deriving the glottal air flow waveform during voicing. Lomakkeen yläreuna. *J Acoust Soc Am*. 1973;53:1632–1645.
20. Alku P. Glottal wave analysis with pitch synchronous iterative adaptive inverse filtering. *Speech Commun*. 1992;11:109–118.
21. Airas M. TKK Aparat: an environment for voice inverse filtering and parameterization. *Logopen Phoniatr Vocol*. 2008;33:49–64.
22. Alku P, Vilkman E. Amplitude domain quotient of the glottal volume velocity waveform estimated by inverse filtering. *Speech Commun*. 1996;18:131–138.
23. Alku P, Bäckström T, Vilkman E. Normalized amplitude quotient for parametrization of the glottal flow. *J Acoust Soc Am*. 2002;112:701–710.
24. Pulakka H Analysis of human voice production using inverse filtering, high-speed imaging, and electroglottography [master's thesis]. Helsinki, Finland: Helsinki University of Technology. 2005:37–38.
25. Klatt DH, Klatt LC. Analysis, synthesis, and perception of voice quality variations among female and male talkers. *J Acoust Soc Am*. 1990;87:820–857.
26. Holmberg EB, Hillman RE, Perkell JS. Glottal airflow and transglottal air pressure measurements for male and female speakers in soft, normal, and loud voice. *J Acoust Soc Am*. 1988;84:511–1787.
27. Childers DG, Lee CK. Vocal quality factors: analysis, synthesis, and perception. *J Acoust Soc Am*. 1991;90:2394–2410.
28. Gobl C, Ní Chasaide A. Amplitude-based source parameters for measuring voice quality. VOQUAL'03. 2003. Available at: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.385.540&rep=rep1&type=pdf. Accessed July 21, 2015.
29. Yamaguchi H, Shrivastav R, Andrews ML, et al. A comparison of voice quality ratings made by Japanese and American listeners using the GRBAS scale. *Folia Phoniatr Logop*. 2003;55:147–157.
30. Ghio A, Weisz F, Baracca G, et al. Is the perception of voice quality language-dependant? A comparison of French and Italian listeners and dysphonic speakers. INTERSPEECH. 2011. Available at: http://despho-apady.univ-avignon.fr/documents/GHIO-2011-Interspeech.pdf. Accessed August 6, 2015.
31. Yiu EML, Murdoch B, Hird K, et al. Cultural and language differences in voice quality perception: a preliminary investigation using synthesized signals. *Folia Phoniatr Logop*. 2008;60:107–119.
32. Laukkanen AM, Ilomäki I, Leppänen K, et al. Acoustic measures and self-reports of vocal fatigue by female teachers. *J Voice*. 2008;22:283–289.
33. Rantala L. Ääni työssä. Naisopettajien äänenkäyttö ja äänen kuormittuminen (Voice at work. Female teachers' use and loading of voice). In Finnish, abstract in English [academic dissertation] Oulu, Finland: Department of Finnish, Saami and Logopedics and Otorhinolaryngology/Phoniatrics. University of Oulu; 2000.
34. Pegoraro Krook MI. Speaking fundamental frequency characteristics of normal Swedish subjects obtained by glottal frequency analysis. *Folia Phoniatr (Basel)*. 1988;40:82–90.
35. Laver J. Phonetic evaluation of voice quality. In: Kent RD, Ball MJ, eds. *Voice Quality Measurement*. San Diego, CA: Singular Publishing Group; 2000:37–48.