An electrolaryngographic study of dysphonic Portuguese speakers

by

Isabel Cristina Ramos Peixoto Guimarães

A thesis submitted for the degree of Doctor of Philosophy in the University of London

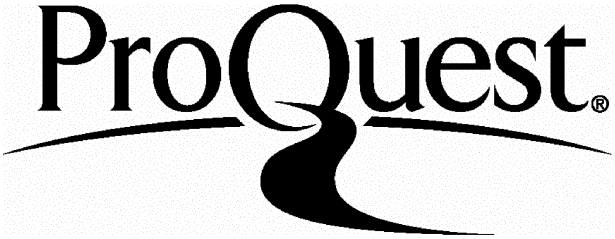
2002

ProQuest Number: U642749

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted. Also, if material had to be removed,
a note will indicate the deletion.



ProQuest U642749

Published by ProQuest LLC(2015). Copyright of the Dissertation is held by the Author.

All rights reserved.
This work is protected against unauthorized copying under Title 17, United States Code.
Microform Edition © ProQuest LLC.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Abstract

Fifty-two dysphonic Portuguese adult speakers were studied in a prospective exploratory study with fifty-seven healthy volunteers for control purposes. A standard complex evaluation protocol including: larynx examination; perceptual voice evaluation; subject's history questionnaire; self-rated level of stress and psychosocial voice impact; and electrolaryngographic measurement were consistently carried out.

The specific aims of this study were (a) normative, to establish tentative normative data for Portuguese speakers; (b) comparative, to investigate whether statistically significant differences exist between two groups with different voice conditions, on performance of different speech tasks and self-rating voice impact; and (c) predictive, to test for the physical variables that best predict aspects of the voice data (fundamental frequency (F0) and contact closed quotient derived measures) and the subjects' self-rated psychosocial voice impact data.

This study concentrates on the electrolaryngographic measurement and subjects' self-ratings psychosocial voice impact. From sustained vowel phonation, reading and conversation several fundamental frequency measures (mean, standard deviation, range and perturbation) and contact derived measures (closed contact quotient mean and standard deviation) were extracted. Also from the psychosocial voice questionnaire measures such as emotional, functional and physical impact were derived. Inferential and regression analyses were applied to these measures.

The normative data proved to be a valuable reference and within the limitations of the study, the following conclusions may be drawn: (1) ANOVA data analysis showed that dysphonia has a significant effect (a) on speech for the F0 standard deviation and for the percentage of F0 irregularity; (b) there is an interaction with age for the 90% F0 range; (c) dysphonia has a main effect on the psychosocial voice impact; (2) when a multivariate analysis was used, evidence suggest that the models (combining predictor variables) contribute differentially between high prediction of the mean F0 variance and moderate to low prediction of the other voice measures.

Acknowledgements

I must first thank the Calouste Gulbenkian Foundation for financial support through a scholarship without which the research described in this thesis would not have been possible.

I would like to express my appreciation to the Department of Phonetics and Linguistics at the University College London, the ENT unit at the S. José Hospital (Lisbon) and the Laryngograph Trust for their support in this project. The Department of Phonetics and Linguistics provided equipment facilities that made possible data gathering in Portugal. The ENT unit (S. José Hospital) through Dr. Cabral Beirão and his staff co-operated in the recording facilities and also acted as subjects. The Laryngograph Trust provided an equipment grant that made data gathering possible during the main study.

I gratefully acknowledge Dr. Evelyn Abberton who supplied motivation and the encouragement and guidance throughout the study.

I am very grateful to Dr. Ezequiel Barros and Dr. Ilídio Gama for their invaluable help in the ENT examinations, for giving me access to their patients, and for spending their time 'out of hours' performing stroboscopy on 'healthy' volunteers.

I also thank Prof. Adrian Fourcin (UCL) for his endless patience and helpful advice on the electrolaryngographic analyses, and Dr. Nuno Santos (PSE-Produtos e Serviços de Estatística, Lda.) for his support throughout the long process of statistical analyses and computing.

Finally, I must stress the precious help, encouragement and unconditional emotional support of my family throughout the 'special' years of this study.

Table of Contents

ABSTRACT **2**
ACKNOWLEDGEMENTS **3**
TABLE OF CONTENTS **4**
LIST OF FIGURES **7**
LIST OF TABLES **8**
CHAPTER I - INTRODUCTION **10**
1.1 – GENERAL AIMS AND PLAN OF THE THESIS 13
CHAPTER II – THE LARYNX **16**
2.1 - INTRODUCTION 16
2.2 - THE SOURCE-FILTER THEORY 16
2.3 - FRAMEWORK OF THE LARYNX 17
2.4 - MUSCULATURE 19
2.5 - LARYNGEAL CAVITY 22
2.6 - HISTOLOGY AND BIOMECHANICS OF THE VOCAL FOLD 24
2.7 - PHONATION 26
2.8 – AGE AND GENDER VARIATIONS IN 'NORMAL' LARYNX 29
CHAPTER III - DYSPHONIA **33**
3.1 - INTRODUCTION 33
3.2 - VOICE QUALITY CONTINUUM BETWEEN 'NORMAL' AND 'DYSPHONIA' 33
3.3 - CLASSIFICATION OF DYSPHONIA 39
3.3.1 – The functional versus organic classification *39*
3.3.1.1 – Minor structural and physiological alterations of the larynx 40
3.3.1.2 - Discrete mass lesions of the larynx 42
3.3.1.3 - Distributed tissue changes 44
3.3.1.4 - Neurological disorders 45
3.3.1.5 - Non-organic or muscular tension disorders 46
3.3.2 – Classification based on perceptual/acoustic phenomena *48*
3.3.2.1 - Types of voice quality 48
3.3.2.2 - Pathological voice types 50
3.4 - INCIDENCE AND PREVALENCE OF VOICE PROBLEMS 51
3.5 - VOCAL FUNDAMENTAL FREQUENCY 54
3.5.1 - Sampling criterion for voice measurement *55*
3.5.2 - Fundamental frequency according to age and gender *57*
3.5.3 - Speech material influences on fundamental frequency *60*
3.5.4 - Cultural influences on fundamental frequency *67*
3.5.5 - Variation in F0 during the day *68*
3.5.6 - Occupational risk factors: effects on fundamental frequency *69*
3.5.7 – Effects of cigarette smoking on fundamental frequency *70*
3.5.8 – Alcohol effect on fundamental frequency *72*
3.5.9 – Effect of the menstrual cycle on fundamental frequency *73*
3.5.10 - Fundamental frequency measurements for dysphonic populations *74*
3.6 - FUNDAMENTAL FREQUENCY PERTURBATION (JITTER) 78
3.6.1 - Jitter formulas *79*
3.6.2 – Jitter: instrumental and procedural influences *80*
3.6.3 – Jitter variability factors *82*

3.6.4 - Data on jitter *87*

CHAPTER IV -VOICE QUALITY EVALUATION **91**

4.1 - INTRODUCTION 91
4.2 - PATIENT'S MEDICAL HISTORY 91
4.2.1 - Case history questionnaire *91*
4.2.2 - Self-assessment of life events *92*
4.3 - PERCEPTUAL EVALUATION 95
4.4. - PHYSICAL EXAMINATION AND MEASUREMENT 100
4.4.1 - Techniques for laryngeal examination *100*
4.4.2 - Stroboscopy *101*
4.4.3 – Electrolaryngography (ELG) *105*
4.4.3.1 - Lx waveform interpretation and voice quality types 106
4.4.3.2 - Deviations from the Ideal Lx waveshape 109
4.4.3.3 - Voice source parameter measurements derived from the Lx waveform 111
4.4.3.4 - Presentation of data at the micro level using Speech Studio 113
4.4.3.5 - Presentation of data at the macro level using Speech Studio 113
4.4.3.5 – Data for vocal fold contact quotient 120
4.4.3.6 - Advantages of ELG 121
4.4.3.7 - Limitations and problems of ELG 122
4.5 - VOICE-RELATED QUALITY OF LIFE ASSESSMENT 125

CHAPTER V - METHODOLOGY **134**

5.1 - SUBJECTS 136
5.2 - MATERIALS 143
5.3 - EQUIPMENT 148
5.4 - EXPERIMENTAL ENVIRONMENT 150
5.5 - EXPERIMENTAL PROCEDURE 150
5.6 - DATA ANALYSES 153
5.6.1 - Screening data analyses *153*
5.6.2 – Laryngographic (Lx) data analyses *156*
5.6.3 - Subjects' self-reported data analyses *158*
5.7 - STATISTICAL ANALYSES 159

CHAPTER VI – RESULTS **162**

6.1 – INTRODUCTION 162
6.2 – VOICE CHARACTERISTICS 162
6.2.1 – Mean Fundamental frequency (F0) 162
6.2.2 – F0 standard deviation 169
6.2.3 - F0 range 177
6.2.4 – Percentage of jitter 182
6.2.5 - Percentage of Irregularity 186
6.2.6 – Mean contact closed quotient (QFx) 192
6.2.7 –QFx standard deviation 198
6.3 – SUBJECTS' SELF-RATED PSYCHOSOCIAL VOICE IMPACT 203

CHAPTER VII - DISCUSSION **213**

7.1 – INTRODUCTION 213
7.2 – VOICE CHARACTERISTICS 213
7.2.1- Mean fundamental frequency (F0) *213*
7.2.2- F0 standard deviation *217*
7.2.3- F0 range *220*

7.2.4 – Percentage of jitter *222*
7.2.5- Percentage of irregularity *224*
7.2.6-Mean contact closed quotient (QFx) *226*
7.2.7- QFx standard deviation *228*
7.3 – PSYCHOSOCIAL VOICE IMPACT 229
7.4 – STRENGTHS AND LIMITATIONS ENCOUNTERED IN THE STUDY 231
7.5 – SUGGESTIONS FOR FUTURE RESEARCH 236
7.6 - CONCLUSIONS 240

CHAPTER VIII - REFERENCES 242

CHAPTER IX – APPENDICES 268

APPENDIX A – PILOT STUDY 269
APPENDIX B – CONSENT FORM 294
APPENDIX C – ENT RATING FORM 297
APPENDIX D – QUESTIONNAIRE 300
APPENDIX E – VOICE HANDICAP INDEX 306
APPENDIX F – SOCIAL READJUSTMENT RATING SCALE 311
APPENDIX G – THE STORY OF ARTHUR THE RAT 314
APPENDIX H – FREQUENCE DISTRIBUTION OF VHI RATINGS 317

BIOGRAPHICAL SKETCH 320

List of Figures

FIGURE II.1 ANTERIOR AND POSTERIOR VIEW OF THE LARYNX 17
FIGURE II.2 – VIEW OF THE ARYTENOIDS 18
FIGURE II.3 - THE EXTRINSIC MUSCLES OF THE LARYNX 19
FIGURE II.4 - INTRINSIC MUSCLES OF THE LARYNX 20
FIGURE II.5 - LARYNGEAL CAVITY 23
FIGURE II.6 - NON-SCALED STRUCTURE OF THE VOCAL FOLD 24
FIGURE II.7 - SCHEMATIC PRESENTATION OF THE VOCAL FOLD VIBRATORY PATTERN 28
FIGURE III.1 - THE GENERAL TRENDS OF F0 ACROSS THE LIFE-SPAN 58
FIGURE III.2 – EFFECTS OF SAMPLING RATES ON THE JITTER MEASUREMENTS 81
FIGURE IV.1 – GELFER & BULTEMEYER'S (1990) RATING SYSTEM 103
FIGURE IV.2 – SÖDERSTEN & LINDESTAD'S (1990) RATING PROTOCOL 104
FIGURE IV.3 - ELECTRODES APPLIED TO THE THYROID LAMINAE 106
FIGURE IV.4 - IDEALISED MODAL LX WAVEFORM FOR AN ADULT 107
FIGURE IV.5 - LARYNGOGRAPHIC (LX) WAVEFORMS 108
FIGURE IV.6 - MEASUREMENTS DERIVED FROM THE ELG WAVEFORM 111
FIGURE IV.7 – VOCAL PROFILE ANALYSIS REPORT FOR A 'NORMAL' FEMALE SPEAKER 113
FIGURE IV.8 - DFX AND CFX HISTOGRAMS IN THE SPEECH STUDIO ANALYSES FOR A DYSPHONIC FEMALE SPEAKER 115
FIGURE IV.9 - QXFX SCATTERGRAM IN THE SPEECH STUDIO ANALYSES FOR A DYSPHONIC FEMALE SPEAKER 119
FIGURE V.1 – AGE DISTRIBUTION ACCORDING TO GENDER AND GROUPS. 138
FIGURE V.2 – SRRS MEANS ACCORDING TO GROUPS 141
FIGURE V.3 - ENT LARYNX DIAGNOSIS 141
FIGURE V.4 - SLT VOICE DIAGNOSIS 142
FIGURE V.5 - DIAGRAM OF EQUIPMENT CIRCUITRY FOR DATA COLLECTION 149
FIGURE V.6 – F0 CENTRAL TENDENCY MEASURES 158
FIGURE VI.1 – MEAN F0 FOR THE CONTROL SPEAKERS ACCORDING TO GENDER 163
FIGURE VI.2 – MEAN F0 FOR THE DYSPHONIC SPEAKERS ACCORDING TO GENDER 164
FIGURE VI.3 – MEAN F0 STANDARD DEVIATION FOR THE CONTROL SPEAKERS 170
FIGURE VI.4 – MEAN STANDARD DEVIATION F0 FOR THE DYSPHONIC SPEAKERS 171
FIGURE VI.5 – 90% MEANS OF F0 RANGE FOR THE CONTROL SPEAKERS 177
FIGURE VI.6 – 90% MEANS OF F0 RANGE FOR THE DYSPHONIC SPEAKERS 178
FIGURE VI.7 – % JITTER FOR THE CONTROL SPEAKERS 182
FIGURE VI.8 – % JITTER FOR THE DYSPHONIC SPEAKERS 183
FIGURE VI.9 – % OF IRREGULARITY FOR THE CONTROL SPEAKERS 187
FIGURE VI.10 – % OF IRREGULARITY FOR THE DYSPHONIC SPEAKERS 188
FIGURE VI.11 – % OF CONTACT QUOTIENT FOR THE CONTROL SPEAKERS 192
FIGURE VI.12 – % OF CONTACT QUOTIENT FOR THE DYSPHONIC SPEAKERS 193
FIGURE VI.13 – VHI BOXPLOTS FOR THE CONTROLS AND DYSPHONICS 203
FIGURE VII.1 – DFX1 DISTRIBUTIONS FOR THE SPEAKER 54 238
FIGURE VII.2 – DFX2 DISTRIBUTIONS FOR THE SPEAKER 54 238
FIGURE VII.3 – CFX DISTRIBUTIONS FOR THE SPEAKER 54 239
FIGURE VII.4 – QXFX DISTRIBUTIONS FOR THE SPEAKER 54 239

List of Tables

TABLE II.1- VOCAL FOLD LENGTH 30
TABLE III.1 –MEDICATION AND DRUGS' EFFECTS ON THE LARYNX AND VOICE 36
TABLE III.2 - KOUFMAN & ISAACSON'S CLASSIFICATION OF PROFESSIONAL VOICE USE. 37
TABLE III.3 - VILKMAN'S VOICE CLASSIFICATION 38
TABLE III.4 – SOME MINOR STRUCTURAL AND PHYSIOLOGICAL ALTERATIONS OF THE LARYNX 41
TABLE III.5 - DISCRETE MASS LESIONS OF THE LARYNX 43
TABLE III.6 – SOME COMMON DISTRIBUTED TISSUE CHANGES 44
TABLE III.7 – RECURRENT LARYNGEAL NERVE PARALYSIS 45
TABLE III.8 – SOME COMMON NON-ORGANIC OR MUSCULAR TENSION DISORDERS 47
TABLE III.9 - INTRINSIC MEAN F0 FOR ADULT FEMALES 61
TABLE III.10 - INTRINSIC F0 FOR ADULT MALES 61
TABLE III.11- MEAN F0 FOR ADULT FEMALES DURING READING 64
TABLE III.12 - MEAN F0 FOR ADULT MALES DURING READING 65
TABLE III.13 - MEAN F0 FOR ADULT FEMALES DURING CONVERSATION 66
TABLE III.14 - MEAN F0 FOR ADULT MALES DURING CONVERSATION 66
TABLE III.15 – MURRY (1978) ACOUSTIC DATA 76
TABLE III.16 – PEPPARD, BLESS & MILENKOVIC (1988) ACOUSTIC DATA 76
TABLE III.17 – MEAN AND STANDARD DEVIATION JITTER FOR ADULT FEMALES 85
TABLE III.18 – MEAN AND STANDARD DEVIATION JITTER FOR ADULT MALES 86
TABLE III.19 – OMORI ET AL. (1997) DATA 90
TABLE IV.1 – VERDOLINI ET AL.'S (1998) DATA 121
TABLE IV.2 - VHI SCORES (JACOBSON ET AL., 1997) 129
TABLE IV.3 - VHI SCORES (STEWART, CHEN & STACH, 1998) 130
TABLE IV.4 - MEAN VHI SCORES (ROSEN & MURRY, 2000) 130
TABLE IV.5 – WHO (1980; 1997; 1999) CONCEPTUAL FRAMEWORKS 132
TABLE V.1 - SUBJECTS' DEMOGRAPHIC INFORMATION 137
TABLE V.2 - DYSPHONICS AND CONTROLS' SOCIAL AND VOICE HABITS HISTORY 138
TABLE V.3 - DYSPHONIC AND CONTROL GROUPS' HEALTH HISTORY 139
TABLE V.4 – DYSPHONIA HISTORY 142
TABLE V.5 – PORTUGUESE VOWELS USED IN THE STUDY 146
TABLE V.6 – QUESTIONNAIRE CATEGORICAL DATA ANALYSES 154
TABLE V.7 – SPEECH SAMPLES USED FOR THE LX ANALYSES 157
TABLE V.8 – PERIODS/SAMPLES USED FOR THE LX ANALYSES 159
TABLE VI.1 – MEANS OF F0 FOR ALL THE SPEECH TASKS ACCORDING TO GROUPS 162
TABLE VI.2 – BETWEEN SUBJECTS' EFFECTS FOR THE MEAN F0 165
TABLE VI.3 – WITHIN SUBJECTS' EFFECTS FOR THE MEAN F0 165
TABLE VI.4 – RELATED T-TESTS FOR THE MEAN F0 SPEECH TASKS PAIRS ACCORDING TO GENDER 166
TABLE VI.5 – MULTIPLE REGRESSION ANALYSIS FOR VARIANCE IN THE MEAN F0 167
TABLE VI.6 – MEANS OF F0 STANDARD DEVIATION ACCORDING TO THE GROUPS AND AGE 170
TABLE VI.7 – BETWEEN SUBJECTS' EFFECTS FOR THE F0 STANDARD DEVIATION 172
TABLE VI.8 – RELATED T-TESTS FOR THE MEAN F0 SPEECH TASKS PAIRS ACCORDING TO AGE 172
TABLE VI.9 – WITHIN SUBJECTS' EFFECTS FOR THE F0 STANDARD DEVIATION 173
TABLE VI.10 – RELATED T-TESTS FOR THE F0 STANDARD DEVIATION SPEECH TASKS ACCORDING TO GROUPS 173

TABLE VI.11 - MULTIPLE REGRESSION ANALYSIS IN THE F0 STANDARD DEVIATION 174
TABLE VI.12 – 90% MEANS OF F0 RANGE (IN OCTAVES) ACCORDING TO GROUPS 177
TABLE VI.13 – BETWEEN SUBJECTS' EFFECTS FOR 90% F0 RANGE 179
TABLE VI.14 – WITHIN SUBJECTS' EFFECTS FOR 90% MEAN F0 RANGE 180
TABLE VI.15 - MULTIPLE REGRESSION ANALYSIS IN F0 RANGE 180
TABLE VI.16 – MEANS OF JITTER PERCENTAGE FOR ALL THE SUSTAINED VOWELS 182
TABLE VI.17 – BETWEEN SUBJECTS' EFFECTS FOR JITTER 184
TABLE VI.18 – WITHIN SUBJECTS' EFFECTS FOR JITTER 184
TABLE VI.19 - MULTIPLE REGRESSION ANALYSIS FOR JITTER VARIANCE 185
TABLE VI.20 – MEANS AND STANDARD DEVIATIONS OF PERCENTAGE OF IRREGULARITY 187
TABLE VI.21 – BETWEEN SUBJECTS' EFFECTS FOR PERCENTAGE OF IRREGULARITY 188
TABLE VI.22 – WITHIN SUBJECTS' EFFECTS FOR IRREGULARITY 189
TABLE VI.23 – PERCENTAGE OF IRREGULARITY ACCORDING TO GENDER AND AGE 190
TABLE VI.24 - MULTIPLE REGRESSION ANALYSIS IN F0 IRREGULARITY 190
TABLE VI.25 – MEANS OF QFX (%) 192
TABLE VI.26 – BETWEEN SUBJECTS' EFFECTS FOR QFX 194
TABLE VI.27 – WITHIN SUBJECTS' EFFECTS FOR QFX 194
TABLE VI.28 – RELATED T-TEST FOR THE QFX SPEECH TASKS PAIRS 195
TABLE VI.29 - MULTIPLE REGRESSION ANALYSIS IN THE MEAN QFX 196
TABLE VI.30 – MEANS OF QFX STANDARD DEVIATION FOR ALL THE SPEECH TASKS 198
TABLE VI.31 – BETWEEN SUBJECTS' EFFECTS FOR QFX STANDARD DEVIATION 199
TABLE VI.32 – WITHIN SUBJECTS' EFFECTS FOR THE MEAN QFX STANDARD DEVIATION 199
TABLE VI.33 – RELATED T-TEST FOR THE QFX STANDARD DEVIATION SPEECH TASKS PAIRS 200
TABLE VI.34 - MULTIPLE REGRESSION ANALYSIS IN THE QFX STANDARD DEVIATION 201

Chapter I - Introduction

Over the centuries the human speaking voice has been a subject of interest throughout society especially within the science and health communities. This interest applies to all types of vocal skill, to speakers of all ages, to 'normal' and 'pathological' voices, and to both professional and non-professional voice use. It has inspired a great deal of individual research concerning voice definition, assessment and therapy. Nevertheless, the study of voice is still fragmented in a huge body of concepts related to the different areas (e.g. medicine, linguistics, engineering, speech and language therapy) and consequently gaps and ambiguities in the knowledge of 'normal' and 'pathological' voice quality evaluation still exist limiting the development of theoretical frameworks (Ludlow, 1981; Bless & Baken, 1992; Bless et al., 1992; Gerrat & Kreiman, 2000).

Although epidemiological studies of the dysphonia incidence in the general population are almost absent, there is evidence that a significant proportion of speakers have some kind of voice problem across their life and look for help (Herrington-Hall et al., 1988; Arnoux-Sindt et al., 1994, Enderby & Emerson, 1995; Fritzell, 1996; Guimarães, 1997; Newman & Kersner, 1998; Simberg et al., 2000; Verdolini & Ramig, 2001). Therefore, voice evaluation and diagnosis have turned out to be important areas for study and research.

The evaluation of the voice's health is an ongoing process whereby clinicians routinely gather all the vital information needed to determine the diagnosis. Historically, the most prominent methods of voice evaluation, in the health field, are almost all based on 'direct observations'; data collected from the patient's medical history, the ENT (Ear Nose and Throat) surgeon's laryngeal observations and SLT's (Speech and Language Therapist) perceptual evaluation (Koike, Takahashi & Calcaterra, 1977; Piccirillo et al., 1998). Although imaging of the vocal fold enables ENT surgeons and SLTs to have essential views of the larynx inlet (anatomy and pathology) and its phonatory action, perceptual methods for evaluating a person's voice are of primary importance because patients usually seek for help based on their own perception of a voice quality deviation (Gerrat & Kreiman, 2000), are favoured for being quick, convenient and cost effective compared to other types of measures and may be the only available mean for clinicians to judge voices (Fex, 1992); however, given their low reliability (Cordes, 1994) their sole use is likely to be inadequate.

It would seem imperative, therefore, that the perceptual judgments of vocal pitch (correlate of fundamental frequency) and quality be associated with measurements that can support a discriminative evaluation of voice range and variability of performance with a satisfactory degree of reliability (Laver, Hill & Beck, 1992; Behrman & Orlikoff, 1997; Baken & Orlikoff, 2000).

More recently, advances in technology and its growing availability have not only increased the stock of knowledge available in several voice areas that otherwise were impossible to evaluate without instrumentation, but they have also brought a new emphasis on standardisation and quantification (Koike, Takahashi & Calcaterra, 1977; Fitch, 1990). The rationale for standardisation is outlined by several authors (Titze, 1994a; Laver, Hiller & Beck, 1992; Baken & Orlikoff, 2000) and societies (e.g. American National Standards Organizations (ANSI), and the International Standards Organization (ISO)) as a process that unifies, simplifies and establishes bases of knowledge that ultimately contribute to diagnosis and to improving health care.

Much effort has been spent on determining measures for 'normal' and 'disordered' voice quality such as fundamental frequency mean, range, variability and perturbation (Horii, 1979; Mysak, 1959; Stoicheff, 1981; Krook, 1988; Hollien, Hollien & De Jong, 1997) as well as on establishing a correlation with the most traditional observational methods of evaluation (Eskenazi, Childers & Hick, 1990; Wolfe, Cornell & Palmer, 1991; Wolfe, Fitch & Cornell, 1995; Kroom, 1995). However, though these forms of evaluation have resulted in an important step forward in voice science, they have also brought several potential problems and limitations (e.g. equipment standardisation) that have not yet been solved (Koike, Takahashi & Calcaterra, 1977; Stone & Rainey, 1991).

Several researchers have pointed out the urgent need of developing different types of databases, which should be recorded under standard conditions (e.g. acoustic environment, speech samples, speakers involved, equipment, procedures, analysis systems) for the development of acoustic indices which may be useful to facilitate pathology detection, assessment of severity or differential diagnosis of dysphonia (Koike, Takahashi & Calcaterra, 1977; Laver, Hiller & Beck, 1992; Bielamowicz et al.,

1996; Titze, 1994a). In the absence of standard assessment measures diagnosis and treatment recommendations are severely based on clinical experience and beliefs.

Unfortunately, to date, information available in the literature regarding details of Portuguese voice is modest. For example, the Portuguese ENT journal presents, regularly, articles where the emphasis is on larynx examination, diagnosis and surgery and since its beginning only two articles (Ferreira et al., 1996; Gomes et al., 2000) were found about the prevalence of psychogenic aphonia and dysphonia in patients seeking help. Also, in the journal of the Portuguese Association for Speech and Language Therapists (APTF) articles about voice are scarce and usually related to therapy and case studies (Ferreira & Cruz, 1986; Cruz, 1986). In other journals and books (Cruz & Barros, 1989; Cruz, 1991, 2000; Guimarães, 1992, 1995, 1997; Guimarães & Cruz, 1995, 1997a; Correia, Quintel & Pimentel, 1995; Carvalho et al., 1995; Cruz & Guimarães, 1997) voice is described from a theoretical point of view (evaluation and prevention) or once again according to individual perspectives and experiences. Consequently, contemporary knowledge of voice in Portuguese speakers is based primarily on theories derived from the observation of the larynx and perceptual analyses of the speech of a small number of patients with dysphonia although more recently, unpublished works on voice measurement for a group of teachers (Guimarães & Cruz, 1997b) and smoking women (Rodrigues, 1998) have started to appear.

Due to the fact that by definition voice quality is multidimensional most clinicians and researchers acknowledge that, optimally, voice evaluation involves the use and analyses of data from a multidimensional perspective (rather than a single measure) focusing on both qualitative and quantitative aspects to obtain a reasonable holistic perspective of the subject's voice characteristics and variance. Nevertheless, investigations of voice quality tend to use unidimensional perspectives of voice evaluation (e.g. perceptual or acoustic and sometimes the correlation between those) (Hirano & Bless, 1993; Verdolini, 1994; Piccirillo et al., 1998; Orlikoff et al., 1999; Andrews, 1999; Hill et al., 2000; Gerratt & Kreiman, 2000).

Moreover, independently of each tool, and of the validity and reliability of methodologies used, most of those approaches and tools used within the clinical setting rely heavily on 'expert' knowledge and opinion (estimates of impairment and functioning related to diagnosis) and usually do not reflect the 'patient's opinion'.

Although measurement of quality of life has not been part of the traditional voice evaluation protocol within the framework of health measurement, it has been regarded as a useful and important supplement to the clinical decision-making process by providing valuable insights into the impact of the disease on the subject's quality of life from his/her own perspective (Smith et al., 1996; Hogikyan & Sethuraman, 1999; Verdolini & Ramig, 2001). Although the combination of several measures in the voice evaluation study exists (Peppard, Bless & Milenkovic, 1988; Biever & Bless, 1989; Kreiman et al., 1992; Verdolini & Palmer, 1997; Murdoch et al., 1997) it rarely includes the subjects'self-assessment (Carlson, 1995a; Bough et al., 1996; Verdonck-de Leeuw, 1998; Wuyts et al., 2000).

Wuyts et al. (2000) report that within the clinical setting, a scientifically relevant model for voice assessment should consider: larynx examination, perceptual evaluation (e.g. GRBAS), patients' self-assessment (e.g. VHI) and objective measure of vocal quality (e.g. aerodynamic, acoustic measurement).

1.1 – General aims and plan of the thesis

Given the reasons expressed above it appeared crucial to work towards the development of multidimensional voice measurement that could give insight into several aspects of the Portuguese voice. Therefore, the main aim of this study was to investigate the voice characteristics of 52 dysphonic Portuguese speakers together with 57 Portuguese healthy volunteers as controls. Electrolaryngography (ELG) (Fourcin & Abberton, 1971; Colton, 1996; Abberton & Fourcin, 1997; Baken, 1992; Orlikoff, 1998; Baken & Orlikoff, 2000), a non-invasive instrumental technique, was used to obtain derived voice source characteristics (e.g. fundamental frequency and closed contact quotient) during different types of speech tasks (sustained vowels, oral reading and conversation). Electrolaryngography (ELG) is an electrical conductance method of estimating vocal fold contact during phonation, which is designed to allow the examiner to investigate laryngeal microfunction (cycle-to-cycle periodicity and contact). Although some caution must be used in interpreting ELG results, since the waveform obtained is unaffected by the acoustic resonances of the vocal tract, a number of authors acknowledge that the ELG signal provides a useful estimate of vocal fold contact during the glottal cycle, and gives some insight into the regulation, maintenance and quality of

phonation (Haji et al., 1986; Colton and Conture, 1990; Orlikoff, 1991; Baken, 1992; Garner & Howard, 1999).
The major questions to be answered were: (a) What are the quantifiable voice characteristics of Portuguese speakers? (b) are there any statistically significant differences between two groups of different voice conditions on performance of different speech tasks? (c) which variables contribute most importantly to predicting variance in the quantifiable voice measures?

In addition to tentative normative aspects of Portuguese voice measurement; another source of voice quality information (the subjects' self-rated voice psychosocial impact) was used to get insight in the impact of voice on the quality of life according to emotional, functional and physical aspects. The major questions to be answered were: (a) What is the dysphonic Portuguese speakers' self-rated voice psychosocial impact? (b) are there any statistically significant differences compared to a group of healthy Portuguese speakers? (c) which variables contribute most importantly to predicting variance in the quantifiable data?

The starting point was a pilot study during a period of three and a half months (to validate the research design, to practise the ELG application and get insight about Portuguese speakers) followed by the main study during a period of nine months. All subjects (dysphonics and controls) were submitted to an evaluation protocol (stroboscopy, case history questionnaire, perceptual evaluation and self-rating stress scale) in order to get insight into variables that may be related to voice quality as well as to safeguard the medical and psychological well- being of the subjects.

The present study developed along these general lines is one of the first studies for Portuguese, and it is hoped that it may contribute to the knowledge of Portuguese voice providing in depth data about dysphonia and its psychosocial impact. Consequently it provides the clinician with a more accurate tool for diagnosis and treatment of adult dysphonic speakers.

In the present thesis a brief description of the larynx anatomophysiology and the process of phonation is given in chapter II and also a discussion of some 'normal' age-gender variability in larynx structure and function. The next chapter (III) gives overviews of voice definition, aetiology, classification and incidence of voice problems,

followed by a critical review of fundamental frequency data. A general overview of voice quality evaluation methodology describing the most commonly used tools in the health field (data from case histories, perceptual evaluation, instrumental measurement and subject'self-assessment) is carried out in chapter IV. The chapter also includes a detailed description of the electrolaryngographic technique. Chapter V deals with the methodology of this research. Chapter VI comprises a detailed presentation of the Lx and Voice Handicap Index (VHI) measurement results, including a comparison of the groups according to the main effects in the study, and also the prediction of variance results. In chapter VII a general discussion focuses on the voice characteristics studied and psychosocial voice impact is presented in relationship to the statistical significance and the influence of the factors investigated and compared to the studies reported in the literature review. Also, strengths and limitations encountered in the study are pointed out, conclusions of this thesis and suggestions for future research are also presented.

Chapter II – The Larynx

2.1 - Introduction

In this chapter aspects that provide fundamental information to the nature of the voice source (the larynx and its surrounding structures) will be presented here regardless of the relative importance of each system involved in the production of the human voice.

The larynx system is described from a functional perspective, deliberately avoiding too much technical detail, with the concepts and ideas reproduced from the insights of several available textbooks such as Kaplan (1971); Prater & Swift (1984); Tucker (1987); Le Huche et Alalli, (1991); Jackson-Menaldi (1992); Morrison et al (1994); Titze (1994); Clark & Yallop (1995); Dworkin & Meleca (1997), Sataloff (1997, 1998) and Stevens (1998).

2.2 - The source-filter theory

Speech is the product of complex activity that involves several systems (respiratory, phonatory, resonance and articulatory) including the neurophysiological system responsible for their coordination.

Essentially, speech is assumed to be the result of the interaction between: (a) the egressive (occasionally ingressive) pulmonic airstream; (b) the vocal folds vibration (the source system), for some sounds but not all, and; (c) the supraglottic structures representing the vocal tract (the filter system) (Cummings & Clements, 1995; Kent, 1977; Ladefoged, 1996; Stevens, 1998).
The application of the source-filter theory gave a major advance in the knowledge of speech production. Although the final output (the radiated pressure spectrum) has a harmonic pattern distribution resulting from the interaction of the vibratory pattern of the vocal folds (in some sounds) and the resonance characteristics of the vocal tract, each element may be analysed and modelled independently for the experimental study of their contribution to the final product.

2.3 - Framework of the larynx

The larynx is a skeletal and neuromuscular system that generates sound but its primary purpose is to allow airflow to and from the lungs, to protect the upper respiratory tract, to assist in pressure-valving functions (to allow activities such as weight-lifting, defecation, vomiting and coughing) and to aid in swallowing. It is located in the neck, in front of the oesophagus and in continuity with the trachea below and the pharynx above; blood vessels, nerves, glands and other structures surround it. The skeletal support of the larynx is provided by five principal cartilages, three median individual cartilages (cricoid, thyroid and epiglottis), two paired, lateral (arytenoids), two adjunct cartilages (corniculate and cuneiform) and a bone (hyoid) connected to each other by a series of articulations, membranes, ligaments and specialised joints as can be seen in Figure II.1.

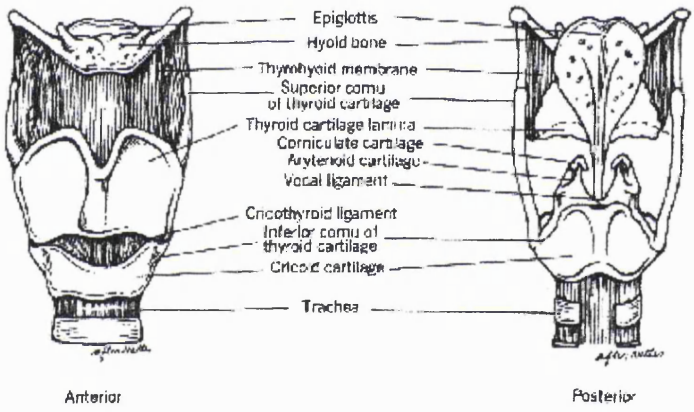


Figure II.1 Anterior and posterior view of the larynx - From: Sataloff (1997:112)

The *cricoid cartilage* lies in front of the sixth cervical vertebra and may be regarded as the base and support for the entire larynx. It is a complete ring between the trachea below and the thyroid cartilage above, as shown in Figure II.1

The *thyroid cartilage* is the largest of the laryngeal cartilages. It consists of two flat laminae with an angle in the medial ventral line of approximately 90 degrees in males and about 120 degrees in females. In addition, the thyroid prominence projects very far forward in males. The thyroid cartilage is located above the cricoid and articulates with its sides through the cricothyroid joints. The upper border of the thyroid

cartilage is connected by the thyrohyoid membrane and ligament to the hyoid bone, as shown in figure II.1. The superficial electrodes employed for electrolaryngographic voice measurements in this study are applied superficially at the level of the thyroid laminae.

The *arytenoid cartilages* are two small bilateral pyramid-like structures located on the top of the posterior surface of the cricoid cartilage via the cricoarytenoid joints, as illustrated in figure II.2. The base has an anterior extension, the vocal process, which forms the point of insertion for the vocal ligament; and a posterolateral extension, the muscular process. The arytenoids can glide in all four cardinal points direction; rotate slightly, but with restricted movement, with respect to the cricoid which controls the vocal folds position.

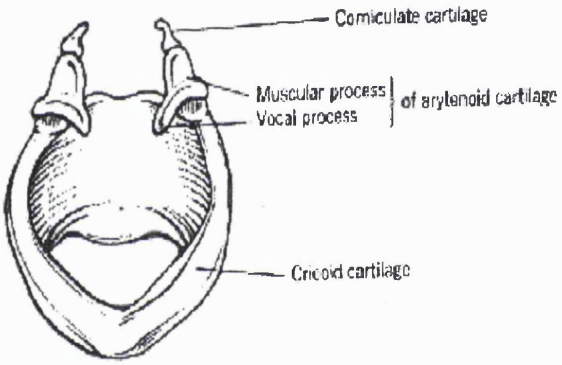


Figure II.2 – View of the arytenoids - From: Sataloff (1997:112)

The *epiglottis* (figure II.1) is a leaf-shaped cartilage attached to the medial surface of the thyroid cartilage, projecting upwards towards the tongue. Its surface is penetrated by multiple small holes that transmit tiny blood vessels and fibrous tissue attachments from its laryngeal surface to the pre-epiglottic space. The epiglottis has an indirect influence on voice quality by producing changes in the shape of the laryngeal cavity.

Although the *hyoid bone* (fig II.1) is primarily a support for the tongue, it may be considered part of the laryngeal framework because it is the point of attachment for muscles and ligaments essential to larynx function. It is a more or less U-shaped bone consisting of an anterior central body with the limbs of the U pointing posteriorly. It is located above the thyroid cartilage and inferiorly to the base of the tongue.

2.4 - Musculature

The muscles of the larynx may have one attachment within the larynx and another attachment to some non-laryngeal structure (called extrinsic muscles) or both their origins and insertions on the laryngeal cartilages (called intrinsic muscles). The extrinsic muscles' function is to support the larynx and change its spatial relationships, and on the other hand, the intrinsic muscles' function is vocal fold movement for phonation, assistance in respiration, containing air below the larynx (glottal effort closure reflex) or for preventing foreign particles entering the lungs.

Although the extrinsic muscles of the larynx are more related to swallowing they also have some indirect effect on phonation. Changing the position of the larynx in the neck by either raising it (suprahyoid muscles) or lowering (infrahyoid muscles) may alter the tension and the angle between the laryngeal cartilages, thereby influencing the support or activity balance.

The suprahyoid group of muscles (Figure II.3), including the digastric, stylohyoid, geniohyoid, mylohyoid and stylopharyngeus muscles, are all laryngeal elevators because, indirectly, they raise the larynx via its attachment to the hyoid bone and consequently may be associated with raising the voice pitch.

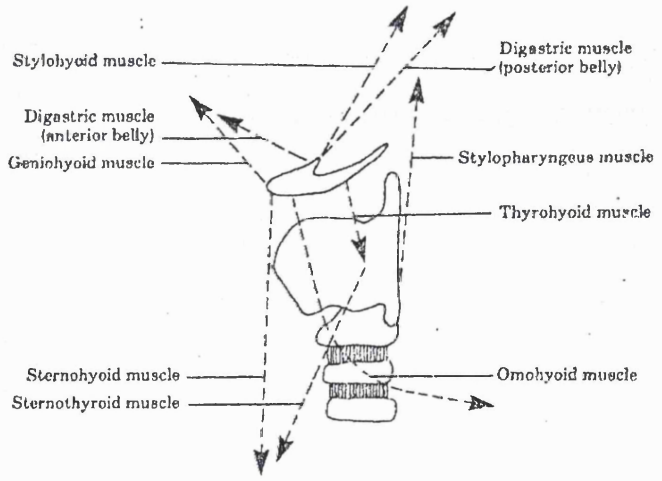


Figure II.3 - The extrinsic muscles of the larynx (supra and infrahyoid groups)
From: Prater and Swift (1984: 5)

The infrahyoid group (Figure II.3) includes the omohyoid, sternothyroid, thyrohyoid and sternohyoid muscles. They are capable of moving downward the entire laryngotracheal complex as much as one whole vertebral body, thereby lowering the pitch. They are also capable of providing a small degree of increased tension in the vocal folds because when they are very tense, they tend to pull the thyrotracheal complex somewhat forward in relation to the cricoid cartilage.

The depressor group of muscles displaces the larynx downward during the inspiratory phase of respiration, while the elevators lift it during most expiratory cycles.

For phonation activity the intrinsic muscles (figure II.4) may be considered according to their general effects upon the glottis: those which contribute to adduction and abduction of the vocal folds (posterior and lateral cricoarytenoids, transverse arytenoid and oblique arytenoids) and those which contribute to vocal fold tension and compression (thyroarytenoid and cricothyroid muscles).

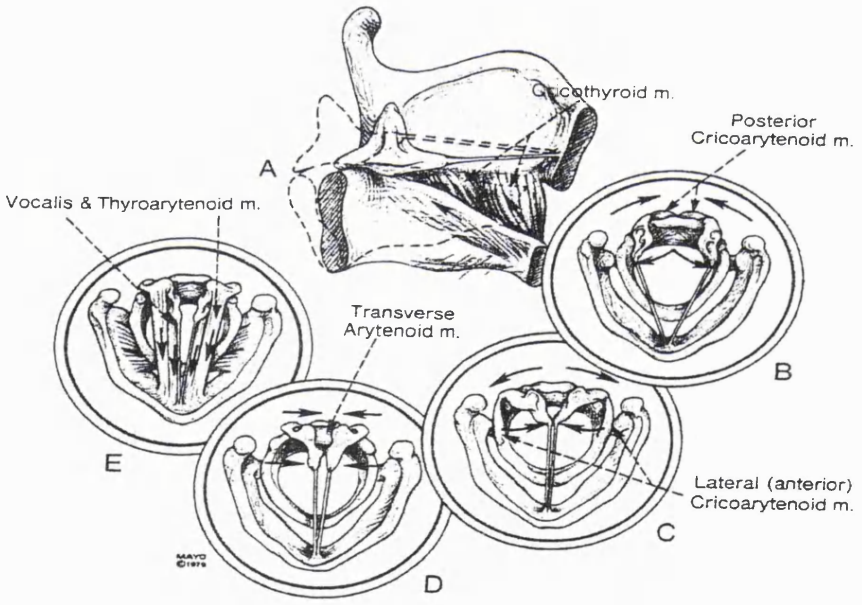


Figure II.4 - Intrinsic muscles of the larynx - From: Aronson (1990:24)

Under normal conditions of speaking and breathing the posterior cricoarytenoid muscles are the only laryngeal muscles responsible for vocal fold abduction, and they are therefore most active during the inspiratory phase of respiration and during the production of voiceless sounds. The posterior cricoarytenoid muscles (fig II.4B) run

from the posterior surface of the cricoid cartilage to the muscular process of the arytenoids. They are direct antagonists to the lateral cricoarytenoids. Although they should be regarded as a primary abductor of the vocal folds, they must also be recognized as a tensor of the vocal folds (elongates and thins the vocal fold by rocking the arytenoid posterolaterally).

The lateral cricoarytenoid muscles (Figure II.4C), transverse arytenoid muscle (Figure II.4D) and oblique arytenoid muscle function as adductors of the vocal folds.

The lateral cricoarytenoid muscles are attached between the upper border of the arch of the cricoid cartilage and the muscular process of the arytenoids. Their action is to adduct and lower the vocal folds complementing the function of the interarytenoideus muscle. Although this activity brings the membranous portions of the vocal folds together, the concomitant adduction of the bodies of the arytenoids provided by the interarytenoideus muscle is necessary to completely close the glottis. In addition to its activity as the major adductor of the vocal fold, the lateral cricoarytenoid muscle can also act as a tensor once the bodies of the arytenoids cartilages have been brought to the midline and the vocal fold edge becomes more angular or sharp.

The transverse arytenoid (also called interarytenoids) (Figure II.4D) is the only unpaired muscle, its fibres extends from the muscular process and lateral border of the arytenoids on each side, crossing the midline. Upon contraction it approximates the bodies of the arytenoids, thus closing the posterior part of the glottis when the vocal processes have already been brought into apposition by the lateral cricoarytenoid muscles. Its most important function, however, is the medial compression of the vocal folds, which consequently affects the stiffness of the membranous portion.

The oblique arytenoid muscles are paired and run from the muscular process of one arytenoid to the apex of the other arytenoid and upward along the aryepiglottic fold forming the aryepiglottic muscle. These muscles acting in concert with the interarytenoideus muscles contribute to vocal fold adduction. Combined action of these two muscles has a sphincter effect on the introitus of the larynx, closing it down during the act of swallowing and effort closure.

The thyroarytenoid muscles (Figure II.4E) consist of an upper portion (ventricular), a lower medial portion (vocalis) and a lateral portion (muscularis). They arise on the posterior surface of the thyroid cartilage, and their lateral fibres insert on the lateral edge and muscular process of each arytenoid. Contraction of these muscles draws the arytenoid cartilages forward and downward towards the thyroid cartilage resulting in shortening of the vocal ligaments and relaxation of the mucous membrane covering them. In fact, Sataloff (1998:10) notes that 'the cover and transition are effectively made slack, while the body is stiffened'. They determine the degree of tension and the bulk of the main body of the vocal folds.

The cricothyroid muscles (Figure II.4A) run between the anterior lateral part of the cricoid cartilage and the lower lateral part of the thyroid cartilage. Upon contraction they move the vocal folds into the paramedian position. Sataloff (1998:12) states that *'it also lowers, stretches, elongates, and thins the vocal fold, stiffening all layers and sharpening the vocal contour'*. Although this is not truly an intrinsic muscle of the larynx, its activity correlates well with pitch control by shortening the distance between the cricoid and thyroid cartilages, so increasing the length and tension of the vocal folds.

The recurrent laryngeal nerve innervates all intrinsic muscles (with the exception of the cricothyroid muscle that is innervated by the superior laryngeal nerve). Both recurrent and superior laryngeal nerves are branches of the vagus nerve (X^{th} cranial nerve).

2.5 - Laryngeal Cavity

The cavity of the larynx is a funnel-shaped tube composed of cartilages connected by ligaments and covered by mucous membrane. Superiorly, the cavity is an anterior diverticulum of the lowest part of the pharynx and continuous below with the cavity of the trachea. The interior may be divided from above downward into three spaces: the vestibule (or supraglottic space), the glottis and the infraglottic (or subglottic space), shown in Figure II.5.

The *vestibule or supraglottic space* as shown in Figure II.5 is the area between the free margin of the epiglottis and aryepiglottic folds above and the lower margin of the ventricular folds (usually referred to as the false vocal folds) below. The vestibular

cavity is wide above, but it narrows from side to side as it approaches the glottis. Colton and Casper (1996: 328) say that: *'...this cavity could potentially act as a resonator of the sound produced by the vibrating vocal folds, although it is usually considered to be a part of the entire vocal tract acoustically'.*

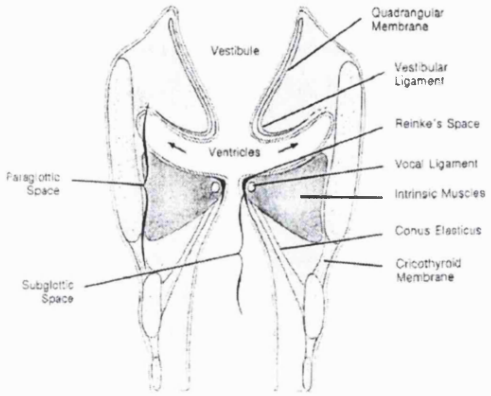


Figure II.5 - Laryngeal cavity - From Tucker (1987:8)

The middle division, the glottic region or ventricle, is the space between the ventricular folds above and the vocal folds below as illustrated in Figure II.5.

The ventricular folds (also called false vocal folds) do not participate in 'normal' phonation. They help to lubricate the true vocal folds during phonation protecting them from the effects of friction, and they also assist in pressure-valving functions by holding high pulmonary (subglottal) pressures.

The vocal folds consist of two bands that play a direct role in the voice source. They are roughly triangular in cross-section and they include the upper part of the conus elasticus, the vocalis muscle, and the mucous membrane lining the laryngeal airway. The vocal folds run from half way between the thyroid notch and inferior border of the thyroid cartilage in females and 1/3 the distance in males. Along their length, the posterior two fifths are cartilaginous (lying between the arytenoids). Some authors (Zemlin, 1988; Greene and Mathieson, 1989) refer to it as the *'respiratory glottis'* because it serves primarily as a conduit for respiration although it participates in swallowing and low frequency vibrations (e.g. creaky voice quality). The anterior three fifths are membranous (and extends anteriorly

from the tip of the vocal process to the anterior commissure tendon) and is also called the '*phonatory glottis*' because of its main function in phonation.
The horizontal space between the true vocal folds and the arytenoid cartilages is the laryngeal valve aperture known as the *glottis* (rima glottidis or glottic chink).

The lowest division of the larynx, referred to as the subglottic or *infraglottic space*, (Figure II.5) includes the area between the vocal folds above (rima glottidis) and the inferior border of the cricoid cartilage below.

2.6 - Histology and biomechanics of the vocal fold

Hirano (1981) has proposed that the adult vocal fold is a five-layered structure with progressive density and viscosity of great significance in the knowledge of vocal fold vibratory dynamics, (Fujimura, 1981; Hirano & Bless, 1993).

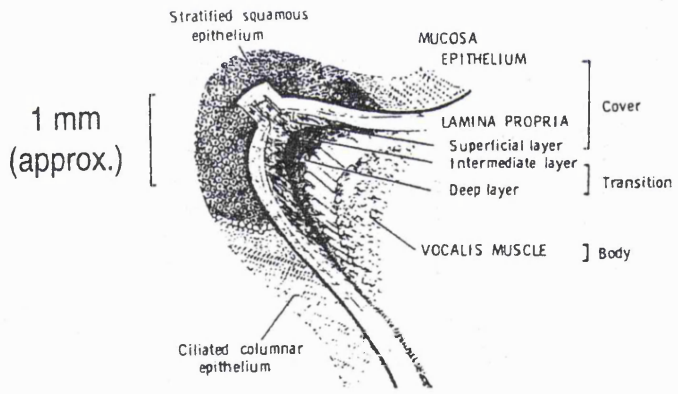


Figure II.6 - Non-scaled structure of the vocal fold - From Stevens (1998:9)

The vocal fold is composed of surface tissue (mucosa) and the underlying muscle (figure II.6). The mucosa consists of epithelium and the lamina propria (superficial also known as Reinke's space).

The epithelium can be regarded as a thin capsule, relatively stiff, inelastic and isotropic (the same stretchability longitudinally and crosswise) responsible for maintaining the shape of the vocal fold (Laver, Hiller & Beck, 1992). Stemple, Glaze

and Gerdeman (1995:36) state that *'it offers no mass, and is totally compliant, but needs a thin layer of mucous lubrication to oscillate best'*.

The lamina propria can be divided into three layers: superficial, intermediate (or middle) and deep as can be seen in Figure II.6.

The superficial layer (also known as Reinke's space) is a network of elastic and collagen fibres that acts as Laver, Hiller & Beck, (1992) pointed out '...like a liquid with a very high surface tension, and forms the medium through which the surface wave of the mucosal ripple travels in phonatory vibration.'
From a mechanical point of view, according to the cover-body model (Hirano, 1981), the epithelium and the superficial layer of the lamina propria form the *'cover'*, a totally passive structure, that moves smoothly over the deeper layers of the vocal folds during vibration, (Fujimura, 1981; Colton and Casper, 1996). Moreover, Colton & Casper (1996) note that the degree of coupling between the epithelium and the lamina propria is greater for high frequency speech than for low frequency speech activities.

The intermediate layer of lamina propria consists predominantly of elastin fibres (arranged parallel to the free border of the vocal fold), anisotropic (more stretchability longitudinally than crosswise) therefore incompressible (Laver, Hiller & Beck, 1992).

The deep layer is similar to the intermediate layer, but the fibres are mostly formed of collagen (flexible, but difficult to stretch) (Laver, Hiller & Beck, 1992). The intermediate and deep layers are significantly stiffer than the superficial layers. Together they form the vocal ligament, a thickened tissue located at the free edge of the conus elasticus as illustrated in Figure II.5. This structure is important for its elastic biomechanical properties during phonation and is designated *'transition'*.

The deepest layer of the vocal fold consists of striated muscle fibres, the most viscous (stiffest) of the mechanical structures. The most medial muscle section is often referred to as the vocalis muscle (anisotropic), and the more lateral fibres comprise the thyro-arytenoid muscles (figure II-4). This layer constitutes the *'body'* of the vocal fold and provides stability and mass (Laver, Hiller & Beck, 1992; Hirano and Bless, 1993).

Recent research (Gauffin and Hammarberg, 1991) has shown the importance of another histological area, the basement membrane, for the surgical field. However, because the basement membrane is responsible for the connection between the epithelium and the superficial layer, Sataloff (1998:10) suggests that additional research is needed *'to show its great importance in other matters such as the ability to heal following trauma, possibly the development of certain kinds of vocal fold pathology, and probably in histopathologic differential diagnosis'.*

The characteristics of vocal fold vibration depend on their elastic properties, mass distribution and their length. The parallel distribution of the elastic and collagenous fibers of the lamina propria and the muscle fibers of the vocalis muscle into the vocal edge facilitates the vibration of the vocal folds. All the layers are passively controlled by the laryngeal muscles and by the subglottal pressure effects. However, the muscle itself also actively controls the 'body'. Also, posteriorly and anteriorly the layered structure of the vocal folds is less flexible than medially, which contributes to different mechanical properties of the folds and protection of possible mechanical damage caused by vibration.

Two other important features advantageous to the maintenance of vocal fold vibration are the mucus and the vascular arrangement. The mucoserous secretion originating from the glands outside the vocal fold is an important feature for vocal fold vibration because otherwise the vocal fold surface would be completely dry and consequently could not vibrate properly, (Hirano and Bless, 1993).
Fourcin (1981) comments that the mucous blanket should be considered a separate layer, as it may influence glottal measurements using electrolaryngography (described in more detail in chapter IV).

Within the vocal folds the blood vessels are very small and run parallel to the edge of the membranous portion. Colton and Casper (1996:64) suggest that *'very few vessels enter the mucosa directly from the underlying muscle.'*

2.7 - Phonation

The study of vocal fold vibration (also known as the sound source) has been the subject of research of a wide range of voice scientists and clinicians. It is currently

accepted that this phenomenon results from a dynamic alternation of forces that requires knowledge of the histological pattern of the vocal fold (the cover-body theory from Hirano, 1981), as well as the knowledge of the combination of aerodynamic force (lung pressure and Bernoulli effect) and the elastic properties of the tissues (the myoelastic-aerodynamic theory from Van den Berg, 1958) (Hirano and Bless, 1993; Borden, Harris & Raphael, 1994; Stemple, Glaze & Gerdeman, 1995).

The subglottal pressure affects the intensity and the fundamental frequency of the sound source. According to Broad (1973) several experimental studies (van den Berg, 1956; Ladefoged and McKinney, 1963; Isshiki, 1964; Cavagna and Margaria, 1965) have shown that the acoustic intensity of voicing is approximately proportional to the subglottal pressure. He also states, that '*every doubling of the subglottal pressure adds 9 to 12 decibels to the intensity of the voice.*'

Besides the subglottal pressure, the adjustment of the larynx itself (the vertical difference, the longitudinal tension of the vocal folds, the mucosal wave and the viscous forces of vocal fold tissues) is another major source of voice control.

At the beginning of one cycle for neutral mode of phonation (also referred as modal voice quality), the vocal folds are approximated and start to blow apart (from bottom up and outward) until a maximum opening is achieved. In a coronal-section perspective the glottis is in 'convergent' shape (wide space at the level of the lower lip of the vocal folds to no space at the upper lip) (Figure II.7A).

The vertical phase difference is known as the temporal difference (between the upper and lower portions of the vocal folds) in the inward movement of vocal fold vibration. So, the vocal folds act as if they were independent but interconnected masses and at the maximum opening the 'upper lip' moves laterally while the 'lower lip' moves medially.

As the subglottal air escapes between the folds, subglottic air pressure decreases and the tissue elasticity of the vocal folds as well as the Bernoulli effect cause the folds to once again snap back and close the glottis from the bottom up as might be seen in Figure II.7.

The number of times per second the glottal cycle is repeated corresponds to the fundamental frequency (correlating with vocal pitch) of an individual's voice.

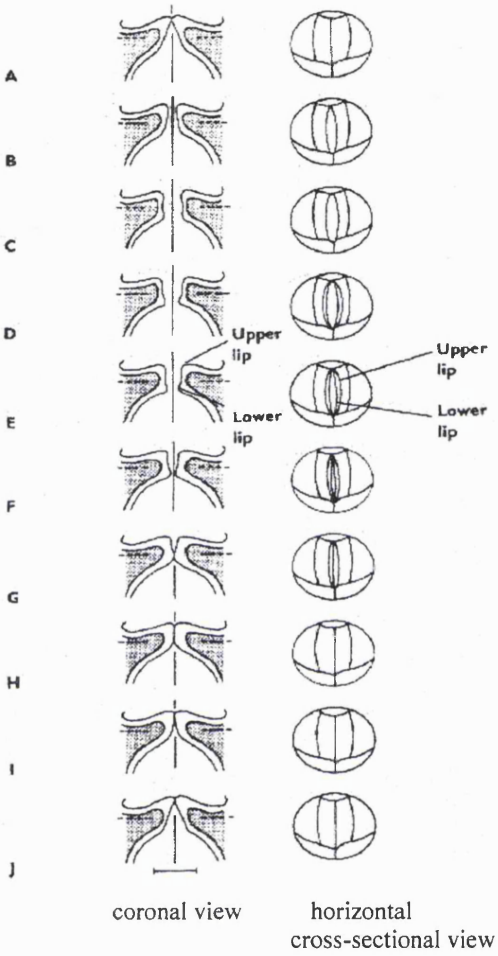


Figure II.7 - Schematic presentation of the vocal fold vibratory pattern
From: Hirano & Bless (1993:24)

Hirano and Bless (1993) mention that displacement and changes in the vocal fold cover create the '*mucosal wave*', to which Perkins & Kent (1986) add that it may be more essential to phonation than the muscle that mainly controls the shape of the vocal fold.

The viscous forces of vocal fold tissues determine how well they dissipate force applied to them and how much resistance they offer to aerodynamic forces. Perkins and Kent (1986:101) state that '*the more viscous the vocal folds, the more they decrease the*

velocity of any movement they are forced to make, and hence the more resistance they offer to aerodynamic forces.'

Resistance to air pressure force comes also from the longitudinal tension of the vocal folds because of the stiffness of the folds and thickness of the vocalis muscle (Perkins and Kent, 1986). Broad (1973) notes that there is some experimental evidence (van den Berg and Tan, 1959; Sonninen, 1954; Moore, 1937; Arnold, 1961; Fink, 1962; Hollien, 1960, 1962; Hollien and Moore, 1960; Luchsinger and Pfister, 1961) confirming the theoretical hypothesis that '*the lengthening of the vocal folds should be particularly effective in increasing the fundamental frequency because of its double function to increase the longitudinal tension and to decrease the effective vocal fold mass*'.

The duration of the closed phase is dependent on the pattern of collision and absorption of the vocal folds force impact. Perkins and Kent (1986:96) comment that '*The more compliant they are, the more they absorb the impact, and the longer they remain closed.*'

2.8 – Age and gender variations in 'normal' larynx

Although normative data is scarce and no consensus has been reached as to where normal ends and abnormal begins, it is accepted that the adult laryngopulmonary-resonance system shows a broad range of interpersonal variations in both structure and function according to age and gender and those are reflected in the acoustic properties of the human voice (described in detail in chapter III).

As a general rule laryngeal physical differences in terms of size, mass and volume are more evident between genders. On average male larynx length and height are about 40 percent greater than for the average female larynx. The thyroid cartilage is only about 20 percent larger in the anterior-posterior dimension in men than in women. It is also a matter of common observation that male vocal folds are larger than female vocal folds, although consensus about its length is no been reached (table II.1). Titze (1994b) found that the membranous length in the male larynx is 60 percent longer than in the female, which clearly explains the primary gender difference in fundamental frequency (F0). However, other features may also contribute to this gender difference in

the F0, such as glottis aperture; he notes that the glottis width of men is 5% larger at rest and its maximum is 30% greater than that of women.

Table II.1- Vocal fold length

Females	Males	Source
10 mm	15 mm	Tucker (1987, p.9)
16mm-2cm	2-2.4 cm	Becker, Naumann, Pfaltz (1994, p.388)
13 mm	18 mm	Maue (1970 cited by Rubin, 1998, p.19)
12-17 mm	17-23 mm	Hirano & Bless (1993); Mueller (1997 cited by Mathieson, 2001, p.94)

Sulter, Schutte & Miller (1996) in a study with 92 females and 47 male subjects without voice problems found that male subjects have thicker vocal folds than the female subjects.

Other important observations are that laryngeal function asymmetry and aperiodicity are common among 'normal' speakers. For example, Bless et al. (1993 cited by Andrews, 1999) found that normal adult speakers presented deviations from what might be considered normal (e.g. slight oedema and incomplete glottal closure).

Casper, Brewer and Colton (1987) examined fiberoptically 20 female and male subjects aging between 21 and 33 years judged to have normal voices and found that the most frequent variability was size, shape and symmetry differences in the arytenoid complex especially during phonation. Stroboscopic examination results on the 20 normal males (aging between 51 and 81 years) in Verdonck-de Leeuw (1998) study showed two speakers with slight supraglottic oedema and one speaker with vascular injection and supraglottic tension.

A posterior glottal chink in the cartilaginous glottis (thought to be pathological) has been found to occur frequently especially in women during modal phonation without being perceptually abnormal. Moreover, women tend to have shorter closed phase than do male counterparts (Biever & Bless, 1989; Södersten & Lindestad, 1990; Rammage, Peppard & Bless, 1992; Hirano and Bless, 1993; Morrison et al., 1994;

Hertegård, 1994; Linville, 1995; Söderstern, Hertegård & Hammarberg, 1995; Andrews, 1999).

Söderstern, Hertegård & Hammarberg (1995) found that the posterior triangular chink was the predominant pattern of closure in a group of seventeen healthy women (mean age, 52 years old) during normal phonation. Nevertheless, Stemple, Stanley & Lee (1995) observed anterior glottal chinks in six of the 10 healthy females (mean age, 25,3 years) after prolonged speaking task . None of those females presented the glottal chink before the experimental reading task. Also, Eustace, Stemple & Lee (1996) found that 37% of the women from 88 patients complaining of laryngeal fatigue in the absence of laryngeal pathology presented with an anterior glottal chink.

Larynx changes in the female (e.g. tissue changes such as oedema, hyperemia, microvarices, micronodules and larynx hypotension during phonation) were found to be dependent on the hormonal changes during stages of the menstrual cycle and menopause (Abitbol et al., 1995; Chernobelsky, 1998). Moreover, Higgins and Saxman (1989a) stress that 'given the evidence of a hormonal influence on neural functioning, motor movement, and sensory thresholds, ovarian hormones might affect vocal behaviour by altering the afferent and efferent processes involved in laryngeal neuromotor control'. Nevertheless, it should be noted that between 20% and 50% of the females suffer from emotional and behavioural symptoms diagnosed by the DSM-IV (1994) as premenstrual syndrome (Chae et al., 2001).

Several studies recognise that physical changes in the structure of the larynx are dependent on the ageing process. Atrophy of the intrinsic laryngeal muscles, thinning and dehydration of the laryngeal mucosa, loss of elasticity of ligaments, calcification of cartilages, flaccidity and bowing of vocal folds have been found (Ramig & Ringel, 1983; Zemlin, 1988; Biever & Bless, 1989; Linville, 1996; Hirano and Bless, 1993; Russel, Penny & Pemberton, 1995; Stemple, Glaze & Gerdeman, 1995).

Zemlin (1988:177) comments that '*the thyroid and cricoid cartilages begin to ossify in the early twenties and by 65 years the entire laryngeal framework, except for the elastic cartilages, has usually turned to bone.*' Moreover, Hirano and Bless (1993) also suggest that significant changes in supraglottic activity, vibratory amplitude, mucosal wave, and phase symmetry appear to exacerbate with age.

Those predictable changes of ageing are once again gender related and show great individual differences. In the male larynx, histological changes may be noted in the fourth decade, whereas female ageing signals do not generally appear until the fifth decade of life. Changes in the male larynx include muscle atrophy, reduced elasticity, and fibrotic changes of the lamina propria. Reductions in bulk, muscle tone, and elasticity may explain a tendency for the natural fundamental frequency of the male to rise in later years. In the female, atrophic changes may be less significant, but greater bulk of the vocal folds often results from increased oedema in the superficial layer of the lamina propria. The overall effect of phonation is to produce a potentially lower fundamental frequency of vibration. On the other hand vocal fold bowing is more common in males in the senior age group than in geriatric females. In addition to these gender-specific changes, systemic diseases such as arthritis, respiratory ailments, and neuromuscular disease can influence characteristics of the ageing voice (Kaplan, 1971; Morrison et al., 1994; Colton & Casper, 1996; Sataloff, 1998).

Nevertheless, the male-female coalescence model stresses that chronological age on its own does not account fully for the observed physiological changes but rather additional human physical fitness, mental alertness and emotional equilibrium factors play a role (Krook, 1988; Yamazawa and Hollien, 1992).

Chapter III - Dysphonia

3.1 - Introduction

In chapter II an outline was given of the framework and muscular adjustments of the larynx that are responsible for the creation of the sound source in speech. Moreover, it was pointed out that these functions involve a physiological continuum dependent on the size and shape of the vocal folds, their speed of vibration and relative proximity.

The following chapter presents a contemporary overview of voice definition, classification, aetiology and incidence of dysphonia. Also, a critical review on F0 data will be presented.

3.2 - Voice quality continuum between 'normal' and 'dysphonia'

In a general sense, researchers define dysphonia (voice disorder) as an abnormality in the sound of the voice, where the quality, pitch, loudness or flexibility are interpreted as being unpleasant or inappropriate to the age or gender of the speaker and unacceptable in social, professional or other contexts (Enderby & Emerson, 1995) to which Verdolini (1994) adds '...and for which the SLT or other practitioner generally finds corroborative evidence.'
Verdolini & Ramig (2001) point out that 'In most or all contemporary epidemiological studies on voice conducted to date voice disorder (or voice problem) is explicitly or implicitly defined as a condition of sufficient concern for the bearer to report it, register functional disruption because of it, and/or seek treatment because of it.'

Obviously, problems arise with the definition of several concepts such as quality, acceptability and abnormality, which depend on notions of normality. Knowledge of 'normal' voice characteristics and variations is crucial for both the early detection and the differential diagnosis of changes resulting from disruption or dysfunction of voice quality (dysphonia/voice disorder). The location of the threshold that separates the 'normal' from 'disordered' is judged by each listener on the basis of his cultural standards, education, environment and vocal training (Russel, Penny & Pemberton, 1995).

Although researchers have been interested in voice quality for at least a century, a unanimous definition of what constitutes a 'normal' voice has still not been achieved. Due to the multiple psycho-acoustic complexities of the human voice and its diverse psychophysiological, social and cultural involvement and also considering the different perspectives of the professionals involved in voice care, an enormous diversity of definitions and terminologies exists.

Some define it on the basis of certain parameters (quality, intensity, fundamental frequency, nasality, time and rhythm), speech variability constituents (Pindzola, 1987; Wilson, 1987; Mathieson, 2000) and anatomophysiological aspects (Hirano, 1981); some take cultural factors into account (Moore, 1971 cited by Fourcin, 2000), and others include aesthetics in their definitions (Hollien, 2000). Moreover, in considering voice quality some researchers mean qualities that concern laryngeal parameters only ('source') (Hirano, 1981; Hammarberg & Gauffin, 1995) whereas others propose that is the result of laryngeal ('source') and supralaryngeal ('filter') features (Laver, 1991; Mathieson, 2000).

In judging the normal voice Aronson (1990) stresses its adequacy to carry language intelligibility to the listener; its aesthetically acceptable acoustic properties and its occupational and social functional requirements satisfaction. For Wilson (1987) the normal voice should exhibit '...pleasing voice quality, proper balance between nasal and oral resonance, appropriate loudness, modal fundamental frequency suitable for age/size/sex, and appropriate inflections in fundamental frequency and vocal intensity'. Mathieson (2000) defines normal voice as '... audible in a wide range of acoustic settings, even with relatively high levels of ambient noise. It is appropriate for the gender and age of the speaker and is capable of fulfilling its paralinguistic and linguistic functions to the speaker's satisfaction.' Also, '... [it] should not deteriorate with use nor should there be any discomfort or pain associated with phonation.' Fourcin (2000:290) comments that '...normality is partially subjective, it can be culturally determined (e.g., Moore, 1971), and thus it may be phonetically defined as a function of the language environment in question (e.g. Maddieson, 1984), or may simply exist in the ear of the particular listener (e.g., Aronson, 1985).'

Each voice is markedly individual because it is dependent on the functioning and structure of the subject's larynx and on his/her age and gender (described in detail in chapter II) as well as cultural influences and speaker's vocal demands. Moreover, each

voice within a normal-disordered continuum may vary temporarily to suit the occasion related to emotional situations, time of the day and ambient noise, to different communication demands and attitudes (e.g. whisper a comment about something, dialogue and vocal loading), to different social habits influences (e.g. smoking, alcohol and caffeine) and to different health related influences (e.g. stress, medication consumption and diseases).

Creaky voice (resulting from vocal fold vibration disperiodicity) is pervasive in some speakers but it is also an acceptable voice type when associated with low terminal falling intonation patterns or associated with some cultures (e.g. English and Swedish male speakers) (Fourcin, 1993; Clark & Yallop, 1995; Abberton et al., 1999; Mathieson, 2000). Also, Cooper and Yanagihara (1971 cited by Mathieson, 2000) found that the morning or post-sleep voice uses a basal or near basal pitch level due to the relaxed state of the individual's muscular system, following a period of sleep . Heaton (1989 cited by Holmberg et al., 1994) notes that breathiness, within a certain degree, in females, is considered attractive. Also, Drew & Sapir (1995) note that USA women appear to appreciate low-pitch voice as being dramatic, sensual and wordly.

Slight episodic phonation irregularities are evident in all 'normal' voices at voicing onset and offset due to the momentary 'catastrophic' myoelastic and aerodynamic vocal fold vibration and/or due to momentary variations in the mucoserous secretion (Baken & Orlikoff, 1988; Laver, Hiller & Beck, 1992).

Hormonal changes during the menstrual cycle and menopause suggest variations on the females' voice (slight hoarseness and vocal fatigue) during the premenstrual and menstrual phases (Sataloff, Spiegel & Hawkshaw, 1995; Abitbol et al., 1995) and lowering of the fundamental frequency at the menopause (Krook, 1988; Greene & Mathieson, 1989; Russel, Penny & Pemberton, 1995).

Also, a dysphonia of short duration caused by upper respiratory tract inflammation is generally accepted with no justification for formal voice evaluation. Nevertheless, within the expected range of variability, persistent or an increasing number of episodes of voice disorders, clinicians and speakers habitually classify abnormalities as dysphonia (Gordon, 1991).

Prescribed medication has been described as having potential minor to significant clinical impact on the larynx and voice although those effects are dependent on the individual biological response (e.g., age, body composition, vocal demands, stress), the dose-response relationship, multiple drug effects and the concurrent use of other medication (Martin, 1983; 1984; Biever & Bless, 1989; Harris, 1992; Sataloff et al., 1994b; Colton & Casper, 1996). Colton & Casper (1996) note that research on medication and voice effects (Leonard & Ringel, 1979; Watkin & Ewanowski, 1979; Williams et al., 1983; Ludlow, Schulz and Naunton, 1988) is almost absent; however, the voice-related effects of medication and drugs were classified into seven categories by Martin (1983; 1984) (presented in table III.1).

Table III.1 –Medication and drugs' effects on the larynx and voice.

Drug effect	Drug type
Coordination and proprioception	Central nervous system stimulants (e.g. amphetamines, cocaine, caffeine, appetite depressants) and depressants (e.g. barbiturates; tranquillisers; alcohol) and anaesthetics (e.g. cocaine).
Airflow	Bronchial dilators and constrictors.
Fluid balance	Diuretics, decongestants and corticosteroids.
Secretion of the upper tract	Drying (e.g. antitussives, antidepressants, antihypertensive) and wetting agents.
Structure of the vocal folds	Androgens (frequently used in treatment of hormonal imbalance in postmenopausal women).
Irritation of the vocal fold mucosal	Lower esophageal sphincter drugs treatment (e.g. atrophine, dopamine, sedatives, calcium blockers), smoking, alcohol, highly fat and spiced foods.
Miscellaneous	Ototoxic drugs (e.g. antibiotics), herbal teas, aspirin, tobacco and other smoked or inhaled drugs.

Voice fatigue often related to symptoms of laryngeal aching, throat thickness, neck tightness, and dryness is a common complaint of those who require great daily voice use (Stemple, Stanley & Lee, 1995; Eustace, Stemple & Lee, 1996). However, as Vilkman (1996) notes voice disorder prevention from the point of view of occupational health is poorly developed as compared to occupational hearing disorders and insurance companies rarely accept a voice disorder as an occupational disease. Studies of the experiences of some professions, which are more directly dependent on vocal

communication than others, confirm that some are at greater risk for voice disorders than others (Sapir, Keidar & Mathers-Schmidt, 1993; Arnoux-Sindt et al., 1994; Sapir, Mathers-Schmidt & Larson, 1996; Fritzell, 1996; Vilkman, 1996, 2000). Although the working vocal loading demand is the basic risk factor for voice problems there are other potential environmental and ergonomic factors as Vilkman (1996) stresses that '...unfavourable biomechanical factors, the quality of the working environment, psychological factors, especially stress, and inadequate or neglected examination and treatment of early symptoms' contribute to voice problems.

Despite this interest, some difficulties emerge when attempting to establish a precise definition of professional voice. Among the published literature available there is a reference to Koufman & Isaacson's (1991 cited by Stemple, 1993) classification of professional voice use, and an attempt to classify different professions according to voice quality demands and estimated vocal loading from Vilkman (2000).

Table III.2 - Koufman & Isaacson's (1991 cited by Stemple, 1993) classification of professional voice use.

Level 1	**Elite vocal performer** - a slight vocal difficulty may cause serious professional consequences, e.g. professional singers and actors.
Level 2	**Professional voice user** – even a moderate vocal difficulty would prevent adequate job performance, e.g. clergy, public speakers, lecturers, telephone operators, airline reservationists.
Level 3	**Nonvocal performer** – a severe dysphonia would prevent him/her from work while mild or moderate dysphonias may be inconvenient but would not preclude adequate job performance, e.g. doctors, lawyers, business persons and sales persons.
Level 4	**Nonvocal nonprofessional** – would not be prevented from work if experiencing vocal disability, e.g. factory worker, labourer and clerk.

Koufman & Isaacson's classification (1991 cited by Stemple, 1993) summarised in Table III.2 presents four levels of voice use according to handicap results, while the Vilkman (2000) system (summarised in table Table III.3) has four different characterisations according to quality (high, moderate or low voice quality) and vocal loading (factors affecting voice production such as the length of time speaking, the acoustic conditions at work, the size of working space, reverberation properties, background noise) as well as the characteristics of speech communication (size of groups). Koufman & Isaacson (1991 cited by Stemple, 1993) add that the history and needs of each person will determine his/her personal vocal level.

Table III.3 - Vilkman's (2000) voice classification

Voice quality precision	Load	Profession
High	High	Actors, singers.
High	Moderate	Radio and TV journalists.
Moderate	High	School and kindergarten teachers, telephone operators, telemarketers, military, clergy, cantors.
Moderate	Moderate	Bank, business and insurance personnel, physicians, lawyers, nurses.
Low	High	Foremen, welders, platers.

So, although many authors (Hirano, 1981; Laver, 1991; Hirano and Bless, 1993; Colton & Casper, 1996) agree that dysphonia may be partially dependent on the pathology itself (histological pattern, localization, size of lesion, extent and degree), laryngeal activity (glottal competence, symmetry, mass and stiffness of both vocal folds and compensatory behaviour) as well as the balance between mechanical and aerodynamic factors of phonation, the wide range of individuals' anatomical and physiological differences should also be considered, because of their capacities to compensate for the deficit as well as for the differences in phonatory patterns, (Milutinovic, 1996). Therefore, it is not possible to establish a unique and perfect relationship between laryngeal conditions and dysphonia. Moreover, the existence of laryngeal pathology may not result in vocal impairment or vice-versa because of the individual's characteristics or the demands made on his/her voice.

Another great constraint, related to what has been described above, is the non-existence of standard unambiguous terminology. Some researchers (Fex, 1992; Orlikoff et al., 1999; Carding et al., 2000) state that the terminology used is based on a wide range of positive and negative aspects, basically defined in acoustic and physiological terms. However, some of the terminology applied is based on impression alone. We have, for example, visually-based concepts (e.g. 'bright', 'clear', 'dark'), kinaesthetic terms (e.g. 'tense', 'forced', 'tight', 'vibrant'), anatomical references (e.g. 'nasal', 'guttural', 'throaty', 'head'), material concepts (e.g. 'metallic', 'velvety', 'golden'), spatial terms (e.g. 'broad', 'narrow', 'flat'), aesthetic values (e.g. 'beautiful', 'unpleasant'), auditory impression (e.g. 'loud', 'high', 'rough', 'breathy'), qualities based on taste (e.g. 'sweet', 'mellow'), and characteristics which are linked to personality (e.g. 'hesitant', 'expressive'). Moreover, the terms 'dysphonia', 'voice

disorder' and 'pathological voice' are used interchangeably in the enormous amount of published literature.

Attempts to overcome this serious problem of terminology have been made by several organizations, (the Japanese Society of Logopedics and Phoniatrics, the International Association of Logopedics and Phoniatrics, and the British Voice Association), and researchers around the world (Sonninen & Damste, 1971; Fex, 1992; Tveteras, 1992; Sonninen & Hurme, 1992; Hammarberg, 2000). However, it is difficult to establish a uniform and systematic terminology for 'normal/disordered' voice qualities in 'normal' and 'pathological' voices, especially due to the different semantic values of the voice-describing words in each language (Sonninen & Damste, 1971; Fex, 1992; Tveteras, 1992; Sonninen & Hurme, 1992; Hammarberg, 2000).

3.3 - Classification of dysphonia

An inspection of the numerous published textbooks (Prater & Swift, 1984; Tucker, 1987; Le Huche & Allali, 1990; Aronson, 1990; Laver, 1991; Haynes, Pindzola & Emerick, 1992; Hirano & Bless, 1993; Titze, 1994b; Verdolini, 1994; Morrison et al., 1994; Stemple, Glaze & Gerdeman, 1995; Milutinovic, 1996; Colton & Casper, 1996; Benninger et al., 1998; Mathieson, 2001) on the classification of dysphonia ('voice disorders') reveals a wide variety of criteria used, either related to the underlying pathology (processes acting within the tissues), to the presumed aetiology (factors arising externally to the tissues), to vocal fold mobility (adduction-abduction; tensing-relaxing), or according to the perceptual voice characteristics (Fawcus, 2000). Habitually, the classification ('label') used will depend on who is applying it, for example, the ENT surgeon on the basis of the larynx examination or the aetiology of the problem, the SLT according to perceptual evaluation or to acoustic measures, other professionals (e.g. linguists, singing teachers).

3.3.1 – The functional versus organic classification

The simplest approach used is the traditional aetiology division into two categories, organic and non-organic or functional. Organic voice disorders are those with an identified specific structural, neurogenic, endocrinal, or laryngeal disease cause, whilst functional dysphonia is normally defined as a voice disorder which cannot be

explained by any organic laryngeal lesion, (Fawcus, 2000; House & Andrews, 1987; Martin, 1993; Titze, 1994b; Morrison et al., 1994; Enderby & Emerson, 1995; Mathieson, 2000). For this reason, according to Titze (1994b), some clinicians have labelled functional dysphonia as an idiopathic phenomenon (of unknown cause) often associated with abnormal muscle activation patterns. Also, others like Aronson (1990) use the term 'psychogenic' as synonymous with 'functional'.

Obviously, this classification poses several problems, especially in terms of the classification of functional voice disorders and the boundaries between the two categories. In many instances, voice disorders are caused by a mixture of a broad number of factors, some physical (organic) and some not strictly physical (behavioural/functional). Behavioural disorders range from muscular tension disorders, with or without laryngeal mucosal changes such as nodules and polyps (Morrison et al., 1994) to psychogenic causes of various types and severity (Aronson, 1990).

For the purpose of this study, dysphonia is categorized according to a system considered by researchers (Verdolini, 1994) to be more useful for assessment, namely (a) minor structural and physiological alterations of the larynx, (b) discrete mass lesions of the vocal folds, (c) distributed vocal fold tissue changes, (d) neurological disorders, and (e) non-organic or muscular tension disorders. Structural or physiological pathologies outside the larynx system will not be considered although it is recognised that they may, direct or indirectly, affect phonatory behaviour.

Detailed descriptions of the aforementioned types of dysphonia can be found in various textbooks (Prater & Swift, 1984; Frèche et al., 1984; Tucker, 1987; Greene & Mathieson, 1989; Aronson, 1990; Colton & Casper, 1996; Le Huche & Allali, 1990; Hirano & Bless, 1993; Verdolini, 1994; Titze, 1994; Morrison et al., 1994; Dworkin & Meleca, 1997; Mathieson, 2001).

3.3.1.1 – Minor structural and physiological alterations of the larynx

Pontes, Gonçalves & Behlau (1999) refer that minor alterations are subtle deviations from 'normal' laryngeal anatomy that may occur independently or in association with another larynx pathology whilst Hirano & Bless (1993) designate it as physiological variations. As has already been mentioned physiological variations are also gender and age related (described in detail in chapter II). A summarised description

of some common minor structural and physiological alterations of the larynx and presumed primary cause is presented in Table III.4.

Table III.4 – Some minor structural and physiological alterations of the larynx

Type of alterations and description	Presumed primary cause and possible contributors
Systemic changes	
Dehydration Increased viscosity of secretions on the vocal folds and other respiratory linings.	Dehydration, drying agents such as medication, alcohol and caffeine and environmental dryness.
Hyperemia Increased vascularisation of the vocal folds.	Vocal abuse, upper respiratory tract infection, medications (e.g. analgesics and corticosteroids), persistent cough and menstruation.
Oedema Increased volume (by fluid retention) of the vocal folds.	Voice use or abuse, excessive coughing and throat clearing, menstruation, menopause, pregnancy, smoking and drugs (steroids, oestrogen and testosterone).
Inflammation Laver (1991) notes that '... inflammation can involve capillary dilation, an infusion of white blood cells, collection of oedematous fluid in the intercellular space, a proliferation of collagen fibres and granulation tissue, and the deposition of hyaline.'	Sensitivity to the propellants used in inhalers, smoking and smoky environments, acute upper respiratory tract infection and sinusitis.
Structural and laryngeal vibration asymmetries and/or glottal closure Slight differences in the position, shape, mass, tension, elasticity and/or viscosity of the vocal folds. Differences in the amplitude phase of vocal fold vibration and/or glottal closure. Detailed information can be found in Hirano (1981); Hirano & Bless (1993) and Poburka (1999).	Gender, age, professional voice use, voice registers and unknown causes.

Depending on the location and the degree of the abnormalities and the subject voice use the voice impact may be clinically insignificant or not and this is the reason

why some researchers call it simple dysfunctional dysphonia (Le Huche & Allali, 1990).

Elias et al. (1997) found 42% of larynx alterations (erythema and/or oedema due to reflux laryngitis, varicosity without evidence of prior haemorrhage, asymmetry, incomplete glottal closure, pseudo nodules) in 65 healthy singers without voice complaints.

Sulter, Schutte & Miller (1996) observed that horizontal phase differences are more frequent in females whilst vertical phase differences are more frequent in males. The group consisted of 92 female and 47 male subjects without voice problems.

3.3.1.2 - Discrete mass lesions of the larynx

Discrete or circumscribed mass lesions include lesions that affect the membranous part of the vocal folds, the posterior glottis and the ventricular space. These lesions may be benign (nodules, polyps, cysts, papilloma, contact ulcers, granuloma and laryngocele), pre-malignant (keratosis, leukoplakia, pachydermia) or malignant tumors (laryngeal cancer). Moreover, lesions may be unilateral or bilateral. A summarised description of the abovementioned pathologies and presumed primary cause is presented in Table III.5.

Laver (1991) notes the importance of taking into account the location of any increase in mass in the final contribution to vocal fold vibration pattern. He also stresses that the mass effects will be more evident at voice onset.

Although with exceptions, 'hoarse' voice quality is very common in all types of discrete mass lesions because these usually do not allow complete vocal fold closure during phonation (resulting in breathiness) and, furthermore, because the pattern of vibration is irregular due to the uneven mass distribution (resulting in harshness). In addition to this, some behavioural factors such as smoke, alcohol, caffeine, diuretics, drying nasal sprays, dry air, and certain (drying) medications may cause vocal fold dryness and stickiness, therefore potentialising the pathological voice quality (Verdolini, 1994).

Voice quality may be nearly normal if the lesion is on the posterior glottis and does not interfere with the 'phonatory glottis'. However, complaints such as effort or pain during phonation or swallowing are commonly reported (Verdolini, 1994).

Table III.5 - Discrete mass lesions of the larynx

Discrete mass lesions	Description	Presumed primary cause and possible contributors
Membranous folds		
Nodules	Oedematous and/or collagenous protrusions at midpoint of membranous folds. Usually bilateral (Colton & Casper, 1996).	Mechanical stress (high vocal fold impact force and viscosity).
Polyps	Localised (pedunculated) or distributed (sessile) protrusions at midpoint of membranous folds; may be haemorragic, fibrous or oedematous; unilateral or bilateral usually asymmetrical. Colton & Casper (1996:69) note that 'the histological features of the lesion are oedema, intratissue bleeding, hyaline degeneration, thrombosis, proliferation of collagenous fibers, and/or cell infiltration.'	Mechanical stress (high vocal fold impact force and viscosity) in combination with air pollutants (smoke or chemicals). Glandular blockage.
Cysts	Protrusions (fluid-filled sacs of epithelium) along membranous folds, often at midpoint. May be epidermoid or retention cysts. The lesion is either uni or bilateral and with different sizes.	Glandular blockage.
Papilloma	Solitary or multiple small epithelial benign tumors that may invade the lamina propria and the muscle.	Viral infection.
Keratosis leukoplakia	White, plaque-like lesions of fibrous protein substance (keratin) on the epithelial surfaces of the vocal folds.	Smoking habits and irritating airbone substances (smoke, dust, chemical alterations of air quality).
Cancer	White, malignant grainy lesions that not only invade the surface (cover) tissues but also penetrate into the vocal fold body. The lesion is originally unilateral but can develop to the contralateral side. The most current classification is the TNM-staging where the larynx is divided into three anatomical sites. The T-division describes the primary tumour, the N-division the metastases to lymph nodes in the neck , and the M-division distant metastases (Verdonck-de Leeuw, 1998).	Unknown aetiology but closely related to heavy smoke and alcohol consumption. The risk ratio increases in a multiplicative rather than in an additive manner (Verdonck-de Leeuw, 1998).
Posterior glottis		
Contact ulcers	Protrusions on medial surface of arytenoids, possibly with contralateral concavity.	Gastric reflux; 'pressed' phonation; chronically low pitch.
Granuloma	Grainy vascularised tissue, on the medial arytenoid surface, bilaterally or unilaterally. Colton & Casper (1996:71) note that 'histologically, granulomas consist chiefly of proliferated capillaries, fibroblasts, collagenous fibers, and leukocytes.'	Gastric reflux; 'Pressed' phonation; chronically low pitch; intubation.
Ventricular space		
Laryngocele	Enlargement of the ventricular space.	Probably related to excessive pressures in the laryngeal region (e.g. trumpet players or oboe players).

3.3.1.3 - Distributed tissue changes

Titze (1994b) states that vocal fold tissue changes may result from 'viral or bacterial infection, and may be subject to chronic or acute mechanical stress, irritated by pollutants, or become dehydrated.' A summarised description of some common distributed tissue changes pathologies and presumed primary cause is presented in Table III.6.

Table III.6 – Some common distributed tissue changes

Distributed Tissue Changes	Description	Presumed primary cause and possible contributors
Reinke's oedema (also known as polypoid degeneration, chronic polypoid corditis, chronic oedematous hypertrophy and polypoid vocal fold)	Widespread accumulation of fluid, or swelling (oedema) along membranous folds (Reinke's space). Usually bilateral and asymmetrical.	The aetiology is unknown but is related to older subjects with chronic exposure to irritants (e.g. smoke).
Laryngitis	Acute or chronic distributed inflammation. The vocal folds may appear reddened and swollen.	Bacteria or viral agents.
Sulcus vocalis also known as vergeture (Frèche et al., 1984:168).	Uni or bilateral longitudinal groove along the membranous vocal fold. According to Hertegård (1994) the amplitude of vibration and mucosal wave are decreased, during phonation, in the majority of cases.	Unclear, suspicious of congenital and developmental factors; secondary to repeated or long duration chronic inflammatory processes; heavy voice use.
Vocal fold scar	Stiffer tissue (dense collageneous fibers) with variable vocal folds localization.	Trauma, surgery, burn or inflammation.
Trauma	Appearance dependent on the nature of trauma and the extent of the lesion.	Mechanical, thermal, or chemical trauma.
Haemorrhage	Rupture of blood vessels in the cover or the muscular portion of the vocal folds.	Severe screaming (e.g. ballgames, rock concerts or in other extremely noisy environments) or accidental injury.

Le Huche et Allali (1990:113) note that Reinke's oedema was first used in the German literature by Hajak (1926) whilst in the English literature the terms used were polypoid corditis (Jackson, 1930), polypoid degeneration (Holinger, 1950; Wallner, 1963), chronic hypertrophic laryngitis (Putney, 1953; Clerf, 1946) and oedematous fibrome (Myerson, 1950) and in the French literature were chronic pseudo-myxomatous laryngitis and pseudo-myxome (Moulonguet & Girand, 1943).

3.3.1.4 - Neurological disorders

A series of neurological diseases may produce paralysis (loss of muscle action), paresis (muscle weakness such as in Parkinson's disease, motor neurone disease and myasthenia gravis) and tremors and seizures. However, the most common neurologically based dysphonia (Titze, 1994b; Colton & Casper, 1996) results from recurrent laryngeal nerve (RLN) lesions (and is the only neurological condition considered in this study, and therefore the only one which will be focused on).

A summarised description of recurrent laryngeal nerve paralysis and presumed primary cause is presented in Table III.7 and detailed information may be found in Prater & Swift (1984:125); Greene & Mathieson (1989: 293); Hirano & Bless (1993:167); Becker, Naumann & Pfaltz (1994: 405); Crary & Glowaski (1996:301) and Colton & Casper (1996:74).

Table III.7 – Recurrent laryngeal nerve paralysis

Neurological disorders	Description	Presumed primary cause and possible contributors
Recurrent laryngeal nerve paralysis	Failure or limited (medial) movement of affected fold (usually unilaterally) contributing to a glottal incompetence.	Local trauma (accidents or surgery), virus or heart disease.

From a neuropathological perspective there are three types of paralysis, neuropraxis (nerve impulse temporarily blocked), axonotmesis (partial section of the axons of the nerve) and neurotmesis (complete section of the axons of the nerve) (Hirano & Bless, 1993; Colton & Casper, 1996).

According to Crary & Glowaski (1996) the unilateral recurrent laryngeal nerve paralysis dysphonia is characterised by breathiness, hoarseness, reduced loudness, possible diplophonia and high pitch whilst in the bilateral lesion the dysphonia is characterised by mild breathiness, mild hoarseness and reduced loudness.

3.3.1.5 - Non-organic or muscular tension disorders

As has already been mentioned from the long-standing debate about voice classification, in general, researchers agree that 'functional' and 'non-organic' voice disorders are those associated with symptoms of vocal fatigue (phonasthenia) due to vocal loading (progressively worse at the end of the day), throat discomfort in the absence of apparent laryngeal structural lesions but almost always associated with signs of laryngeal muscle tension. Nevertheless, in the literature, there is an abundance of synonymous terms: hyperfunctional voice disorders (Hirano & Bless, 1993), functional dysphonia (Kotby et al., 1993), psychogenic dysphonia (Aronson, 1990) or more recently, muscle tension dysphonia (MTD) (Stemple, 1993) in spite of the classical terms like non-organic dysphonia (Eustace, Stemple & Lee, 1996; Carding, Horsley & Docherty, 1998). Moreover, according to Fawcus (2000) attempts to classify voice on a continuum of overadduction and underadduction gave rise to synonymous terms such as 'hypo/hyperkinetic' (Luchsinger & Arnold, 1965; Greene, 1980), 'hypo/hypervalvular' (Brackett, 1971) and 'kinesiologic' (Aronson, 1980).

Freeman (2000) comments that debate about the psychological and physical mechanisms involved in functional disorders leads to the use of another term 'psychogenic' to distinguish within the functional disorders those with predominate psychological aetiologies. In fact, Aronson (1990) says that a 'psychogenic dysphonia' is synonymous with 'functional' but has the advantage of being more specifically related to its cause which is the manifestation of one or more types of psychogenic disequilibrium such as anxiety, depression, conversion reaction or personality disorder. However, other clinicians (Morrison et al., 1994; Colton & Casper, 1996; Rosen & Sataloff, 1997) are in disagreement with Aronson's conceptualisation because as Freeman (2000) reports '... predictions about personality cannot be made on the basis of the vocal and laryngeal symptoms alone.'

According to Eustace, Stemple & Lee (1996) some researchers (Jackson, 1940; Hirano et al., 1969; Greene, 1972) describe vocal fatigue as a weakness in the thyroarytenoid muscle due to vocal loading or straining and this may cause several different laryngeal features/presentations (e.g. glottal gap, ventricular phonation, mutational falsetto).

A summarised description of some common non-organic or muscular tension disorders and presumed primary cause is presented in Table III.8.

Table III.8 – Some common non-organic or muscular tension disorders

Non-organic disorders	Description	Presumed primary cause and possible contributors
Hypofunctional (hypokinetic) dysphonia	Weakness in the adductory action of the vocal folds (but not in arytenoid adduction) and, possibly, a slight deformation in membranous vocal fold tissue (bowing).	Chronic heavy voice use. Old age.
Hyperfunctional (hyperkinetic) dysphonia	Hirano & Bless (1993) define it as an excessive intrinsic and extrinsic laryngeal muscle activity that may be associated with organic pathology or not.	Functional compensation to postural misalignments (Morrison et al., 1994); inefficient respiratory patterns, situational, faulty learning (e.g. imitating others), medical (e.g. allergies), laryngeal irritant (e.g. tobacco) compensatory and psychogenic factors (Oates, 2000).
Posterior glottal gap	Persistent posterior glottal gap during phonation, and complaints of voice problems.	Simultaneous contraction of laryngeal adductors and abductors. Idiosyncratic habitual adjustments of the phonatory musculature (Laver, Hiller & Beck, 1992).
Ventricular phonation also known as Plica ventricularis or Hypertrophy of the false fold.	Titze (1994b) attests that 'the false vocal folds (ventricular folds) are either adducted more than the true vocal folds, thus creating the primary vibration in the ventricular regions, or they are adducted together with the true vocal folds, creating a complicated vibration pattern between the two systems.' Hertegård (1994, p.20) shows that ventricular dysphonia is associated with pressed/strained voice quality.	Compensatory function of the true vocal folds (e.g. after surgery or performance-related activities such as in instrument players and in the theatre) (Hertegård, 1994).
Mutational falsetto also known as Puberphonia and delayed pubertal voice change.	Persistent high pitch falsetto voice during spontaneous speech in post-pubertal male, with seemingly normal larynx and the ability to produce lower, more appropriate pitches on stimulation.	Psychological conflict; learned behaviour.
Conversion (or hysterical) aphonia/dysphonia	Lack of voice (aphonia) or hoarse voice although with seemingly normal larynx. Titze (1994b) explains that 'vocal fold tension may be excessively large, or the folds may be put into a prephonatory position that does not allow vibration to take place, even when the air pressure is raised to an appropriate level.' Normal voice can be elicited during vegetative functions such as gargling or coughing.	Psychological mechanisms not well understood; extreme emotional stress.

Neverthless, whether or not this situation will lead to damage of the epithelium and lamina propria of the vocal folds (e.g. oedema, inflammation, haemorrhage, nodules) will depend on the individual' susceptibility and the frequency, extent and severity of the vocal behaviour (Oates, 2000).

3.3.2 – Classification based on perceptual/acoustic phenomena

The attempt to arrive at a consensus with regard to the definition and classification of types of voice quality, and the terminology involved, has given rise to much discussion over the years, either by individual researchers or committees (eg. Hollien, 1983; Collegium Medicorum Theatri). On the whole, researchers agree that there are multiple sources for those perceived types of voice quality but the larynx is the primary source (Laver, 1980, Hollien, 1983; Perkins & Kent, 1986; Seikel, King & Drumright, 1997).

According to Ladefoged (1971), the adjustment of the vocal folds during speech varies in different ways, which have come to be known as phonation types. These include the laryngeal actions referred to by Peterson & Shoup (1966 cited by Broad, 1973) and phonatory settings defined by Laver, (1991). Several researchers (Daniloff, Schuckers & Feth, 1980; Perkins & Kent, 1986; Titze, 1994b) add that the different modes of larynx biomechanics, which are perceived to yield different sounds, are traditionally called registers. According to Titze (1994b) voice register is the term used to describe the perceptual variations of voice quality (related to manner of vocal fold vibration) that can be maintained over a range of pitch and loudness levels.

3.3.2.1 - Types of voice quality

It is generally agreed that there are three key types of voice quality or registers, in the natural speaking voice, each one defined in terms of a different degree of vocal fold opening and opening-closing relationships and tension during the expiratory airflow, and consequently characterised by different (but somewhat overlapping) fundamental frequency ranges (Hollien, 1983; Seikel, King & Drumright, 1997; Mathieson, 2000; Baken & Orlikoff, 2000).

Laver (1980) considers as the neutral mode of phonation what several others refer to as *'modal voice'*, or *'chest voice'* (Hollien, 1974, 1983; Perkins & Kent, 1986; Baken & Orlikoff, 2000) and as *'habitual pitch'* (Van Riper & Irwin, 1958; Boone, 1982), to designate the range of fundamental frequencies and manner of vocal fold vibration (94-300 Hz) and vocal quality most commonly found in conversation and reading samples, although the possibility of it co-occurring with others does exist (Hollien & Michel, 1968; Carlson, 1995a; Seikel, King & Drumright, 1997; Mathieson, 2000; Baken & Orlikoff, 2000). 'Habitual pitch' to Coleman & Markham (1991) is one of the most common terms used to describe the vocal characteristics of a speaker or group and it is similar to 'average pitch' (Baken, 1996), 'modal fundamental frequency' (Fitch, Holbrook & Tallahassee, 1970) and 'speaking fundamental frequency' (Michel, Hollien & Moore, 1966).

Physiologically, in this mode of phonation the vocal folds vibrate with periodic, and large amplitude of vibration and also complete rapid closure under moderate forces of muscular dynamism. Perceptually, the sound is efficient in terms of utilisation of air, and without audible friction. In terms of acoustics, the generated voice is rich in harmonics, and with efficient energy (loudness) (Laver, 1980; Seikel, King & Drumright, 1997). Nevertheless, difficulties arise in considering habitual pitch variation due to the voice changes to suit the occasion and due to the emotional content of speech that affect F0 (Coleman & Markham, 1991).

Loft voice, also known as falsetto, an old musical term according to Baken & Orlikoff (2000), corresponds to the higher range of fundamental frequencies (276-364 Hz) of the vocal continuum (Hollien & Michel, 1968). Physiologically speaking, and when compared to modal voice, this register reveals not only longer, thinner and stiffer vocal folds but also a small amplitude of vibration, similar open and closing speed phases and maybe with incomplete closure of the folds. Acoustically, it has fewer higher harmonics, and is less efficient at converting the energy from the lungs into sound energy (Perkins & Kent, 1986).

Perkins & Kent (1986) comment that pulse voice or pulsation is essentially synonymous with vocal/ glottal fry (North American usage according to Laver, Hiller & Beck, 1992), creaky voice (British usage according to Laver, Hiller & Beck, 1992) and laryngealization (Broad, 1973). Moore and von Leden (1958 cited by Laver, Hiller & Beck, 1992) called it dicrotic dysphonia. Clinicians, Van Riper and Irwin (1958)

designate the pulse voice as 'basal pitch, 'the lowest note on which (one)...can sustain an utterance'.

Physiologically, it is characterised by short, thick, and lax vocal folds and sometimes by a complex vibratory pattern in which the folds come together twice in each closing phase. Perceptually, this is extremely low in pitch and sounds rough for the majority of the speaking situations, and therefore results perceptually in monotonic voice. Acoustically, it is in a very low frequency register, 20-90Hz, (Broad, 1973; Perkins & Kent, 1986).

Although pulse voice constitutes a normal mode of vibration which can occur in some speaking situations (e.g. at the ends of falling intonation contours in which fundamental frequency of the voice drops rapidly) and conditions (e.g. post-sleep voice) (as already mentioned) its consistent and habitual use is atypical and may be considered misuse of the voice.

3.3.2.2 - Pathological voice types

The traditional model of pathological voice types described by Fairbanks (1960) and Zemlin (1988) accounts for three distinctive perceptual pathological voice quality types, '*breathy*' and '*harsh or rough*' (represent two distinctive voice types) and '*hoarseness*' (a third type, combining the features of the other two) (Pindzola, 1987; Haynes, Pindzola & Emerick, 1992; Wolfe, Fitch & Martin, 1997).

The quality of '*Breathiness*' varies over a wide range that is determined by the ratio of the turbulent noise in the sound caused by inefficient vocal fold closure (due to mass, muscular and neurological conditions) or habitual pattern of speaking (as already mentioned) (Hirano, 1981; Perkins & Kent, 1986; Pindzola, 1987; Haynes, Pindzola & Emerick, 1992; Holmberg et al., 1994; Omori et al., 1997).

There are at least two different possible laryngeal adjustments, which contribute to breathiness. One type manifests vibration of the anterior part of the vocal folds with a permanent posterior glottal chink (cartilaginous glottis) that according to Holmberg et al. (1994) '...results in relatively high levels of energy in the high-frequency regions'. In the other adjustment, there is a longitudinal glottal chink which results in an excessively large glottal area in each cycle, through which large volumes of air flow at low velocities resulting in low levels of high frequency energy (Holmberg et al., 1994).

'*Harshness*' is a quality which results from the irregularity of the vocal fold vibration caused by asymmetrical structural changes of the vocal folds or by atrophy of the vocalis muscle (Perkins & Kent, 1986; Omori et al., 1997). According to Haynes, Pindzola & Emerick (1992) harshness 'is usually perceived in a phonatory milieu of hard glottal attacks, low pitch, intensity problems, and strain or tension in speaking.' Moreover, terms like 'roughness', 'unpleasantness', 'strident', 'coarse', 'grating', 'rasping', 'rough', 'mettalic', and 'guttural' have been used as synonymous with harshness (Pindzola, 1987). Fawcus (2000:57) writes that 'Depending on the physical vulnerability of the vocal folds, such habitual vocal misuse may eventually lead to tissue changes of chronic laryngitis, vocal nodules or contact ulcers.'

'*Hoarseness*' is the auditory effect that results from the simultaneous irregular vibratory patterns (harshness) and from turbulent airflow at the glottis due to incomplete closing phase of the vocal folds vibration (breathiness). The glottal chink may be caused by idiosyncratic laryngeal muscular adjustments or by any mechanical obstruction into the glottis (e.g. nodules). Pindzola (1987:6) writes that in the hoarse voice '...turbulent air flow, rough/aperiodic vibrations, low pitch, and neck muscle strain may be evident'. Although hoarseness can be produced with a normal larynx, it is usually considered to be an early warning sign of laryngeal pathology (Perkins & Kent, 1986; Laver, Hiller & Beck, 1992; Omori et al., 1997).

3.4 - Incidence and prevalence of voice problems

Epidemiological studies of the incidence of dysphonia in the general population are practically unknown, and, as regards the population as a whole, it is not possible to generalise. The criterion used for existing estimates is usually biased by the type of samples used (e.g. mainly those who seek help or specific populations like university students, teachers and singers) or by other factors (e.g. the growing development of voice clinics in past years, a greater public awareness of the importance of voice problems) (Mathieson, 2000).

Nevertheless, several researchers in different countries have published some figures. Verdolini and Ramig (2001) estimate that in the USA from 3 to 9% of the general population has some type of voice problem at any given moment in their life.

Enderby and Philipp (1986) recorded an estimated 28 dysphonias annually per a million people, in a calculation based on five years of referrals for dysphonia to one general English hospital. Gordon (personal communication, 1994 cited by Enderby & Emerson, 1995 and Mathieson, 2000) gives a figure of 89 per 1,000 000 annual referrals for dysphonia in Glasgow. Boone (1983) suggests a prevalence of 3% phonation disorders in the USA and La Guait (1972) found 7.2%, out of his sample of 428 adult patients (Enderby & Emerson, 1995). Verdolini (1994) states that 3% to 6.5% of adults seem to be hoarse. Guimarães (1997) reports 48% of dysphonic (child and adult) patients out of her 228 new patients in a private Portuguese clinic (during a four and a half year period).

Although speculative, it seems that dysphonia correlates with gender. Kinzl, Biebl & Rauchegger (1988) reported a female: male ratio of 22:3 for functional aphonia at the department of Logopedics & Phoniatrics of the University Hospital of Innsbruck while Smith et al. (1996) reported a ratio over 2:1 of women to men seaking treatment for a voice disorder in Iowa and Utah University ENT clinics. Arnoux-Sindt et al. (1994) reported that 83% of the subjects in a survey study with 1200 teachers in Montpellier (France) from 1991 to 1993 were female. Fritzell (1996) states that 66% of the new patients, from a working population, at Swedish hospitals (in a six month period) were female. Guimarães (1997) found that 61.8% of the new dysphonic patients at her private clinic were females (in the period between 1992 to June 1996). Greene (1972 cited by Tucker, 1987) reports that psychogenic dysphonia is seven times more common in women than men. Stemple, Glaze, Gerdeman (1995) stress that the majority of conversion disorders occur in women. A study by Herrington-Hall et al. (1988) investigating laryngeal pathology in a sample of 1,262 individuals seeking treatment, found that the most frequent were females. Verdonck-de Leeuw (1998) reports that in the Netherlands laryngeal cancer occurs mainly in elder males (prevalence of 73 per 1000.00), is rare for females (prevalence of 6 per 1000.000) and 3% of the male laryngeal cancers are under the age of 45 years, 35% is between 45 and 60 years, 49% between 60 and 75 years, and 13% is older than 75 years. Ferreira et al. (1996) found that in two SLT departments of Portuguese Hospitals, during a period of 6 years, the majority (92 speakers) of the 100 psychogenic dysphonics were female. Carvalho et al. (1995) reports that in the Garcia de Orta Hospital, in Portugal, during a period of 3 years, 74% of the patients with vocal fold nodules were female. Moreover, in the same

Portugese hospital, Gomes et al. (2000) report 75% of females with functional dysphonia (without larynx abnormalities) during the period between 1992 and 1999.

While some authors (Vilkman, 2000) raise the possibility that this happens due to anatomophysiological variability (described in detail in chapter II), others (Fritzell, 1996) suggest that women might have a greater need to speak than men.

There are also some reasons to believe that the incidence of dysphonia correlates with occupation. Verdolini & Ramig (2001) indicate an estimate of 25% or more of the working USA population in which the voice plays an important rule. Vilkman (1996) showed that demands made on the voice and its quality had an increased importance on occupation performance in Finland; about one fourth of the total labour force needs to use the voice. Verdolini (1994) estimates that approximately 'one-third of some groups of professional voice users (including public school music teachers and singing voice teachers) may complain of voice problems and /or have some abnormality in the larynx at any given time.' Research carried out at Swedish hospitals (in a six month period), shows that 1212 working patients (out of a total 1484 of patients with voice disorders) represented twenty occupational groups, the largest ones being teachers (197 - in many different areas and at all levels), office workers (183), manual workers (117), students (101), healthcare occupations (92), salespersons (76) and managing directors (69) (Fritzell, 1996).

Published data concerning teachers shows that they constitute a significant proportion of those who seek help for voice problems (Herrington-Hall et al., 1988; Sapir, Keidar & Mathers-Schmidt, 1993; Arnoux-Sindt et al., 1994; Fritzell, 1996; Smith et al., 1996; Vilkman, 1996, 2000; Morton & Watson, 1998; Simberg et al., 2000). Smith et al. (1996) state that in Iowa and Utah within the working population seeking medical care, teachers comprise about 1%. Vilkman (1996) states that from questionnaire surveys it is known that 20-30% of teachers do not present with voice complaints, more than 10% reported one or more symptoms weekly and as many as 50% to 80% of the teachers occasionally experienced voice problems. Klingholz et al. (1978 cited by Vilkman et al., 1998) refer to a nearly 100% prevalence of voice complaints among kindergarten teachers. Newman & Kersner (1998) have identified in a questionnaire survey that 52% of the female aerobics instructors reported having voice

problems and 27% currently had a voice problem. Unfortunately, epidemiological information on the non-clinical teacher population is scarce (Arnoux-Sindt et al., 1994).

Many studies have focused on the relationship between laryngeal pathologies and dysphonia and gender, occupation and lifestyle. Fritzell (1996) reports that some laryngeal pathologies are more common among some occupational groups (e.g. vocal fold oedema in office workers, nodules in students and chronic laryngitis in patients who are manual workers). Forty seven subjects out of 226 students in Finland who underwent laryngoscopic examination showed some pathology: the most prevalent was laryngitis (60%) followed by nodules (21%), functional disorders (without organic lesion, 11%), minor findings (not specified, 6%) and polyps (2%) (Simberg et al., 2000). Carvalho et al. (1995) refer that 43% of the 133 patients observed in the ENT department of Garcia de Orta Hospital, in Portugal, during a period of three years, presented vocal fold nodules.

3.5 - Vocal fundamental frequency

The listener's judgment of pitch, widespread in the clinical field is an important factor in diagnosis and therapy; hence it is crucial that fundamental frequency, the major physical correlate of perceived pitch be measured. In fact, fundamental frequency is one of the most reported acoustic parameters in voice research, mainly because is the single most important acoustic cue for intonation (Atkinson, 1976; Graddol & Swann, 1983) and also because there is a need to develop 'objective' outcomes for dysphonia (Krook, 1988; Murry, Brown & Morris, 1995; Bough et al., 1996; Rantala, Määttä & Vilkman, 1997; Ryan, 1998; Baken & Orlikoff, 2000).

Vocal fundamental frequency (F0) or speaking fundamental frequency (SFF) is the term used to refer to the physical sound produced by the periodic vibration of the vocal folds per unit time in sustained phonation and connected speech respectively. F0 reflects the competency of the phonatory system, larynx biomechanics and the interaction with aerodynamics (this is described in detail in the chapter II).

Some authors (Boone, 1982) use the cycles per second (cps) as a frequency unit but the Hertz (Hz) or semitones (ST) are nowadays universally adopted; however the

use of the semitone scale could be considered preferable to Hz because it is associated with human perception (Baken & Orlikoff, 2000).

There is a huge body of research devoted to normative data for fundamental frequency (F0) ranging from the newborn to the elderly, showing variability associated with age, gender, race, time of the day, using a variety of speech tasks (sustained vowels, reading, conversation, singing, counting), different languages, studying the effects of behavioural factors (e.g. drinking and smoking), occupational voice use (e.g. speaking and singing) and dysphonia, using different data gathering and analysis systems (e.g. acoustic and electrolaryngographic), inter and intraspeaker variability (Atkinson, 1976; Linville, 1988; Bough et al., 1996) and reporting different F0 measures (e.g. mean, mode, range). Nevertheless, there are some areas where reasonable agreement has been reached (e.g. as regards the speakers' age and gender) (Mysak, 1959; Hollien & Shipp, 1972; Stoicheff, 1981; Krook, 1988) in others, however, controversy still exists (e.g., language, education, type of speech samples used, special behaviours such as smoking) (Murry, Brown & Morris, 1995; Hollien, Hollien & De Jong, 1997). As would be expected, comparability between studies is often problematic because of differences in sample analysis methods, criteria for subject selection, recording and analysis techniques, the number of subjects and reported measurements (Krook, 1988; Carlson, 1995a; Murry, Brown & Morris, 1995; Hollien, Hollien & De Jong, 1997; Baken and Orlikoff, 2000).

3.5.1 - Sampling criterion for voice measurement

Attempts to develop normative data for fundamental frequency (using measures such as mean, mode, median, standard deviation and mid-90% range to characterise it) are highly reported in the voice literature; however, some methodological issues remain unresolved. For example, speech samples often differ from one study to another, in type (sustained phonation, oral reading or speaking) and length (time or sentence unit) with necessarily consequences in the validity and reliability of the measurement (e.g. the magnitude of associated sampling error) (Horii, 1975; Titze, Horii & Scherer, 1987; Karnell, 1991; Scherer, Vail & Guo, 1995).

Traditionally, sustained vowels have been chosen as the speech sample to obtain data on voice quality and the principle reason given was the limitation of the equipment

used (Verdonck-de Leeuw, 1998; Askenfelt and Hammarberg, 1986; Yiu et al., 2000). Still, other reasons involve the fact that sustained vowels are more or less 'stable' and do not contain intonational changes and effects of coarticulation, they may therefore, be considered advantageous for 'jitter' measurement and may be more advantageous in a practical sense (e.g. typically easily understood by the subject and easy to imitate) (Askenfelt and Hammarberg, 1986; Laver, Hiller & Beck, 1992; Scherer, Vail & Guo, 1995). However, vowels have the major disadvantage of not being as representative of speech communication (Klingholtz, 1990; Yiu et al., 2000; Fourcin, 2000) as reading or spontaneous speech tasks and may mask the effects of dysphonia because of its restrictive production (Laver, Hiller & Beck, 1992).

Continuous speech has the advantage of more realistic data as it represents the result of the dynamic voice action with multiple onsets and offsets. However, it has more artifactual distortion possibilities (e.g. emotional features) and implies more sophisticated analysis techniques (Klingholtz, 1990; Laver, Hiller & Beck, 1992).

Fitch (1990) comments that oral reading has high test-retest consistency, and the obtained F0 values result were similar to those for spontaneous speech. Also, Baken & Orlikoff (2000) suggest the use of reading for F0 measurements because its content may be compared and it is a repeatable task.

An inspection of the published research shows little agreement in the length magnitude of the analysis window to be used in the F0 perturbation measurement (Horii, 1975; Titze, Horii & Scherer, 1987; Deem et al., 1989; Karnell, 1991; Scherer, Vail & Guo, 1997).

Horii (1975) studied the magnitude of sampling errors, for oral reading (Rainbow Passage), associated with two different measures, fixed time windows and sentence unit of sampling. In 10 'normal' adults (six males and four females) the variability of the F0 measures (mean, standard deviation, mid-90% range) decreases as the sampling size (time) increases. Using the first paragraph of the Rainbow Passage (to test the sentence unit as a sampling) by 65 adult males (mean age 54.1 years) higher correlation coefficients and smaller errors of F0 estimates were found the longer the sentences were. Moreover, correlations tend to be higher for middle sentences than initial and final sentences.

Hiller (1985 cited by Laver, Hiller & Beck, 1992) found that a 40-second sample of reading provided relatively stable long-term speaker-characterizing parameters of perturbation for healthy male and female speakers.

Titze, Horii & Scherer (1987) recommended as many as 30 cycles length window for a reliable analysis of jitter and shimmer in 'normal' speakers. However, Deem et al., (1989) refer that a sample duration of 40 cycles or greater yielded more consistent measures (jitter) than sample durations of 10 or 20 cycles.

Karnell (1991) indicates as many as 190 cycles for an accurate minimum value of jitter in dysphonic adult subjects, balanced for sex and severity of hoarseness (no hoarseness to moderate hoarseness) although the use of 110 cycles would be adequate. This length analysis window is similar to the suggested window length (100 cycles) by Horii (1979) and Scherer, Vail & Guo (1997) for most subjects.

3.5.2 - Fundamental frequency according to age and gender

A large body of research has been devoted to an examination of age and gender-related changes in voice and relative consensus about fundamental frequency (F0) level variation through the life span for both male and females has been reached (Mysak, 1959; Hollien & Shipp, 1972; Linke, 1973; Krook, 1988; Fitch, 1990; Higgins and Saxman, 1991; Shipp et al., 1992; Russel, Penny & Pemberton, 1995; Linville, 1996; Hollien, Hollien & De Jong, 1997).

There is a high degree of consensus that the female F0 is generally higher than male F0 although reported data show considerable discrepancy among reported values. Linke (1973) has pointed out that the pitch level of the female voice is somewhat less than one octave higher than that of the male speaking voice and that this tends to confirm what was found by Cowan (1936) and Snidecor (1940).

The general trends of F0 versus age and gender may be seen in Figure III.1. Additionally, specific values for reading and spontaneous speech according to age and gender for adults are presented from table III.9 to III.14. Moreover, Krook (1988);

Carlson (1995a) and Russell, Penny & Pemberton (1995) present extensive lists of previous research on F0 of adults according to age and gender.

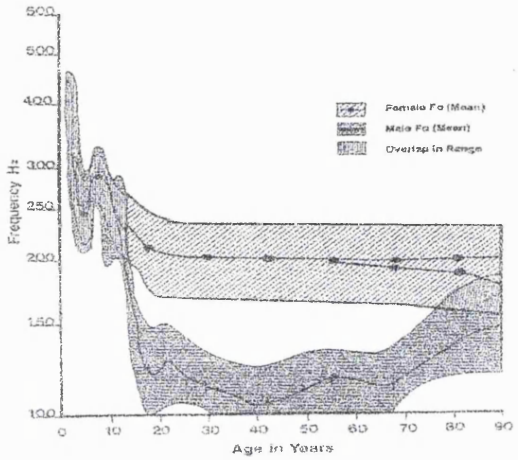


Figure III.1 - the general trends of F0 across the life-span
From: Hollien (2000:20)

Investigation of the various acoustic aspects of voice production associated with aging shows that, although both genders demonstrate a decrease of F0 mean level with increasing age and then a rising trend again with the onset of old age a consensus about the magnitude of the differences and the corresponding ages was still not achieved (Honjo & Isshiki, 1980; Linville & Korabic, 1987; Krook, 1988; Biever & Bless, 1989; Russel, Penny & Pemberton, 1995).

For male voices the published data shows a progressive lowering of mean F0 in middle age after which a change in the trend occurs with a rising of the mean F0 during senescence (Mysak, 1959; McGlone and Hollien, 1963; Fitch, Holbrook & Tallahassee, 1970; Hollien & Shipp, 1972; Krook, 1988; Higgins & Saxman, 1991). For example, Shipp et al. (1992) found a significant different mean F0 and standard deviation during reading between a group of young male speakers (mean age 25.3 years) and old males (mean age 83.7 years) and between the middle-age speakers (mean age 57.7 years) and the old subjects. However, no differences were found between young and middle-age speakers. Nevertheless young speakers presented higher F0 means (120.67Hz, s.d.=10.87) than middle-age (106.22Hz, s.d.=2.27) and lower than old speakers (149.23Hz, s.d.=19.97). Moreover, Higgins & Saxman (1991) reported a slightly, but

not significantly, higher F0 in the elderly (aged 75.3, s.d.=4.9) than in young (aged 24.1, s.d.=2.8) males during normal sustained vowel [æ] and the syllable [bæp].
According to investigators this phenomenon (increasing F0 with ageing) is presumably associated with the body and larynx aging process (e.g. vocal fold atrophy and tissue stiffening) (described in detail in chapter II).

For female voices a reported statistically significant lowering of average fundamental frequency through adulthood was found (Honjo and Isshiki, 1980; Stoicheff, 1981; Pinto and Hollien, 1982; Krook, 1988; Higgins and Saxman, 1991; Russell, Penny & Pemberton, 1995) in contrast with reports of minimal change with no statistical significance (Biever and Bless, 1989; Hollien, 1987; McGlone & Hollien, 1963; Yamazawa & Hollien, 1992).

Moreover, some researchers (McGlone & Hollien, 1963; Saxman and Burk, 1967; Gilbert and Weismer, 1974; Stoicheff, 1981; Krook, 1988; Biever & Bless, 1989; Russell, Penny & Pemberton, 1995) reported an F0 increase again during the seventh and eighth decades with stability in these two decades while others found the reverse (Stoicheff, 1981; Brown, Morris & Michel, 1989).

Stoicheff (1981) found that when the smoking variable is controlled the F0 in females only lowers with the completion of the menopause and is maintained in old age. Brown, Morris & Michel (1989) found significantly lower F0 for a group of 25 women aged 75-90 (mean 79.4 years) when compared to a group of 25 young women aged 20-32 (mean, 27.5 years).

Also, research into age-related changes in F0 suggests that elderly speakers present more intraspeaker variability in F0 stability in comparison to young speakers (Stoicheff, 1981; Linville, 1988). Still, data on F0 of females' voices are somehow conflicting and do not show a clear pattern with aging (especially in the elderly females) (Stoicheff, 1981; Hollien, Hollien & De Jong, 1997). It is not yet known if the above described F0 discrepancies are random in nature or follow specific trends justified by the male-female coalescence model (described in chapter II) (Hoonjo and Isshiki, 1980; Stoicheff, 1981; Krook, 1988; Greene & Mathieson, 1989; Yamazawa & Hollien, 1992; Hollien, Hollien & De Jong, 1997).

A further comparison between the genders' pitch usage characteristics shows that during reading the reported 90% F0 range is 9.5 semitones for 15 males with a mean age of 47.9 years (Mysak, 1959) and 7.95 semitones for 65 males with a mean age of 54.1 years (Horii, 1975) whereas for 27 university women the reported value was 9.32 semitones (Linke, 1973). During conversation Mysak (1959) found a 90% F0 range of 9.4 semitones for 15 males with a mean age of 47.9 years. Ramig & Ringel (1983) note that F0 variability between reading and speech is not large. Moreover, studies show a tendency toward greater pitch variability as a function of age (Mysak, 1959; Linke, 1973; Stoicheff, 1981; Linville, 1988).

3.5.3 - Speech material influences on fundamental frequency

Normative data on adults is extensively available for sustained vowel phonation (Bless et al., 1993; Gelfer, 1995; Martin, Fitch & Wolfe, 1995; Wolfe, Cornell & Fitch, 1995; Kingston et al, 1997; Higgins, Netsell & Sculte, 1998; Whalen et al., 1998; Debruyne & Decoster, 1999; Yiu et al., 2000), and for vowels in word context (Zawadzki & Gilbert, 1989; Fischer-Jorgensen, 1990; Hillenbrand et al., 1995) and shows that regular differences in F0 occur according to the vowel type produced.

Several investigators (Viana, 1987; Whalen & Levitt, 1995; Kingston et al., 1997; Whalen et al., 1998; Fischer-Jorgensen, 1990) refer to the universality of intrinsic F0 of vowels for a great majority of the reported languages (e.g. English, Portuguese, Dutch, Italian, Finnish, German, French, and Southeast Asian) but not all (e.g. East and West African). This means a correlation between F0 and the tongue root position in the type of vowel, where F0 is higher for [+high] vowels such as [i] and [u] than for [+low] vowels such as [ɑ].

Data on the mean F0 for sustained vowels have been suggested in several studies and are presented in Table III.9 for adult females and in Table III.10 for men.

Table III.9 - Intrinsic mean F0 for adult females

N	Age Average/ range	[u] Hz	[i] Hz	[a] Hz	Source
10	(16-30)	------	212 ± 27 range 81-263	-----	Peppard, Bless & Milenkovic (1988)*
10	25.7± 3.9 (22-34)	231	228	215	Sussman & Sapienza (1994)*
20	37.1 (27-49.11)	204.60	205.5	198.75	Sorensen & Horii (1982)*
13	51. 6 ± 22.2	-------	-----	206.3 ± 37.9	Hertrich & Ackerman (1995)**
3	adult	253	250	212	Peterson (1961 cited by Whalen & Levitt, 1995)*

Language: *American English; **German

Table III.10 - Intrinsic mean F0 for adult males

N	Age average/ range	[u] Hz	[i] Hz	[a] Hz	Source
10	24 ± 2.7 (19-28)	128	122	115	Sussman & Sapienza (1994)*
20	27.5 (22.4-37.3)	128.8 ± 17.7	128.5 ± 18.5	125 ± 17.6 [ɑ]	Pinto & Hollien (1982)*
20	39.7 (29.1-49.11)	123.20	125.60	110.85	Sorensen & Horii (1982)*
12	48.4 ± 16.5	--------	--------	125.4 ± 17.3	Hertrich & Ackerman (1995)**
4	adult	128	124	119	Peterson (1961 cited by Whalen & Levitt, 1995)*

Language: *American English; **German

Explanations for this universal fact have been given on the basis of a theoretical articulatory-based framework in which there is a covariation between the phonatory (source) and the articulatory (filter) systems in the production of vowels. This means the F0 varies with tongue root position because this structure is linked by cartilages and ligaments to the larynx inlet, although creating increased external tension in the larynx. Changing the position of the larynx by either raising it (as during the production of vowel [u]), by pulling it in an anterior-vertical direction (as in vowel [i]), or by lowering it (as in vowel [ɑ]), alters the position of the laryngeal cartilages and consequently the tension of the vocal folds. Therefore, it would seem reasonable that perceptually tense (high) and lax (low) vowels will have different intrinsic fundamental frequencies (Sapir,

1989; Sussman & Sapienza, 1994; Gelfer, 1995; Kingston et al, 1997; Higgins, Netsell & Schulte, 1998; Whalen et al., 1998).

Nonetheless, Zawadzki & Gilbert (1989) claim that mandible position is more closely related to F0 than tongue position. Whalen & Levitt (1995) notes that Steele (1986) stresses the importance of subglottal pressure and others (Diehl & Kluender, 1989; Diehl, 1991 and Kingston, 1993) suggest that it is a deliberate change to perceptually differentiate vowel categories.

Whalen & Levitt (1995) present an extensive analysis of the published data values for intrinsic F0 in [u], [i] and [ɑ] vowels in 32 different languages and in different contexts (e.g. isolated vowels and in words). Their analyses of those studies show a highly significant intrinsic F0 vowel effect ($p < 0.0001$) across all the languages and a high F0 variability between the front and back vowels in some languages. The authors explain the F0 variability as a result of the inevitable sampling error.

Also, Ewan (1975 cited by Fischer-Jorgensen, 1990) notes that [u] generally has a slightly higher F0 than [i].

Additionally, several other researchers (Fitch, 1990; Higgins, Netsell & Schulte, 1994; Murry, Brown & Morris, 1995) have reported that the variability of F0 outcomes is not only dependent on the vowel type but also on the different speech samples used.

Fitch (1990) noted a relatively high F0 in sustained vowels, compared to that in reading (Rainbow Passage) and spontaneous speech in 12 young adults of both genders. He also stated that the most consistent F0 measurements came from reading. Also, Murry & Brown (1982 cited by Ryan, 1998) indicated significantly higher mean F0 for the sustained vowels than that for the reading and speaking whereas no significant difference was found between reading and conversation in 16 female college students.

Murry, Brown & Morris (1995) present a possible explanation for this fact: i.e. the relative static laryngeal adjustment during sustained phonation or the subject's attitude when asked to produce sustained phonation.

Another issue, namely whether or not oral reading will necessarily result in different F0 levels from those present in normal spontaneous speech, has been the subject of much research. Studies found a consistently slightly higher F0 during reading than spontaneous speech (Snidecor, 1943; Mysak, 1959; Saxman & Burk, 1967; Hollien & Jackson, 1973; Brown & Hollien, 1981; Ramig & Ringel, 1983; Fitch 1990; Higgins and Saxman, 1991; Drew & Sapir, 1995; Hollien, Hollien & De Jong, 1997).

Mysak (1959) noted that oral reading was characterised by a slightly but consistently higher mean F0 level than conversation in 39 male subjects divided in three groups with 47.9; 73,3 and 85 mean ages as did Fitch (1990) for 12 young adults (21 to 26 years old). Brown & Hollien (1981) found that 16 females (aged from 19 to 22 years) used a slightly higher F0 during reading (mean= 207.5, standard deviation=16.5) than during speaking (mean= 203.4, standard deviation=16).

Hollien, Hollien & De Jong (1997) found that the mean fundamental frequency for reading aloud was always higher (between 0.6 st. to 0.9 st.) than that for spontaneous speaking. The study involved 341 young, healthy subjects under precise conditions. However, the authors stress that one subject in five exhibited either no difference at all or a higher F0 for the spontaneous speech when compared to reading aloud. The possible reason presented for this bias was the individual differences such as stress levels experienced by some subjects (e.g. some subjects responded with a neutral attitude while others where more enthusiastic).

Conversely, Britto & Doyle (1990) obtained a lower F0 from reading tasks versus spontaneous monologue in 40 healthy Canadian speakers (20 women and 20 mean with an age range between 20 and 30 years old and with English as their native language). Data is shown between table III.11 and table III.14. However, the mean F0 for the sustained vowel [ɑ] was higher than that of connected speech for both, men (mean=124.4 Hz, s.d.= 17.8) and women (mean=218.4 Hz, s.d.= 27.3).

Data for mean F0 and standard deviation during reading and conversation for female and male speakers is shown between table III.11 and table III.14

Table III.11- Mean F0 for adult females during reading

Mean age	Age Range	N	Mean F0 (Hz)	SD F0	Stimulus	Language	Source
	18-19	10	241.6		Sentence	Brit. English	Abberton (1976)
	18-25	66	224	1.7 ST	Standard sentences (72 words).	Austr.Engl.	Pinto & Hollien (1982)
	20-29	35	195.5	25.51 Hz	Standard sentences (99 words).	Swedish	Krook (1988)
24.6	20-29	21	224.3	3.78 ST	1st paragraph of the RP	Am English	Stoicheff (1981)
	21-26	6	206.4	15.3 Hz	Middle 55 words 'RP' (Rainbow Passage)	Am English.	Fitch (1990)
22.5	21-30	18	198	1.2 T	Rainbow Passage	Am. English.	Hollien, Hollien & De Jong (1997)
23.6	20-30	20	198.6	19.9 Hz	2nd sentence 'Rainbow Passage'	Am. English	Britto & Doyle (1990)
	21-32	10	197	1.4 ST	4 & 5th sentences of 1st paragraph 'RP'	Am. English	Benjamin, 1981 (Russell, Penny and Pemberton, 1995)
26	----	23	202	20.3 ST	Passage (9 seconds)	Am. English	Stone & Rainey (1991)
27.5	20-32	20	192.4	2 ST	2 1st paragraph 'RP'	Am English	Brown, Morris & Michel (1989)
	30-39	100	194.7	18.84 Hz	Standard paragraph (99 words).	Swedish	Krook (1988)
35.4	30-39	18	213.3	3.92 ST	1st paragraph of the RP	Am English	Stoicheff (1981)
33.5	30-40	9	196.3	1.2 ST	1st paragraph 'Rainbow Passage'	Am.Engl.	Saxman & Burk (1967)
37.1	27-49.11	20	188.35		1st paragraph 'Rainbow Passage'	Am. English	Sorensen & Horii (1982)
	40-50	9	181.6	1.4 ST	1st paragraph 'Rainbow Passage'		Saxman & Burk (1967)
46.4	40-49	21	220.8	4 ST	1st paragraph 'Rainbow Passage'		Stoicheff (1981)
	40-49	83	190.9	19.47 Hz	Standard paragraph (99 words).	Swedish	Krook (1988)
	45-50	6	180		Sentence	Brit. English	Abberton (1976)
52	45-61	17	196	------	Swedish text (~ 35 sec.)	Swedish	Södersten, Hertegård & Hammarberg (1995)
54.5	50-59	17	199.3	4.33 ST	1st paragraph 'Rainbow Passage'	Am English	Stoicheff (1981)
	52-60	11	180	1.7 ST	Standard sentences (72 words).	Aust. Engl.	Pinto & Hollien (1982)
	50-59	83	182.1	18.99 Hz	Standard paragraph (99 words).	Swedish	Krook (1988)
	60-69	85	180.9	24.20 Hz	Standard paragraph (99 words).	Swedish	Krook (1988)
65.8	60-69	15	199.7	4.25 ST	1st paragraph 'Rainbow Passage'	Am English	Stoicheff (1981)
66.3	65-68	15	181.2	17.5 Hz	Standard sentences (45 words)	Aust. Engl	Russell, Penny and Pemberton (1995)

Table III.12 - Mean F0 for adult males during reading

Mean age	Age Range	N	Mean F0 (Hz)	SD F0 (ST)	Stimulus	Language	Source
20.3	18-26	157	129	1.6 T	A passage from R.L. Stevenson	Am English	Hollien & Jackson (1973)
20.5	18-25	142	122	1.6 T	A passage from R.L. Stevenson	Am English	Hollien & Jackson (1973)
	21-26	6	112.9	13.6 Hz	Middle 55 words of the Rainbow Passage	Am English	Fitch (1990)
	20-29	7	112.3	7.63 Hz	Standard sentences (99 words).	Swedish	Krook (1988)
23.4	21-30	25	114	1.3 T	The Rainbow Passage	Am English	Hollien, Hollien & De Jong (1997)
24.4	20-29	25	119.5		1st paragraph 'Rainbow Passage'	Am English	Hollien & Shipp (1972)
24.6	20-30	20	114.6	11.3 Hz	2nd sentence 'Rainbow Passage'	Am.English	Britto & Doyle (1990)
25.3	21-35	10	120.67	10.87 Hz	3rd sentence 'Rainbow Passage'	Am. English	Shipp et al. (1992)
34.9	30-39	25	112.2		1st paragraph 'Rainbow Passage'	Am English	Hollien & Shipp (1972)
	30-39	20	109.6	12.41 Hz	Standard sentences (99 words)*.	Swedish	Krook (1988)
39.7	29.1-49.11	20	120.65		1st paragraph 'Rainbow Passage' (98 words)	Am. English	Sorensen & Horii (1982)
47.9	32-63	15	113.2	2.9 ST	1st paragraph 'Rainbow Passage' (98 words)	Am. English	Mysak (1959)
	40-49	12	107.6	14.32 Hz	Standard sentences (99 words).	Swedish	Krook (1988)
45.4	40-49	25	107.1		1st paragraph 'Rainbow Passage'	Am English	Hollien & Shipp (1972)
	50-59	46	119.5	21.49 Hz	Standard sentences (99 words).	Swedish	Krook (1988)
52.5	SD-11.9	20	121.9	4.63 ST	3rd sentence of the 1st paragraph 'Rainbow Passage	Am English	Murry (1978)
54.3	50-59	25	118.4		1st paragraph 'Rainbow Passage'	Am English	Hollien & Shipp (1972)
57.5	46-71	10	106.22	2.27 Hz	3rd sentence 'Rainbow Passage'	Am. English	Shipp et al. (1992)
	60-69	77	118.5	22.51 Hz	Standard sentences (99 words).	Swedish	Krook (1988)
64.6	60-69	25	112.2		1st paragraph 'Rainbow Passage'	Am English	Hollien & Shipp (1972)

Table III.13 - Mean F0 for adult females during conversation

Mean age	Age Range	N	Mean F0	SD F0	Duration	Language	Source
	19-24	10	207			British English	Barry et al. (1990 cited by Carlson, 1995a)
	21-26	6	210.3	18.6 Hz	3 minutes	American English	Fitch (1990)
22.5	21-30	18	190	1.1 T	40 seconds	Am English	Hollien, Hollien & De Jong (1997)
23.6	20-30	20	199	16.8 Hz	30 seconds	Am. English	Britto & Doyle (1990)
37.1	27-49.11	20	186.45	-----	30 seconds	Am. English	Sorensen & Horii (1982)

Table III.14 - Mean F0 for adult males during conversation

Mean age	Age Range	N	Mean F0 (Hz)	SD F0	Stimulus	Language	Source
20.3	17.9-25.8	157	123	1.6 T	2 minutes	Am. English	Hollien & Jackson, (1973)
20.5	18-25	142	116	1.7 T	2 minutes	Am. English	Hollien & Jackson (1973)
23.4	21-30	25	110	1.3 T	40 seconds	Am. English	Hollien, Hollien & De Jong (1997)
24.6	20-30	20	115.9	14.3 Hz	30 seconds	Am. English	Britto & Doyle (1990)
	21-26	6	109.1	13.2 Hz	3 minutes	Am. English	Fitch (1990)
39.7	29.1-49.11	20	115.95	-----	30 seconds	Am English	Sorensen & Horii (1982)
47.9	32-62	15	110.7	2.9 ST	Not specified	Am English	Mysak (1959)

3.5.4 - Cultural influences on fundamental frequency

There is some controversial evidence, which suggests that F0 is culturally influenced by factors such as race, language, accent and level of education. While some studies (Majewski, Hollien & Zalewski, 1972; Loveday, 1981; Hudson & Holbrook, 1982; Pinto & Hollien, 1982; Yamazawa & Hollien, 1992) found some differences in fundamental frequency between speakers from different race, language, accent and level of education others have shown no difference (Hollien, Hollien & De Jong, 1997; Millar & Wagner, 1983).

Majewski, Hollien & Zalewski (1972) found a different F0 average between comparable samples of Polish and American males (respectively 118.9 Hz and 137.6 Hz), as did Loveday (1981) in his study between Japanese and British English female and male speakers. However, it should be noted that Loveday's sample was too small (10 speakers in total).

Yamazawa & Hollien (1992) reported a statistically significant higher F0, for 32 young Japanese women (during the reading of a 88 word Japanese passage) than for a matched group of 24 young American women (during the reading of the first paragraph of the Rainbow passage, 99 words). While it may be possible that these differences are racially related it is probable that results also from other factors such as differences between the two languages or social behaviour. On the other hand, data on Japanese F0 characteristics, according to Yamazawa & Hollien (1992), is consistent with the male-female coalescence theory of vocal aging.

Millar & Wagner (1983) conducted a pilot study with 10 English speakers (5 speakers of general or cultivated Australian English and 5 speakers of Mid-Western or New England English). Although a considerable variation in intonation patterns and rate of speech was reported (during the reading of six sentences), no obvious difference, however, was found between the varieties of English that could differentially affect the measurement of fundamental frequency.

Pinto and Hollien (1982) found higher mean F0 of 66 Australian women aged 18-25 years (224 Hz) than that of North American women of approximately the same age

(200Hz), however, the mean F0 was calculated from recordings made in 1945 and 1950 which may be a bias in their research design.

Hudson & Holbrook (1981) analysed 100 American black men and 100 black women ranging from 18 to 29 years and found that both genders had a lower F0 mode frequency (110.15Hz for males and 193.10Hz for females) compared to the published data for white adults during the reading of the Rainbow Passage (116.65Hz for males and 217Hz for females). Morevover, black subjects in general present a lower modal fundamental frequency range, and standard deviations than the similar white subjects group.

Some researchers suggest that F0 results gathered from university students' samples are somehow biased because students have uncommon characteristics compared to the general population (e.g. high levels of language skill and/or intellectual ability) (Hollien, Hollien & De Jong, 1997).

Hollien, Hollien & De Jong (1997) found a significant statistical difference (over a semitone) for F0 level, during reading, between 157 healthy young adult university students (mean F0=129Hz) and 142 military males (F0=121.5Hz), all displaying identical characteristics, except for the variable which the study dealt with (level of education). However, the authors stress that fact that the military group had a slightly greater proportion of smokers. Although they are young and not long-term smokers the effects of this variable cannot be ignored.

Possibly, the differences found in the above mentioned studies are biased not only by methodological design differences but also by the racial, cultural and linguistic confounding relationships or other unknown variables.

3.5.5 - Variation in F0 during the day

Although variation in F0 during the day has been reported for both women and men by Brown, Murry and Hughes (1976) and just for men by Garrett and Healey (1987) and by Higgins, Netsell & Schulte (1994), Nittrouer et al. (1990) reported the reverse, that is to say, they noted no variation of F0 during the day for men and women

for voice samples obtained in the morning compared to those obtained in the early afternoon (Murry, Brown & Morris, 1995).

3.5.6 - Occupational risk factors: effects on fundamental frequency

The effects of occupational voice use on fundamental frequency have been the subject of research (Stemple, Stanley & Lee, 1995; Buekers et al., 1995; Rantala, Määttä & Vilkman, 1997; Rantala, Lindholm & Vilkman, 1998; Vilkman et al., 1998) from the point of view of vocal loading and environment conditions (working places or laboratory field). However, results are partly contradictory; some showed a rise in fundamental frequency (Gelfer, Andrews & Schmidt, 1991; Stemple, Stanley & Lee, 1995; Rantala, Lindholm & Vilkman, 1998) while others failed to show any specific trend in comparison to a healthy group of subjects not exposed to adverse environmental conditions (Buekers, 1998a).

Gelfer, Andrews & Schmidt (1991) found a more consistent pattern in the acoustic measures (F0, intensity, jitter ratio and signal-to-noise ratio) obtained pre- and post-reading for one hour of loud and constant reading in a group of trained female singers (trained and untrained) than in a group of untrained female singers.

Stemple, Stanley & Lee (1995) studied the effects of vocal loading (reading for a period of 2 hours at 75-80 db of intensity) in ten women without voice problems (mean age 25.3 years) under controlled conditions (e.g. hydration, voice use and nonsmokers). The major findings were a significant F0 increase during reading, laryngoscopic evidence of anterior gap in six of the ten women (which was not observed in the pre-examination) and jitter decrease during high-pitched vowel production.

In a retrospective study with 88 adult patients complaining of voice fatigue Eustace, Stemple & Lee (1996) found that data for F0 and jitter (in vowels and reading) were in agreement with the published data for 'normal' subjects. However, 61% of the patients presented glottal gaps (37% of the women and 28% of the men demonstrated an anterior glottal gap; 3% presented with anterior and posterior gap, and 25% presented a spindle-shaped glottal closure).

Buekers et al. (1998a) in a study with EGG measurements did not find any significant evidence of vocal fatigue after a working day and in a laboratory test of vocal loading comparing 20 women with history of voice fatigue complaint and 12 healthy women.

Rantala, Lindholm & Vilkman (1998) found a F0 increase for three female teachers (mean age 34 years) with no laryngeal pathology after loading under laboratory and field conditions. Moreover, F0 was higher under field condition than in the laboratory.

Discrepancies found are explained by some (Vilkman, 1996; Rantala, Määttä & Vilkman, 1997; Rantala, Lindholm & Vilkman, 1998) on the basis of methodological issues such as the condition of data gathering (laboratory settings or real life situations), duration of the loading, speech task samples (sustained phonation or connected speech), subjects' characteristics and analysis criteria.

Rantala, Määttä & Vilkman (1997) present a brief analysis of some of the published studies on vocal loading.

3.5.7 – Effects of cigarette smoking on fundamental frequency

It has been long known that smoking has a detrimental effect on the vocal fold mucosal and a broad effect on phonation (Gilbert and Weismer, 1974; Stoicheff, 1981; Sorensen & Horii, 1982; Murphy & Doyle, 1987; Sataloff, Spiegel & Hawshaw, 1995; Stemple, Glaze & Gerdeman, 1995; Colton & Casper, 1996; Dworkin & Meleca, 1997).

In a study with smokers and nonsmokers ranging in age from 30 to 54 years Gilbert and Weismer (1974) found that smokers had a significant lower F0 (163.75Hz) than nonsmokers (182.8 Hz) during a reading passage. However, in spontaneous speech although the same trend was evident it was not statistically significant. Vocal fold thickening in 87% of smokers as compared with 7% of nonsmokers was also observed.

Stoicheff (1981) found higher F0 values in 111 nonsmoking female adults in comparison to the published data with similar age and she considers that those differences are due to the control of the smoking variable. However, some methological

limitations were found that may limit the validity of the findings such as: data was obtained from several occupational groups (e.g. secretaries, housewives, nurses and students) and the subjects were instructed to read the passage as though to an audience of 25 persons.

Sorensen & Horii (1982) found that a group of 20 smoker females and 20 smoker male subjects showed lower F0 during sustained phonation with the vowels ([i], [a], [u]), oral reading and spontaneous speech (description of a picture) than the other two equally sized groups of nonsmokers. However, the differences were only statistically significant for the male subjects.

Murphy & Doyle (1987) found a F0 rise (about 18 Hz; 1.63 ST) for two smokers after 40-hours of non-smoking period whilst the two control age and gender matched non smoker speakers did not show a rise in their F0 for all the speech tasks used (sustained vowels, reading and speech). However, Coleman & Markham (1991) postulated that such changes might be attributable to intraspeaker variability over repeated trials, rather than the cessation of smoking.

Commins (1988 cited by Carlson, 1995a) found a highly positive correlation between the levels of carbon monoxide and F0 variance with the heaviest smokers showing greater F0 variability.

Rodrigues (1998) compared 30 Portuguese smoking female students with 30 non-smoking female students of similar age (mean= 20.57 and 20.83 respectively). The smoking habit mean was 5.37 years (s.d.=2.28) and the mean number of cigarettes smoked per day was 9.87 (s.d.=4.31). The acoustic results showed that, for the sustained vowel [ɛ] smokers showed statistically significantly lower F0 (206.78 Hz) and jitter (0.33%) values than non-smokers (F0=217.99 Hz and jitter=0.21%). Although control of variables was taken into consideration such as non-professional voice use, and no respiratory and voice problems, some limitations should be considered (e.g. laryngostroboscopic examination was not performed, phase of the menstrual cycle was not evaluated, recordings were performed in different parts of the day).

3.5.8 – Alcohol effect on fundamental frequency

Despite the increasing role played by alcohol in our society only limited meaningful research exploring its effects on speech and voice has been reported (Trojan & Kryspin-Exner, 1968; Sobell & Sobell, 1972; Sobell, Sobell & Coleman, 1982; Klingholz, Penning & Liebhardt, 1988).

Trojan & Kryspin-Exner (1968) reported changes on pitch as an effect of alcohol in 3 male subjects (mean age, 31 years).

Sobell, Sobell & Coleman (1982) revealed no significant effects of moderate versus high alcohol levels on the acoustic F0 measurement during reading for 16 young males. According to the authors the results may be due to the lack of sensitivity of the method employed. It should also be mentioned that subjects showed a high variance of alcohol blood level concentration although the dose consumed was experimentally controlled.

Klingholz, Penning & Liebhardt (1988) examined 15 non-alcoholic male subjects (mean age, 28.7). Eleven subjects consumed alcohol to induce different blood level concentrations whereas 5 acted as controls (sober). The acoustic results showed an increase of F0 variability with alcohol intoxication. The possible explanation for the fact, presented by the authors, is physiological since blood alcohol influences cerebellum function, causes irritation and desensitisation of the proprioceptors thus altering motor coordination and also causes a swelling of the vocal fold mucosa.

A study carried out by Niedzielska & Pruszewics (1996) on 30 alcoholic and 10 non alcoholic adults found a significant increase on the acoustic measurements of F0 range and jitter mean values in the alcoholics in comparison to controls. Within the alcoholic subjects long-lasting alcohol addiction significantly influenced the increase of jitter values.

The above reported studies illustrate a lack of consistency in their findings and a plausible explanation for that may be the research design discrepancies between the studies. Firstly, the studies used different blood alcohol concentration levels and it is generally agreed that pharmacologically at low levels of concentration alcohol acts as a

stimulant whereas at high levels of concentration acts as a depressant (Martin, 1983, 1984) leading to the assumption that the influence of alcohol on F0 will be dependent on the alcohol concentration level. Secondly, either the method of blood alcohol level measurement is not reported (Trojan & Kryspin-Exner, 1968; Niedzielska & Pruszewics, 1996) or it is different from study to study (for example Sobell, Sobell & Coleman, 1982 used a breath analyser whereas Klingholz, Penning & Liebhardt, 1988 used blood sampling analysis). Thirdly, the number of subjects in study is generally small and not all studies mentioned the use of controls such as the Klingholz, Penning & Liebhardt (1988) and Niedzielska & Pruszewics (1996) studies, so changes observed may be due to other unknown reasons.

3.5.9 – Effect of the menstrual cycle on fundamental frequency

Among all the potential factors that may contribute to F0 variations there is one that is unique to females, the endocrinological variations and consequent physiological changes (described in detail in chapter II) related to the menstrual cycle. Although studies support the fact that during the premenstrual and menstrual period some women (especially those with high vocal demands) exhibit a lowering of the F0 or even hoarseness consensus has still not been reached (Silverman & Zimmer, 1978; Brown & Hollien, 1981, 1982, 1983; Abitbol et al., 1995; Chae et al., 2001). Whether this controversy is due to the methodological differences among studies (e.g. characteristics of the subjects, speech samples, physiological measures used) or due to the small amount of data on the subject is still a question of speculation.

Silverman & Zimmer (1978) in a study with 20 young women with no vocal training found that there were no statistically significant F0 differences between ovulation and premenstruation during the production of sustained vowels, measured acoustically. However, eight subjects presented slight lower F0 at premenstruation, nine a slight higher F0 at premenstruation, and three showed no difference. In a replication study with 27 young women the findings of the original study were confirmed: in the non existence of statistical significant differences and in the appearance of individual F0 differences within the subjects.

Brodnitz (1979) in a letter to the editor criticises the validity of Silverman & Zimmer (1978) results considering that testing was used only for vowels at comfortable level (higher pitches and different speech tasks like connected speech may present different results) and none of the subjects was observed laryngoscopically.

No significant acoustic F0 differences in 16 young vocally untrained females (Brown & Hollien, 1981) and in 8 vocally trained females (Brown & Hollien, 1982) were found for the premenstruation, menstruation and ovulation phases using different speech tasks (reading and speaking). However, individually some singers showed either a reduction in the mean F0 or a change in the singing range related to specific stages of the menstrual cycle.

According to Brown & Hollien (1983) several studies (Coleman & Hyler, 1981; Wilson & Purvis, 1988 and Flach et al., 1980) were unable to demonstrate significant correlations between F0 and menstrual cycle phases.

Higgins & Saxman (1989) found a change in the magnitude of the fundamental frequency perturbation (jitter factor) was associated with the ovulation period in ten young women (aged 19-26 years) during the sustained vowels [i] and [u]. Ten young men were used as controls.

Chae et al. (2001) compared the voice quality during the sustained vowel [a] of 28 females (age ranged from 21 to 30 years old) during the premenstrual and menstrual cycle, using an acoustic measurement. They found no statistically significant difference between the two phases although there was an increase of the measures such as jitter, shimmer and normalised noise energy and an F0 decrease during the premenstrual phase. However, within the group 16 females with premenstrual syndrome (diagnosed by the DSM-IV, Diagnostic and Statistical Manual of Mental Disorders) showed a statistically significant increased jitter value in the premenstrual phase.

3.5.10 - Fundamental frequency measurements for dysphonic populations

Reported F0 values for dysphonic or laryngeal pathology populations (Murry, 1978; Hecker & Kreul, 1971; Peppard, Bless & Milenkovic, 1988; Hillman et al., 1990; Holmberg et al., 1994; Murry, Brown & Morris, 1995; Carlson, 1995a; Orlikoff et al.,

1997; Verdonck-de Leeuw, 1998) reflect the same methodological problems found for 'normal' samples concerning the number of subjects, type of speech tasks used and analysis methods. Furthermore, there is a wide range of laryngeal pathologies and dysphonia severity. Still, in the aetiological categorization of voice disorders, mean F0 measures have not been very successful in distinguishing 'pathological' from 'normal' subjects (Klingholtz, 1990; Baken, 1996); however, it has been found that F0 variabilty (pitch sigma) and range measures are useful for assessing the degree of the vocal pathology (Scherer et al., 1988) and in monitoring the effect of intervention (Hirano, 1989; Kotby et al., 1993; Orlikoff et al., 1997).

As might be expected, there is no consensus on the F0 trends for subjects with dysphonia; however, variability and range seem to be greater in subjects with dysphonia compared to 'normal' controls (Hecker and Kreul, 1971; Murry, 1978; Kotby et al., 1993; Orlikoff et al., 1997; Baken & Orlikoff, 2000).

Hecker and Kreul (1971), for example, found that 5 male patients with laryngeal cancer had a higher mean F0 (although not statistically significant) and a more restricted pitch range than normal speakers when reading the second sentence of the 'Rainbow Passage', whilst Lehman et al. (1988 cited by Carlson, 1995a) found no significant difference in mean fundamental frequency between a group of normal speakers and a group of patients with laryngeal carcinoma during reading of the 'Rainbow Passage', followed by a two minute conversation and also sustained phonation on [α], [i] and [u]. Verdonck-de Leeuw (1998) found that laryngeal cancer males tend to have a higher F0 before radiotherapy than after 6 months or several years and than normal controls. Also, after radiation subjects tend to haver lower F0 than normal controls.

Murry (1978) found that the mean speaking fundamental frequency, the standard deviation and the semitone range of the voices of 20 patients with vocal fold palsy were significantly different compared to a sample of 20 normal speakers (from the extracted third sentence of a reading of the first paragraph of the 'Rainbow Passage'). However, mean speaking fundamental frequency did not distinguish normal speakers from two other groups of pathological speakers (a group with benign mass lesions and a group with laryngeal cancer). Data is shown in table III.15.

Table III.15 – Murry (1978) acoustic data

		Normal	Paralysis	Benign mass	Cancer
		N=20	N=20	N=20	N=20
Age	mean	52.5	53.0	57.6	60.7
	standard deviation	-------	11.35	10.05	8.74
F0	Mean (Hz)	121.9	127.0	133.0	133.3
	S.D. (ST)	4.63	3.64	4.34	3.89
F0 range (st)		10.60	7.55	10.55	9.04

Fritzell, Sundberg & Strange-Ebbesen (1982) discovered the F0 of some women with confirmed vocal fold oedema to be considerably lower than expected. Furthermore, the F0 of their entire patient group rose significantly following vocal fold surgery.

Peppard, Bless & Milenkovic (1988) compared a group of 20 young female singers and nonsingers with nodules with another group of 20 young 'healthy' female singers and nonsingers on a wide range of measures (e.g. acoustic, stroboscopic, perceptual and aerodynamic). Results showed significant acoustic differences in the mean F0 and jitter during the sustained vowel [i]. The group with nodules showed significantly lower mean F0 than the 'healthy' group as well as the nonsingers in comparison to singers. Also, the group with nodules presented significantly higher jitter than the 'healthy' subjects; however mean jitter was not significantly different between singers and nonsingers. Acoustic data for the groups is shown in table III.16.

Table III.16 – Peppard, Bless & Milenkovic (1988) acoustic data

	Normal		Nodules	
vowel [i]	Nonsingers	Singers	Nonsingers	Singers
	N=10	N=10	N=10	N=10
F0 (Hz)				
Mean ± SD	212.0 ± 27	244 ± 18	197 ± 22	207 ± 27
Jitter (ms)				
Mean ± SD	0.0358 ± 0.0152	0.0162 ± 0.0135	0.0405 ± 0.0335	0.0478 ± 0.0193

Klingholtz (1990) notes that the F0 distribution fails to distinguish laryngeal pathologies from 'normal' larynx because, for example, a narrow F0 distribution is common in monotonous voices and in cases of mental depression.

Carding and Horsley (1992) reported a wide range of mean F0 in a group of non-organic moderate to severe dysphonic subjects before treatment. Data was extracted from the reading of 'The Rainbow Passage' and DFx1 means from Laryngograph Ltd. software were used. Twenty one women ranging in age from 18 and 65 years old presented a mean F0 of 295.85 Hz (range: 147-576 Hz) whereas four men aged between 28 and 65 years old showed a mean F0 of 248.5 Hz (range: 87-364 Hz).

Kotby et al. (1993) found unimodal F0 distributions for 40 normal subjects (ranging from 18 to 42 years) in contrast with bi-modal or multi-modal F0 distributions for 155 dysphonic patients (ranging from 8 to 72 years) due to the presence of subharmonics or low frequency modulation (about 65 Hz) during the sustained vowel [a]. They also, affirm that the complex pattern appearance was related to dysphonia severity (severe) and not to aetiology. The aetiological categories of the dysphonic group were functional dysphonia (e.g. phonasthenia, hyperfunctional dysphonia), minor associated pathological lesions (e.g. nodules, polyps) and organic dysphonia (e.g. neurological disorder) with a degree of severity ranging from mild to severe.

Drew & Sapir (1995) found a statistically significant difference in the mean F0 (in both reading and spontaneous speaking tasks) between a group of sopranos with symptoms of vocal attrition and a group of 'healthy' singers and nonsingers; however, no statistically significant difference was found for the F0 standard deviation. Speakers with symptoms of vocal attrition showed lower F0 values for reading (mean=201.83, s.d.=23.95) and spontaneous speaking (mean=190.97, s.d.=22.98) than 'healthy' speakers for reading (mean=230.42, s.d.=25.66) and conversation (mean=218.58, s.d.=22.8) respectively.

Carlson (1995a) using electrolaryngography found, on the whole, a lower mean and range F0 (during reading and speech) in 10 irradiated laryngeal cancer patients in comparison to 10 'normal' subjects although the difference was not statistically significant. However, the irregularity measure and the second order F0 maximum range were significantly lower in the irradiated laryngeal cancer patients.

Orlikoff et al. (1997) report similar mean F0 (117.0 Hz, s.d.=28.3 for reading) between two groups with laryngeal cancer (with different response to chemotherapy)

and that F0 mean was typical of healthy men of similar age. However, F0 variability (pitch sigma) was affected by the disease and changed over the course of chemotherapy (as did jitter). Across all patients (in spite of the intersubject variability) F0 variability increased from a baseline of 1.89 ST to 2.55 ST before the last treatment. Moreover, the authors stress that the values are in agreement with those for 'normal' subjects (2.41 ST) reported by Horii (1975) and below those (4.63 ST) reported by Murry (1978).

Buekers (1998a) did not find statistically significant changes (using EGG and acoustic analyses) in pitch and jitter neither in 20 women with voice complaints and laryngeal abnormalities (e.g. nodules, thickening oedema) nor in 12 healthy women before and after a voice endurance test. Also, no significant differences were found between patients and healthy subjects. Neverthless, the F0 range in the patients' group was wider.

3.6 - Fundamental frequency perturbation (jitter)

Any speaker's voice is unstable to some extent due to a huge number of factors (e.g. neurological, emotional and biomechanics), therefore the presence of small perturbation and irregularity of the vocal fold vibration in the voice signal is to be expected (Heiberger & Horii, 1982; Laver, Hiller & Beck, 1992).

Although researchers refer to the existence of at least three different perturbation measures (jitter, shimmer and harmonic-to-noise ratio), only the short-term (cycle-to-cycle) variability in fundamental frequency (or period), not accounted for by voluntary changes in F0, known as jitter will be described here. As Laver, Hiller & Beck (1992) commented that perturbation in this context is the random variability and not the low-rate oscillations of frequency that form the basis of vibrato and tremolo.

Changes in the magnitude of jitter have been associated with: (1) coupling between the glottic and supraglottic region affecting the acoustic pressure dynamic, (2) mucus distribution during vibration, (3) vocal fold histological or mechanical asymmetry, and (4) unsteadiness of the vocal muscle contraction and (5) laryngeal tactile sensation (Sorensen, Horii & Leonard, 1980; Higgins & Saxman, 1989).

Orlikoff & Baken (1990) stress the importance of accurate characterisation of jitter and a precise understanding of its influence for the reliable study of voice and validity in the voice pathology evaluation.

3.6.1 - Jitter formulas

Deem et al. (1989) explain that ' a jitter extraction formula is a symbolic, arithmetic expression of the average of absolute values of jitter per unit of time, and is implemented as part of a jitter calculation algorithm.'

Heiberger & Horii (1982) note that there are several different statistical measurements for obtaining jitter cited by various investigators, 'mean jitter' (defined as the sum of absolute differences in consecutive periods divided by the number of periods minus one) (Hollien, 1979), 'percent jitter' (which is defined as average jitter in milliseconds divided by the average period in milliseconds times 100), 'jitter ratio' (defined as mean jitter in msec divided by mean period times 1000) (Jacob, 1968; Smith et al., 1978; Horii, 1979; Brown, Morris & Michel, 1989; Laver, Hiller & Beck, 1992), 'jitter factor' (which is in fact the percent jitter times 10) (Hollien, Michel & Doherty, 1973), 'pitch perturbation factor' (defined as the percentage of number of times cycle-to-cycle differences were equal to or greater than 0.5 msec), (Lieberman, 1961, 1963), 'directional jitter' (Hecker & Kreul, 1971); 'relative average perturbation' (Koike, 1973); 'pitch perturbation quotient' (David, 1976).

Ingrisano, Perry & Jepson (1998) report that '*RAP (relative average perturbation), PPQ (pitch period perturbation) and APQ (average perturbation quotient) are based on period-to-period variability using a smoothing or averaging factor over three (RAP) or five (PPQ) or 11 (APQ) periods'.*

Heiberger & Horii (1982) state that *'within the frequency ranges of adult males and females, jitter factor and the percentage of jitter described previously are relatively similar. Thus, the percentage of jitter, jitter factor, and jitter ratio (after dividing it by 10) are essentially comparable'.*

3.6.2 – Jitter: instrumental and procedural influences

Nevertheless, the accuracy of jitter measurement is dependent on a number of instrumental and procedural factors (e.g., the recording system, the temporal resolution and the type of speech sample). Moreover, the paucity of published data on a considerable amount of speech material does not allow any consistent knowledge on this subject.

Several researchers presented studies considering the influence of different jitter collecting methods and extraction procedures. Overall, all researchers consider that cassette recorders may perturb the signal (in the time/frequency and amplitude domains) so that true values of jitter are masked and that digitisation (digital audio tape and direct-to-computer recording) is most desirable (Titze, Horii & Scherer, 1987; Doherty & Shipp, 1988; Deem et al., 1989; 1991; Laver, Hiller & Beck, 1992; Sussman & Sapienza, 1994; Gelfer & Fendel, 1995; Perry, Ingrisano & Blair, 1996).

Doherty & Shipp (1988) analysed different recording systems and found that the worst systems were audio recorders that can introduce artifacts to the real signal. Sussman & Sapienza, (1994) explain the presented differences in jitter data in terms of the different collecting methods used, tape-recorded signals collected from microphones, transducers or accelerometers, because they found higher jitter values in the women's high-tongue vowel production when using microphones transducers compared to accelerometers. However, Horii (1979, 1982) reported similar F0 and jitter in 20 men across vowel gathered by the accelerometer. Deem et al. (1989) report that jitter values may be 0.02% reduced when the signals are directly captured in to an A/D system rather than by a tape-recorder. Gelfer & Fendel (1995) found small statistically significant differences from taped samples in comparison to digitised samples of the sustained vowel [a] by 30 young women. Laver, Hiller & Beck (1992) note that '...analog magnetic tape recorders produce phase distortion caused by the recording and playback amplifiers and preamplifiers.' Perry, Ingrisano & Blair (1996) studied the influence of recording systems in the analyses of simulated voice samples and reported that analog cassette recorders introduce distortion whereas digital audio tape (DAT) is as accurate as the direct-to-computer recording.

Titze, Horii & Scherer (1987) and Deem et al. (1989, 1991) compared the differences in jitter magnitude according to different analysis methods (e.g. positive and negative peak picking without interpolation and positive going zero crossing with interpolation) using a variety of stimuli.

Hollien, Michel & Doherty (1973) analysed three 5-second voice samples (of 150, 210 and 260Hz) using five different sampling rates (5, 10, 20, 40 and 80 KHz) and proved that the temporal resolution in the analysis methods is a critical factor for jitter value. As might be seen in figure (III.2) each sampling condition shows different variability according to the sampling rate (Horii, 1979; Heiberger & Horii, 1982). In fact Titze, Horii & Scherer (1987) suggest the use of 20- to 30-cycle sample duration for jitter analysis whereas Laver, Hiller & Beck (1992) report that sampling frequency rates should be not lower than 20 KHz to detect low levels of perturbation in adult speakers. Moreover, data from Deem et al. (1989) in normal speakers indicate that greater stability occurred when using 40 cycles or more. Whether or not this is applicable to pathological voice analysis is not yet known.

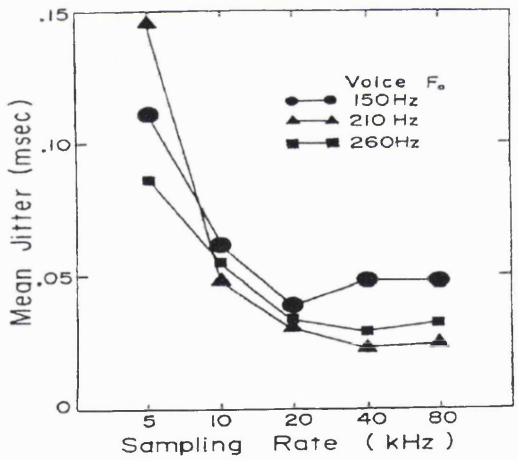


Figure III.2 – Effects of sampling rates on the jitter measurements
From: Heiberger & Horii (1982:310)

Several researchers (Horii, 1979; Heiberger & Horii, 1982; Askenfelt & Hammarberg, 1986; Titze, Horii & Scherer, 1987; Colton & Casper, 1996; Baken & Orlikoff, 2000) recommend the use of only steady, sustained phonation at a comfortable fundamental frequency and loudness for jitter analysis, instead of connected speech, since this last speech task has confounding linguistic and paralinguistic variables. If

connected speech is used derived jitter measures should be interpreted with caution for the reasons already mentioned.

3.6.3 – Jitter variability factors

A considerable number of published studies show that fundamental frequency level (Hollien, Michel & Doherty, 1973; Horii, 1979; Orlikoff & Baken, 1990; Verstraete et al., 1993), gender (Sussman & Sapienza, 1994), ageing (Milenkovic, 1987; Linville & Korabic, 1987; Brown, Morris & Michel, 1989), vowel type (Horii, 1979; Linville & Korabic, 1987; Deam et al., 1989; Sussman & Sapienza, 1994), voice onset and termination (Horii, 1979; Rantala, Määttä & Vilkman, 1997), intensity (Orlikoff, 1991; Orlikoff & Kahane, 1991; Dejonckere, 1998) and dysphonia (Rantala, Määttä & Vilkman, 1997; Dejonckere, 1998; Yiu et al., 2000) are all factors that contribute to different magnitudes of jitter.

Available data on the relationship between F0 level and jitter magnitude is equivocal. Some researchers found a negative correlation between jitter and F0 level (Horii, 1979; Heiberg & Horii, 1982; Linville, 1988; Peppard, Bless & Milenkovic, 1988; Stemple, Stanley & Lee, 1995; Gelfer, 1995), others a positive correlation (Hollien, Michel & Doherty, 1973) while for others this relationship is not straightforward (Orlikoff & Baken, 1990; Gelfer, Andrews & Schmidt, 1991).
For example, in his study with six male adults, Horii (1979) found a tendency for the average jitter (of the sustained vowel [i]) to increase as F0 level decreased as did Linville (1988) in her study with 22 young women during the production of the vowels [i], [u] and [a] and Peppard, Bless & Milenkovic (1988) with two groups of young adult singers and nonsingers with vocal nodules (data is shown in table III.16). Moreover, Horii (1979) note that Beckett (1969) and Jacob (1968) report a decrease in jitter ratio as F0 levels increase. Heiberger & Horii (1982) also report the fact that average jitter is affected by the F0 level (tends to be large for low levels of F0 and smaller for high levels of F0).
Nevertheless, Hollien, Michel & Doherty (1973) found a slight jitter increase as the F0 level increases in four different F0 levels during sustained phonation by four male subjects.

Results on jitter ratio from the study of Gelfer, Andrews & Schmidt (1991) do not correlate with F0 data for both 26 trained and 24 untrained female singers during pre- and posttest measurements. The authors stress that this provides empirical support for the jitter ratio use as an appropriate method in voice evaluation and corroborates Horii's (1979) statement that jitter ratio is not dependent on F0.

Deem et al. (1989) consider that 'percent jitter' is useful for comparing voices with different fundamental frequencies.

Notwithstanding, there is no consensus among researchers about gender differences in jitter measurements. Some authors believe that adult females might have higher jitter values than men (Sorensen & Horii, 1982; Deem et al, 1989; Fitch, 1990), while others argue the reverse (Higgins & Saxman, 1989; Orlikoff and Baken, 1990; Colton & Casper, 1996).

Orlikoff and Baken (1990) reported lower mean absolute jitter for women's voices compared with men's, although the differences were not significant and Colton & Casper (1996) reported an overall mean jitter lower than 0.91msec (standard deviation of 0.63) for males and less than 0.7 msec (standard deviation of 0.59) for females in the sustained vowels [i] and [u] produced by three groups of males and females.

Titze, Horii & Scherer (1987) explains that those gender differences are associated with vocal fold physiology and laryngeal biomechanical gender differences (explained in detail in chapter II).

The effects of ageing on voice are highly documented in the literature (Higgins & Saxman, 1991; Shipp et al., 1992; Linville, 1988; 1996; Russell, Penny & Pemberton, 1995; Baken & Orlikoff, 2000). While some researchers suggest that jitter increases with ageing (Colton & Casper, 1996; Baken & Orlikoff, 2000) others indicate no significant age-related differences (Ramig & Ringel, 1983; Linville & Korabic, 1987; Brown, Morris & Michel, 1989).

Wilcox (1978 cited by Heiberger & Horii, 1982) presents significantly greater magnitude of jitter in sustained phonation for older adult males (.75%) than younger

adult males (.57%). Colton & Casper (1996) reported that jitter values tend to increase as a function of age. Brown, Morris & Michel (1989) found no significant difference between the jitter ratio of sustained [ɑ] at comfortable loudness produced by young and old women. Recent insights, however, question whether the dichotomy chronological/biological age is one of the possible reasons for the different data obtained in the published studies.

Research on jitter variability across adult vowel type remains another unsolved question. Although, several researchers agree that jitter values are significantly influenced by vowel type (Ramig & Ringel, 1983; Milenkovic, 1987; Linville, 1988; Deem et al., 1989) others (Horii, 1982; Heiberger & Horii, 1982; Orlikoff, 1995) claim that there are no physiological reasons for that. Anyway, consensus has still not been reached for the magnitude of those differences. Some have reported significantly greater amounts of jitter for the tongue-low vowel [a] than the tongue-high vowels [i] and [u] (Sussman & Sapienza, 1994; Deem et al., 1989; Linville and Korabic, 1987; Milenkovic, 1987), while others observed a tendency for tongue-high vowels to show greater jitter than low ones (Horii, 1982; Linville, 1988; Nitrouer et al., 1990).

Sorensen & Horii (1982) reported higher percent jitter values for [i] than for [a] and [u] for 20 women studied, with a similar pattern in men reported by Horii (1979).

Linville (1988) in a study with 22 healthy young women demonstrated that there are statistically significant percent jitter differences among vowels [i], [a] and [u]. Vowel [i] has the highest magnitude (mean = 0.97%) and variability (s.d = 0.47) followed by [u] and [a]. Data is presented in table III.17.

Wilcox (1978 cited by Heiberger & Horii, 1982) found greater magnitudes of jitter for [i] and for [ɑ] (respectively .69 and .68%) than for [u] (.55%) in young and old male adults and Horii (1980) also found a small jitter value for [u], intermediate for [i], and greatest for [ɑ]. Also, Jacob (1968 cited by Heiberger & Horii, 1982) found the greatest magnitude of jitter for the sustained vowel [ɑ] at low F0 level and intensity.

Higgins & Saxman (1989) found significant jitter factor magnitudes differences between [i] and [u] productions for 5 young men (aged 19-26 years) respectively (mean

0.69 and 0.62). However, no significant differences were found for 10 women (aged 19-25 years) between the same vowels (mean 0.41 and 0.43).

To examine the differences found among various vowels it is critical to stress some elementary principles that have been already focused on (e.g., the analysis methods employed must be comparable, the F0 level used must be identical). If those conditions are controlled and still jitter differences are found, the possible recognised correlations are the 'inherent F0 levels' of the vowels and some interactions between glottal sound source and vocal tract transfer function (Heiberger & Horii, 1982).

Data on the jitter mean and standard deviation for females and males is shown in tables III.17 and III.18.

Table III.17 – Mean and standard deviation jitter for adult females

N	Age average / range	[u]	[i]	[a]	Jitter calculation	Sound source
20	adult			8.4 ± .3	Jitter ratio	Sorensen & Horii (1982)
10	(16-30)	-------	.036 ± .015	-----	Absolute jitter magnitude (msec)	Peppard, Bless & Milenkovic (1988)
22	20.32	0.84 ± 0.43	0.97 ± 0.47	0.55 ± 0.36	Percent jitter (%)	Linville (1988)
30	21.8 (18-30)	0.31 ± 0.26	0.33 ± .23	0.38 ± 0.17	Percent jitter (%)	Deem et al. (1991)
25	27.5 (20-32)			[ɑ]= 0.78	Jitter ratio	Brown, Morris & Michel (1989)
7	33 (26-48)	0.44 ± 0.13	0.44 ± 0.13	0.47 ± 0.18	Percent jitter (%)	Deem et al. (1989)
6	30.2 (21-44)	--------	------	9.36 ± 4.57	Jitter ratio	Orlikoff & Baken (1990)
13	51.6 ± 22.2	-------	----------	1.15 ± 1.16	Percent jitter (%)	Hertrich & Ackerman (1995)

Table III.18 – Mean and standard deviation jitter for adult males

N	Age average/ Range	[u]	[i]	[a]	Jitter calculation	Sound source
30	21.1 (18-30)	.29 ± .12	.28 ± .12	.37 ± .16	Percent jitter (%)	Deem et al. (1991)
20	23.3 (18-26)	0.51	0.61	------	Percent jitter (%)	Wilcox (1980 cited by Higgins & Saxman, 1989)
20	27.5 (22-37)	.66 ± .26	.68 ± .22	.86 ± .51	Percent jitter (%)	Horii (1982)
6	(28-43)		36		Mean Jitter (msec)	Horii (1979)
6	(23-47)	----------	------	6.24 ± 3.12	Jitter ratio	Orlikoff & Baken (1990)
5	33.2 (26-45)	0.38 ± 0.20	0.33 ± 0.11	0.56 ± 0.16	Percent jitter (%)	Deem et al. (1989)
10	32 ± 5.5	.046 ± .019	.051± .020	--------	Absolute jitter magnitude (msec)	Orlikoff (1995)
11	48.4 ± 16.5	------	--------	.44 ± .36	Percent jitter (%)	Hertrich & Ackerman (1995)

The onset and termination of phonation typically have much greater frequency perturbation than the steady state mid portion of a sustained vowel (Horii, 1979; Rantala, Määttä & Vilkman, 1997).

Koike (1973 cited by Baken and Orlikoff, 2000) found a higher mean relative average perturbation on the first 17 glottal cycles of a normal (soft) vocal initiation (0.0276) than in steady-state phonation (0.0046), in his study of normal men and women.

Rantala, Määttä & Vilkman (1997), in a study of 11 teachers, with symptoms of voice fatigue, found that the mean jitter percent for the first portion of a prolonged phonation of [a] was lower than the mean for the latter portion (1.35% versus 1.47% respectively). The analysed portion had an average of 500 cycles and the first and latter portion were determined symmetrically from the midpoint of the vowel. They also refer a high intra- and intersubject variability in jitter measures. The authors explained the intersubject variability on the basis of the field test circumstances (different school places); however, other confounding variables should be taken into account such as two subjects had been in voice therapy before.

Variations in sound pressure level of voicing can have different consequences on jitter magnitude. Orlikoff (1991) hypothesis that large vibrating mass, in louder voicing, contributes to stabilisation of the oscillator while low voicing is often related to shortened vocal fold contact and consequently associated with great randomness and increased airflow turbulence. Dejonckere (1998) points out that louder voicing is related to lower jitter magnitudes in healthy larynxes or superficial vocal fold pathology '...probably due to the increase of the oscillating mass by larger involvement of normal deeper layers of the vocal fold.' He also explains that in cancer and vocal paralysis louder voicing increases the irregularity of vocal fold vibration and consequently the jitter magnitude.

Dejonckere (1998) in a study with 87 patients (mean age, 36.2 years) with several larynx pathologies (e.g. nodules, chronic laryngitis, vocal fold paralysis) found a significantly larger jitter variance for vowel [a] at louder voicing than for comfortable voicing.

Orlikoff and Kahane (1991) confirmed, in their study using men phonating the vowel [α] in modal voice, at three different sound pressure levels trying to keep the fundamental frequency equal during increase in loudness, that there is a significant negative correlation between jitter ratio and vocal intensity (jitter tends to decrease with increased vocal loudness).

3.6.4 - Data on jitter

There is so little data on jitter that is not possible to consider any normative values (Rantala, Määttä & Vilkman 1997; Yiu et al., 2000). However, the published data on percent jitter (Hollien, Michel & Doherty, 1973; Horii, 1979, 1982; Heiberger & Horii, 1982; Deem et al., 1989) indicates a typical jitter value of 0.5 to 1.0% for normal sustained phonations by young adults while Lieberman (1963) reports that only 15% of normal adult males exceed 0.5 msec in magnitude (Iwata & von Leden, 1970; Laver, Hiller & Beck, 1992). Horii (1979) converted data from the studies of Smith (1972), Hollien, Michel & Doherty (1973) and Jacob (1968) in jitter ratio and shows that for normal phonation the jitter ratio magnitudes are respectively 3.7 to 17 (Smith, 1972), 3 to 14 (Hollien, Michel & Doherty 1973) and 3.4 to 123.2 (Jacob, 1968). However, he

stresses that from his experiment jitter values such as reported by Jacob (1968) were not found.

One of the potential major interests of jitter measures is the differential diagnosis between 'normal' and 'pathological' voice. Although jitter values are valuable for the evaluation and validation of the perceptual quality of dysphonia this should not imply that they can be used as the sole diagnostic criterion for voice disorders (Lieberman, 1961, 1963; Iwata & von Leden, 1970; Hecker and Kreul, 1971, Horii, 1979; Rontal et al., 1983; Murry and Doherty, 1980; Heiberger & Horii, 1982; Askenfelt and Hammarberg, 1980, 1986, Wolfe and Steinfatt, 1987, Karnell, 1991; Hertrich & Ackermann, 1995; Yiu et al., 2000).

Heiberger & Horii (1982) comment that '*The work by Lieberman and his colleagues (Lieberman, 1961, 1962; Lieberman & Michaels, 1962; Smith & Liberman, 1969) probably represents the pioneering studies of laryngeal pathology detection by waveform analysis*'.

Results from Iwata & von Leden's (1970) study of 30 normal subjects and 52 subjects with laryngeal pathology (e.g. chronic laryngitis, tumors, vocal fold paralysis) indicate significant differences in jitter between the two groups in the sustained vowel [a]. The normal subjects presented a jitter range that extended from –0.6 to + 0.5 msec (in accordance with Lieberman's (1963) data already mentioned). Within the 'pathological' group, significant differences were found between patients with chronic laryngitis (jitter ranged from –0.7 to +1.3 msec) and patients with vocal nodules (jitter ranged from -0.7 to +0.7 msec). Also, laryngeal carcinoma patients (jitter ranged from –8.6 to +9 msec), and laryngeal paralysis patients (jitter ranged from –0.9 to +1.0 msec for the untreated patients and -0.3 to +0.3 msec for the treated patients) differed significantly.

Crystal et al. (1970 cited by Heiberger & Horii, 1982) discriminated normal from pathological larynges, especially for vowel productions by men, using a total of 40 vocal quality indexes (including the cycle-to-cycle period and amplitude differences) from connected speech and sustained [a] produced by 172 normal and 52 pathological larynges.

Klingholz & Martin (1985) found that the mean jitter factor of 18 women with hyperfunctional dysphonia (higher laryngeal muscular tension than in normal phonation) is statistically significantly different from the mean jitter factor of 14 women with hypofunctional dysphonia (lower laryngeal muscular tension than in normal phonation) during the production of the sustained vowel [a] at comfortable level. The hyperfunctional patients showed lower jitter values (data is not presented in the article). A control group of 20 women with normal voice was used and the minimal and maximum jitter factor values reported were 0.3% and 1.6% respectively.

Haji et al. (1986) used EGG measurements to evaluate the frequency perturbation and auditory evaluation during sustained phonation in 30 healthy speakers (between 24 and 59 years) and 33 speakers with several laryngeal pathologies (eg. nodules, polyps and vocal fold paralysis) aged between 14 and 77 years. They found a significant difference in EGG frequency perturbation only between the moderate and extreme hoarse groups and no significant difference among normal, slight and moderate hoarse groups. Also, no gender significant differences were found for the frequency perturbation.

Pruszewicz et al. (1991) found a higher mean jitter value (8%) for 15 patients with organic pathology (e.g., polyps, papilloma, pachydermia) than for 15 patients with functional dysphonia without laryngeal pathology (4%). The maximum mean value of jitter for the first group was 24% and for the second was 28%. However, it is not clearly explained to which vowels correspond those values (either [i] or [e]) and patients' ages and genders were not given.

Hertrich & Ackermann (1995) found significant increased values of jitter in 15 men with idiopathic Parkinson's disease as compared to 12 healthy men.

Orlikoff et al. (1997) found that mean jitter was significantly different across chemotherapy cycles (in two groups of laryngeal cancer with different response to chemotherapy) and the average jitter magnitudes for both laryngeal cancer groups, complete response to chemotherapy (mean=1.62%, s.d.= 1.38) and incomplete response to chemotherapy (mean=1.45%, s.d.= 1.50) were higher than the expected values for 'normal' adult males of approximately the same age. Also, patients with a major tumour reduction showed lower jitter values than those with none or minimal tumour reduction.

From a large database of 389 pathologic voices Omori et al. (1997) found 20 roughness voices (females with a mean age of 52.1 years and males with a mean age of 54.2 years) and from those, 5 presented jitter values within normal range (0.15-0.53% for male and 0.07-0.69% for female) although subharmonics were found in their spectrum analysis. Detailed data is shown in table III.19.

Table III.19 – Omori et al. (1997) data

Case nº	Gender	Larynx pathology	Mean F0 (Hz)	Jitter (%)
1	M	Polyp	128	0.82
2	F	Polyp	115	0.34*
3	F	Polyp	132	1.46
4	F	Polyp	240	0.49*
5	M	Polyp	129	2.05
6	F	Atrophy	207	7.00
7	F	Nodule	168	1.17
8	F	Polyp	218	2.11
9	F	Polyp	188	0.91
10	M	Cyst	156	1.12
11	M	Polyp	144	0.76
12	M	Atrophy	140	5.26
13	F	Polyp	161	1.54
14	F	Polyp	232	1.22
15	F	Paralysis	265	3.64
16	F	Polyp	272	0.53*
17	M	Polyp	101	0.45*
18	F	Paralysis	235	0.51*
19	M	Polyp	122	2.12
20	M	Paralysis	172	0.85

* within the normal range

Careful considerations should be made when interpreting data from pathological phonation especially because of the type of speech sample used, either sustained phonation or connected speech, and the effects of the temporal resolution of some of the analysis methods used (as previously mentioned).

Chapter IV -Voice quality evaluation

4.1 - Introduction

The purpose of this chapter is to provide a summarised review of voice evaluation methodology, focusing especially on the tools (case history, self-ratings, stroboscopy and electrolaryngographic measurement) used in the experimental study.

4.2 - Patient's medical history

A standard adult case history follows a methodology that covers areas that may potentially contribute to a holistic view of the origin and development of the voice problem, such as voice and medical history, as well as the life-style (health habits, psychological and sociological information) of the subject (Verdolini, 1994; Koschkee & Rammage, 1997). Usually, data is gathered during an interview where history questionnaires and self-assessments (e.g. the individual's perspective on potential sources of stress, related to either changes within the family or work setting) are used.

4.2.1 - Case history questionnaire

History questionnaire tools, used during an interview, may ensure a more consistent, systematic and comprehensive methodology in gathering data and recording it (Greene & Mathieson, 1989; Morrison et al., 1994).
There are a wide variety of history forms available. Some are designed for professional voice users (Sataloff et al., 1994a), other for adults (Boone, 1982; Guimarães & Cruz, 1995); patients with paradoxical vocal fold dysfunction (Bless, 1994 cited in Koschkee & Rammage, 1997), spasmodic dysphonias (Bless & Ford, 1989 cited in Koschkee & Rammage, 1997), children (Pindzola, 1987; Guimarães & Cruz, 1997a) or non-specified age or pathology (Lockhart & Martin, 1997; Greene & Mathieson, 1989; Morrison et al., 1994; Dworkin & Meleca, 1997). However, it is not uncommon for clinicians to develop their own structured interview tools (questionnaires). That was the case in this experimental study (see chapter V).

It is worth noting that the authors referred to above give comprehensive instructions for the appropriate use of these tools, outlining what factors it is important to assess and how these ought to be assessed. Despite this, clinicians are aware that non-standardised assessments may have some degree of bias.

4.2.2 - Self-assessment of life events

Among the social psychological models for life stress processes, the interaction between personal predisposition and social circumstances is a highly important one clinically (for medical and health-related professionals). It allows important basic understanding about the effects of psychological and social data in the precipitating and perpetuating factors in illness. Since the early 1950's the literature on psychosomatic medicine has offered a large body of research demonstrating the association between stressors and their detrimental effect on the subject's health, (Holmes & David, 1989; Miller, 1989; Rosen & Sataloff, 1997).

Rosen & Sataloff (1997) recognised stress as a factor in illness and disease and estimate 50% to 70% of illness complaints are stress-related.

To Holmes & Rahe (1967) the term stressor is generally taken to mean those life events (either 'negative' in a sense they are socially unfavourable or 'positive', socially favourable) that evoke an adaptative or coping behaviour on the part of the involved individual. Each person's life is characterised by various life events that may occur in social environments (e.g. work, school), home, marriage and health. In many respects, an life event change recently experienced can be considered a 'risk factor' for the development of an illness.

Miller (1989) notes that depressed patients tend to report more recent life change events than do healthy controls. However, Danish, Smyer, and Nowak (1988 cited by Miller, 1989) point out that events per se do not cause change unless factors such as time, sequencing, and duration are considered also.

In spite of the potential biases acknowledged by several authors concerning the methodologies used, (e.g. samples of convenience, mainly retrospective studies and combining both unfavourable and favourable events) not only can an inventory of recent

life events be made, but some estimation of the magnitudes of these events can be ascertained (Miller, 1989).

Though there remains no unanimity of opinion among researchers about the definitive list of stress-causing life events, and as to which life events weigh more significantly in assessing impact on human behaviour, one of the first published lists, and still one of the main methods of assessment most widely used, because of its simplicity, availability and economy is the Social Readjustment Rating Scale (SRRS) developed by Holmes & Rahe (1967) (appendix F and described in detail in chapter V).

The aim of the SRRS was to sample salient events (derived from the life charts of clinic patients), rather than to cover all the possibilities, scaled in terms of the amount of change they produce in a person's adaptability to stress producing events (Holmes & David, 1989; Miller, 1989). It has been subjected to criticism for oversimplification of complex issues, and questioned about its validity, retest and inter-rater reliability (Miller, 1989). Moreover, it is recognised that several factors may affect the individual's responses to life events, namely cultural and socioenomic status, income, and interpersonal support relationships (family and friends).
Cross-cultural validation studies indicate that groups of individuals showed remarkable agreement as to average degrees of life change and readjustment engendered by each of the events on the list though totally culturally diverse (e.g. American, Western European, Japanese) (Masuda & Holmes, 1967; Harmon, Masuda & Holmes, 1970; Holmes & David, 1989).

In the voice field many authors of textbooks (Prater & Swift, 1984; Greene & Mathieson, 1989; Haynes, Pindzola & Emerick, 1992; Aronson, 1990; Jackson-Menaldi, 1992; Morrison et al., 1994, Koschkee & Rammage, 1997; Rosen & Sataloff, 1997; Freeman, 2000) imply that potential sources of stress, related to psychosocial factors, are linked to the onset and maintenance of dysphonia (especially those labelled as 'functional', 'psychogenic' or 'hysterical'). Additionally, a great number of the published articles (Rantala, Määttä & Vilkman, 1997; Buekers, 1998a) mention that in many professionals, high demands on voice are dependent among other factors (e.g. vocal loading) on psychological stress. A review of the literature for information concerning the experimental evidence of psychosocial factors in voice, however, shows

a quantitatively restricted number of published studies (House & Andrews, 1987; Kinzl, Biebl & Rauchegger, 1988; Goldman et al., 1996; Bough et al., 1996; Epstein, 1999).

The Bedford College Life Events and Difficulties Schedule was used by House and Andrews (1987) in a group of patients with 'functional' voice disorders and a normal control group to determine the significance of the life events experienced in both groups. Although a much higher percentage of the voice-disordered group did report experiencing 'a difficulty or event which involved conflict over speaking out' twelve months prior to onset of the disorder, there were no significant differences between the groups in other experienced life events.

Through the use of a life event inventory in twenty-two women with functional aphonia (that did not involve organic changes of the vocal folds), Kinzl, Biebl, and Rauchegger (1988) show that only a few patients (there is no specific reference to the number) with depressive neurosis experienced significant life event stressors during the preceding 6 months. They also found that in half of the women the somatization of symptoms (e.g. upper respiratory tract infection with concomitant hoarseness) was a result of the subjects' inability to deal with actual or chronic stress and keep her psychic stability.

Goldman et al. (1996) studied the role of four psychosocial factors including stress (measured by the Social Readjustment Rating Questionnaire, Holmes & Rahe, 1967) in three groups of adult female subjects (27 with vocal nodules, 17 with hyperfunctional voice disorders and 33 with no history of voice disorders). Once again, according to the stress factor, only the hyperfunctional voice disorders group showed significantly increased scores of stress compared with the group without voice disorders.

Bough et al. (1996) used a self-rating scale of stress (ranging from 1=severe and 5= stress free) in a study with fourteen healthy subjects (eight women, mean age 31 and six men with a mean age of 36) during 15 test sessions for acoustic measurement. A statistically significant negative correlation (although less than 0.25 of magnitude) was found between the stress level and F0 and jitter.

4.3 - Perceptual evaluation

The term perceptual evaluation means, according to Fex (1992), that '...a listener is making a comparison between a (perhaps unspecified) number of qualities that the listener can hear in the speaker's voice and the listener's own opinion about how these different qualities should sound in the normal voice.' Either informally or formally this type of evaluation requires a listener to judge a voice sample according to various parameters of voice quality.

The term *perceptual evaluation* instead of *psychoacoustic' evaluation* was suggested by Fex (1992) and recommended for standard use by the Voice Committee from the International Association of Logopedics and Phoniatrics (IALP) as either *perceptual evaluation or perceptual assessment* (Bless and Baken, 1992).

Within the extensive battery of tests for voice evaluation, perceptual evaluation is still an important and widely used method in the clinical field (Bassich and Ludlow, 1986; Tveteras, 1992; Hammarberg, 2000, Carding et al., 2000) not only because the listener's ear may be the only tool available (Fex, 1992; Orlikoff et al., 1999) but also because the patient's complaints are mainly based on auditory perceptions (Orkiloff et al., 1999). Furthermore, systematic research around the world has shown that perceptual voice parameters may have a quantifiable basis that can correlate with other forms of assessment, namely laryngoscopy observations (Hirano & Bless, 1993), acoustic analysis (Emanuel & Sansone, 1969; Eskenazi, Childers & Hicks, 1990; Wolfe, Cornell & Palmer, 1991; Rabinov et al., 1995; McAllister, Sundberg & Hibi; 1998; Millet & Dejonckere, 1998) and electrolaryngography (Wechsler, 1977).

Despite such arguments, perceptual evaluation is one of the most controversial topics in voice research. It is widely acknowledged that the perceptual rating of voice quality has a number of potential intra- and inter-judge reliability disagreements. A complex interaction of factors like the listener's references, the background education and experience in judging voices, the type and length of voice samples being judged, the type of rating scale being selected, may influence reliability (Bassich & Ludlow, 1986; Fex, 1992; Tveteras, 1992; Orlikoff et al., 1999; Carding et al., 2000).

Fex (1992) mentioned that the existing, quantity and quality, differences between different listeners result from the fact that there is no standardised and universal definition of 'normal' voice quality among clinicians and researchers. Some may define it according to structural and functional features while others emphasise other aspects (e.g. linguistic and aesthetic), (described in detail in chapter chapter III). So, it is not surprising that the type and number of dimensions used for describing a voice vary from listener to listener.

Furthermore perceptual evaluation is likely to be dependent also on the background education and experience in evaluating voices (Bassich and Ludlow, 1986; Gelfer, 1988; Kreiman et al., 1992, 1993; Buekers, 1998b; Verdonck-de Leeuw, 1998). In fact, Buekers (1998b) showed that the variance in the perceptual voice evaluation between experienced professionals (speech therapists and voice experts) was not significantly different, but it was markedly less than in the speech therapy students. Bassich & Ludlow (1986) state that the reported levels of intra-judge reliability for different types of listeners are highly variable – between .78 for speech-language pathologists (Murry et al., 1977) and .40 for inexperienced college students (Kruel & Hecker, 1971). Verdonck-de Leuuw (1998) found that the perceptual voice analysis of cancer speakers following radiotherapy by trained raters (speech and language therapists) is analytical and related best to acoustical and physiological analysis than the perceptual voice analysis by naïve raters that represent best the judgements made by the patients social environment (e.g. family and friends).

Additionally, Kearns & Simmons (1988 cited by Kreiman et al., 1993) pointed out that raters agreed about what constitutes normal phonation or severe pathology but were less confident when rating the severity of pathology in mild to moderately pathological voices.

Moreover, for some dimensions, inter-judge reliability is higher than for others (Bassich & Ludlow, 1986; Buekers, 1998b). Bassich & Ludlow (1986) state that hoarseness ratings vary between .17 and .69 (Shipp & Huntington, 1965), while breathiness ratings are more reliable at .95 (Kruel & Hecker, 1971). Buekers (1998b) found a higher inter-judge correlation for resonator aspects (.57) than voice aspects (.35); however it should be noted that voices were rated by means of video-recordings and not by auditory samples as the majority of the studies. Bassich and Ludlow (1986)

reported poor judge reliability in their 13 dimensional perceptual rating system to assess sustained vowels. However, the authors recognise the type of raters (four inexperienced speech and language therapist raters with 16 hours training for the task) and the type of sample used (vowel prolongations) as a limitation in the study.

Another great constraint, related to the studies described above, is the disparity in the design of the available perceptual rating scales. Some scales use a descriptive system which allows the listener to analyse and describe normal and pathological voices at laryngeal (source) and supralaryngeal (filter) levels (e.g. 'VPA', Vocal Profile Analysis from Laver, 1980), while others are focused only on the laryngeal level (e.g. 'GRBAS', Grade, Rough, Breathy, Asthenic and Strained from Hirano, 1981) and others include general aspects related to voice behaviour (e.g. VAP-Voice assessment protocol from Pindzola, 1987 and Voicing Evaluation Scheme from Buekers, 1998b). According to Carding et al. (2000) the number of parameters used varies between 5 in the GRBAS, 31 in the VPA and 12 in the Buffalo III. Also, the rating system may vary from: categorical, equal-appearing interval scales (e.g. GRBAS, Hirano, 1981) to visual analogue scales, direct magnitude estimations and paired comparison tasks (Verdonck-de-Leeuw, 1998; Carding et al., 2000).

To assure the reliability and the validity of the judgments several expert clinicians and researchers (Fex, 1992; Orlikoff et al., 1999; Carding et al., 2000) recommend the use of unambiguous terminology and definitions; the continuous training of the listeners and the use of running speech samples (Bassich & Ludlow, 1986; Verdonck-de Leeuw, 1998).

Bassich & Ludlow (1986) suggest the use of running speech instead of sustained phonation because the former is physiologically more complex and consequently has more aspects of deviant voice quality. This is corrobated by Verdonck-de Leeuw (1998) in the perceptual analysis of cancer patients following radiotherapy where connected speech was shown to be more suitable than the sustained vowel [a].

Finally, some investigators (Wolfe, Cornell & Palmer, 1991; Carding et al., 2000; Orlikoff et al., 2000) refer to the usefulness and the gain in credibility when the perceptual judgments of voice disorders are complemented by physical measurements. Therefore, attempts to explore the relationship between the perceptual descriptors of

voice quality as well as the dysphonic severity and their acoustic correlates have been the focus of many investigations (Wolfe, Cornell & Palmer, 1991; Kempster, Kistler & Hillenbrand, 1991; Karnell, 1991; Rabinov et al., 1995; Wolfe, Fitch & Cornell, 1995; Hillenbrand & Houde, 1996, Wolfe & Martin, 1997; Verdonck-de Leeuw, 1998). However, results from studies examining the correlation between perceptual qualities and physical measures (jitter, shimmer, or harmonic-to-noise ratio) present contradictory findings because there are divergences in the speech material used (sustained vowels and connected speech), raters (trained or naïve), rating scales (terminology and parameters) and different acoustical analyses (different algorithms).

Firstly, a voice quality may correlate with different acoustic measures or vice versa. For example, judgments of '*breathiness*' were found to be associated with percent jitter (Eskenazi, Childers & Hicks, 1990) and spectral noise (Klatt & Klatt, 1990; Orlikoff et al., 1999). Kempster, Kistler & Hillenbrand (1991) note that judgments of '*roughness*' were found to be related to measures of jitter (Kitjaima & Gould, 1976; Wendhal, 1966; Kreiman et al., 1993), to harmonic-to-noise ratio and spectral flatness of the residual signal (SFR), (Eskenazi, Childers & Hicks, 1990) and noise level (Deal & Emanuel, 1978; Yanagihara, 1967; Yumoto, Gould & Baer, 1982). Both '*breathiness*' and '*roughness*' were found to be associated with harmonic-to-noise ratio (Martin, Fitch & Wolfe, 1995; De Krom, 1995).
Moreover, Emanuel & Sansone (1969) have reported that the level and frequency location of spectral noise in vowels sustained by seven normal speaking males with 'normal' and simulated 'roughness' voices are related to the perceived severity of roughness. Roughness voice presented wider noise spectral ranges than normal voice and the spectral noise levels are related to the tongue height in each vowel. Overall, high tongue vowels (e.g.[i]) showed less spectral noise than the lowest vowels (e.g.[a]).

Although studies have been focused almost exclusively on univariate correlational techniques between the voice quality ratings and the physical measurement it seems obvious for the reasons given above, and because voice quality is an inherently multidimensional phenomenon, that multivariate correlation should be used. In fact, Verdonck-de Leeuw (1998) found that results on single acoustical correlates of perceptual ratings (of laryngeal cancer subjects) were improved by the use of multiple analyses prediction.

Secondly, although studies demonstrated a correlation between the acoustic measures and the perceived deviation from normal voice quality, several authors highlight the problem of the numerous measures used and the degrees of correlation reported (even for the same acoustic measure).

Jitter is one of most commonly used acoustic measures of voice quality because it is available in almost all commercial systems (even with different measures they are easily comparable); it has been studied and correlates with dysphonia severity (Wolfe, Cornell & Palmer, 1991) as well as with breathiness (Eskenazi, Childers & Hicks, 1990), roughness (Heiberger & Horii, 1982) and hoarseness (Yumoto, Sasaki & Okamura, 1984).

Wolfe, Fitch & Cornell (1995) mention that Deal and Emanuel (1978) found positive correlations (from 0.03 to 0.69) between perceptual roughness and jitter for five vowels produced by 20 individuals with clinical hoarseness whilst Wolfe and Steinfatt (1987) found a moderate correlation of 0.51 between roughness and jitter for vowels [a] and [i] produced by 51 individuals with a wide variety of voice problems.

Yumoto, Sasaki & Okamura (1984) demonstrated a moderately high correlation of 0.71 between hoarseness and jitter for vowels produced by 87 individuals varying from nearly normal to severely hoarse.

Multiple regression analysis applied to vowel data, from adult males from 60 to 80 years old, indicated that jitter made a minimal contribution in predicting the roughness levels of the voices (-.01). The strongest predictor was the mean F0 which suggests that perceptual notion of roughness is intrinsically tied with perceived voice pitch (Heiberger & Horii, 1982)

Smith et al. (1978) found nonsignificant correlation between jitter and roughness in oesophageal speech samples although experienced listeners rated reliably the roughness of each voice sample (Heiberger & Horii, 1982; Wolfe, Fitch & Cornell, 1995).

4.4. - Physical examination and measurement

4.4.1 - Techniques for laryngeal examination

A review of literature reveals a wide variety of techniques available to assess vocal function, such as stroboscopy (Hirano, 1981; Faure & Muller, 1992; Hammarberg, 2000) acoustic analyses (Gould & Korovin, 1994; Eustace, Stemple & Lee, 1996; Hammarberg, 2000); electrolaryngography (Abberton, Howard & Fourcin, 1989; Colton & Conture, 1990; Baken, 1992; Fourcin et al., 1995; Abberton & Fourcin, 1997; Orlikoff, 1998; Fourcin, 2000; Baken & Orlikoff, 2000; Hammarberg, 2000); flow glottography (Gould & Korovin, 1994; Hammarberg, 2000), inverse filtering (Hammarberg, 2000) and electromyography (Gould & Korovin, 1994; Hammarberg, 2000). The value of these techniques is dependent on the purposes of the voice evaluation. Some are not always clinically practicable either due to the technique itself (e.g. invasive electromyography), or for economic reasons (e.g. ultra-speed photoglottography) or even because they are very time-consuming (e.g. spectrography). Extensive and important information about the techniques referred to are described in detail in several textbooks and articles (e.g., Gould & Korovin, 1994; Baken & Orlikoff, 2000; Hammarberg, 2000).

Only two techniques, used in this study, will be described below, one for the visualization of vocal fold movement (stroboscopy) and another as a correlate of vocal fold motion (electrolaryngography, ELG).

Stroboscopy is one of the most widely available and recommended routine clinical techniques because it is the only way to observe, more or less directly, the larynx and search for any structural or functional pathology (Laver, Hiller & Beck, 1992; Hammarberg, 2000; Baken & Orlikoff, 2000).

ELG has become a widely used tool for voice study/examination as an aid to diagnosis (Fourcin, 1981; Dejonckere & Lebacq, 1985; Painter, 1988; Orlikoff, 1991; Murty, Carding & Lancaster, 1991; Winstanley & Wright, 1991; Hertrich & Ackerman, 1995; Watson, 1995; Carlson, 1995a; Selby & Wilson, 1997; Verdonck-de Leeuw, 1998; Cheyne, Nuss and Hilman, 1999; Blomgren et al., 1998; Verdolini et al., 1998) and also as an outcome measure for speech and language therapy (Fourcin & Abberton,

1971, 1976; Wechsler, 1977; Carding & Horsley, 1992; Carlson, 1993, 1995b) especially because it is entirely non-invasive.

4.4.2 - Stroboscopy

The history and the principles on which stroboscopy is based (known as Talbot's law) have been extensively presented in the literature (Hirano, 1981; Wendler, 1992; Faure & Muller, 1992; Hirano & Bless, 1993; Gould & Korovin, 1994 and Baken and Orlikoff, 2000).

Hirano (1981: 47) describes the principle of stroboscopy as follows:
'*The light source of the stroboscope emits intermittent flashes of light which are synchronous with the vibratory cycles. The source of the trigger signal for the light flashes is the waveform of the subject's voice. When the flashes are emitted at the same frequency as that of the vocal fold vibration, that is, at an identical phase point in successive vibratory cycles, a sharp and clear image of the vocal folds is observed, assuming that the reappearance of the waveform is maintained constant. When the flashes are emitted at frequencies slightly less than the frequency of vocal fold vibration, giving rise to a systematic phase delay of the consecutive light flashes, a slow motion effect is produced.*'

Furthermore, Faure & Muller (1992) pointed out that '*according to Kitzing, the difference between the phonatory frequency and the rate of light flashes should be approximately 1.5 Hz.*'

Nowadays, stroboscopy can be used with either a rigid (oral route) or flexible (nasal route) fiberoptic endoscope and can be equipped with videotaping or computing facilities (Casper, Brewer and Colton, 1987; Stemple, Glaze & Gerdeman, 1995).

The fiberoptic endoscope, also called fiberscope, has the great advantage of non-interference with speech production. According to researchers (Hirano & Bless, 1993; Stemple, Glaze & Gerdeman, 1995; Baken & Orlikoff, 2000) this allows insights into the co-ordination between the larynx and the rest of the vocal tract; ideal for patients with motor speech disorders and professional voice users.

Although the fiberoptic endoscope is a great clinical aid, it has the disadvantage of being more invasive than the rigid endoscope, and the captured image may have several problems. Firstly, due to the natural movement of the articulators (e.g. in swallowing) visualization may be difficult and it is also difficult to obtain a stable image (Baken & Orlikoff, 2000). Secondly, the image may be darker because of the smaller diameter of the fiberoptic light bundle (Casper, Brewer, & Colton, 1988; Hibi et al., 1988). Thirdly, the image distortion is a significant problem that according to some, arises from the simultaneous use of videotaping (Stemple, Glaze & Gerdeman, 1995) while others (Hibi et al., 1988; Gould and Korovin, 1994) related it either to the fiberscope lens system or to the unstable lens to object distance.

The rigid endoscopes offers larger, more stable, and often brighter images of the vocal folds, and also avoids some optical artefacts due to stable lens to object distance (Yanigasawa, Godley & Muta, 1987), but it does have disadvantages. It requires a sustained vowel for visualization of the larynx. Approximately 5 to 10% of all patients may not be able to tolerate this rigid examination without the use of topical anaesthesia to suppress the gag reflex. However, Peppard and Bless (1991 cited by Stemple, Glaze & Gerdeman, 1995) and Hirano & Bless (1993) have demonstrated that the use of topical anaesthesia does not alter the resulting visual image of the larynx. In Verdonck-de Leeuw (1998) 12% of the speakers refused to participate in the video-laryngo-stroboscopy because of its uncomfortable invasion and 33% of the speakers had hypersensitivity leading to unsuitable recordings.

As with all the observational methods the conditions under which laryngostroboscopic observations are performed should be clearly controlled, especially concerning the indication of the frequency and intensity used (acoustic parameters which are the determinants of the physiological vocal fold vibration as described in chapter II) (Gelfer & Bultemeyer, 1990; Wendler, 1992; Sulter, Schutte & Miller, 1996; Poburka, 1999).

Laryngoscopic observations are usually performed with a sustained sound and sometimes with running speech. However, Wendler (1992) notes that the running speech task provides little visual larynx information because the glottic positions are highly transient.

In an attempt to establish standards and control the potential biases of this observational assessment some investigators (Hirano, 1981; Gelfer and Bultemeyer, 1990; Faure and Muller, 1992; Wendler, 1992; Gould and Korovin, 1994; Colton et al., 1995; Sulter, Schutte & Miller, 1996; Poburka, 1999) proposed rating systems where the differential diagnosis of the glottal function rests on criteria such as: the degree and shape of the glottal closure, the phase of closure, the extent and amplitude of mucosal waves, the symmetry, with respect to amplitude and phase of the movements of the two vocal folds, phonatory standstill and the presence of no vibrating segments, and the overall regularity of vibration. However, a consensus and standard ratings system was still not achieved (Gelfer & Bultemeyer, 1990, Södersten & Lindestad, 1990).

Gelfer & Bultemeyer (1990) presented a 7 scale rated system for the description of glottic closure (figure IV.1) where (A) represents complete glottic closure; (B) posterior chink; (C) anterior chink; (D) spindle; (E) incomplete; (F) hourglass, and (G) irregular.

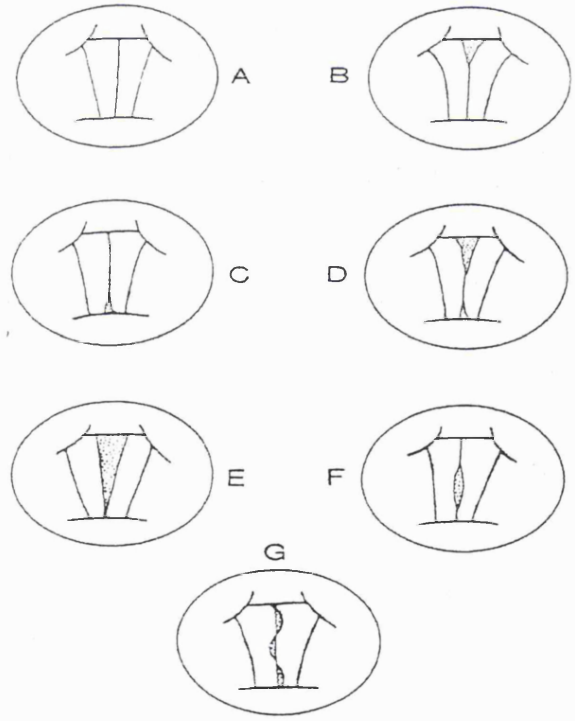


Figure IV.1 – Gelfer & Bultemeyer's (1990) rating system
for judging glottic closure

Södersten & Lindestad (1990) in their study of glottal closure proposed a rating system (shown in figure IV.2) with 6 scale degrees of closure and four types of closure patterns. According to them the points represent (figure IV.2): (1) the complete glottal closure, (2) incomplete closure at cartilaginous glottis; (3) triangular incomplete closure reaching the anterior vocal processes; (4) triangular incomplete closure of the posterior third of the vocal folds; (5) incomplete closure of the posterior two-thirds of the folds; (6) incomplete closure along the vocal folds. The closure patterns were (figure IV.2): '(A) spindle-shaped incomplete closure, closure at the vocal processes; (B) spindle-shaped incomplete closure at the posterior third of the membranous folds, closure at the vocal processes; (C) spindle-shaped incomplete closure at the anterior third of the folds, closure ate the vocal processes; (D) spindle-shaped incomplete closure at the posterior and the anterior thirds of the folds, closure at the vocal processes and at the middle of the membranous portion ('hourglass')'.

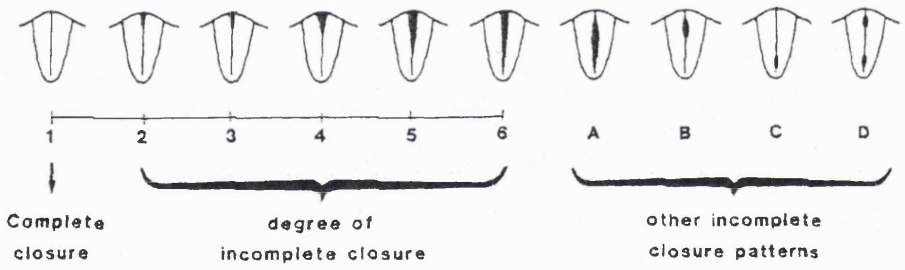


Figure IV.2 – Södersten & Lindestad's (1990) rating protocol for judging glottal closure during phonation

A number of studies intend to create a normative database to expand knowledge of 'normal' larynx apperance and function that can be used as a reference to detect 'deviation' and 'abnormal' laryngeal appearance and function. Still, several potential methodological limitations must be considered (e.g. terminology problems, limited number of subjects and different types of larynx examination techniques) (Casper, Brewer & Colton, 1987; Södersten & Lindestad, 1990; Gelfer & Bultemeyer, 1990; Pemberton et al., 1993; Colton et al., 1995; Sulter, Schutte & Miller, 1996; Elias et al., 1997).

4.4.3 – Electrolaryngography (ELG)

Electroglottography (EGG) was first reported, in the voice field, by Fabre (1957) as a non-invasive technique that gives information about the glottal opening pattern (Baken, 1992; Abberton & Fourcin, 1997; Fourcin, 2000). However, in recent years, several researchers have proved, through the simultaneous use of electroglottography and other techniques (e.g. high-speed photography, xeroradiology and stroboscopy), that EGG is most revealing about vocal fold contact pattern during vibration (Colton, 1996; Abberton & Fourcin, 1997; Howard, 1998; Baken & Orlikoff, 2000; Fourcin, 2000).

For this reason the term electrolaryngography (ELG) instead of electroglottography (EGG) was suggested by Fourcin (Fourcin and Abberton, 1971; 1976; Fourcin, 1982; Abberton, Howard, & Fourcin, 1989, Abberton & Fourcin, 1997; Orlikoff, 1998; Baken & Orlikoff, 2000). However, Orlikoff (1998) states that 'today, this form of electrical glottography is sometimes called electrolaryngography or, more commonly in countries outside the U.K., electroglottography, abbreviated EGG.'

In this study the Laryngograph® device and the ELG technique used was that developed and reported by Fourcin and Abberton (1971). From now on the terms ELG (electrolaryngography) and Lx (the ELG output waveform) and references to research especially using this technique will be used.

Abberton (1976) describes the technique in the following way (pp 45-46):
'*The laryngographic approach enables fundamental frequency measurement to be accomplished in real time and, since the information is derived directly from the vibrating vocal folds, also permits monitoring of certain aspects of mode of vibration, the Lx waveform giving a particularly clear indication of the moment of the vocal fold contact...*'

In this technique, a pair of surface electrodes is positioned superficially on either side of the neck on the wings of the thyroid laminae at the level of the true vocal folds (held to the neck by an elastic throat band), a few centimetres posterior to the angle of the thyroid notch (figure IV.3). A small high frequency current (between 2 and 5 MHz,

typically) at a constant voltage is transmitted from a centre conductor and a guard ring of one electrode to another (that act as a current pick-up). The current flow will be modulated by the vocal folds contact (Fourcin & Abberton, 1971; Fourcin, 1981; Colton, 1996; Abberton & Fourcin, 1997; Orlikoff, 1998; Howard, 1998; Fourcin, 2000).

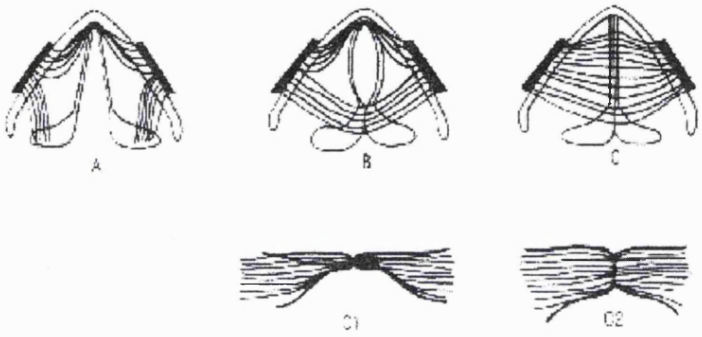


Figure IV.3 - Electrodes applied to the thyroid laminae - From: Baken (1996: 217)

Nowadays, almost all devices, as in the Laryngograph processor and as recommended internationally (Baken, 1992), show, in their output waveform, the increasing conductance due to vocal folds contact represented by an upward deflection of the electrical signal; however in some other devices the opposite tends to be the case (e.g. the Rothenberg electroglottograph, 1992) (Colton, 1996; Howard, 1998; Fourcin, 2000).

4.4.3.1 - Lx waveform interpretation and voice quality types

According to several researchers (Abberton, Howard & Fourcin, 1989; Orlikoff, 1998; Howard, 1998; Baken & Orlikoff, 2000) inferences about the nature of vocal fold motion can be obtained through the direct observation of the Lx waveform features.

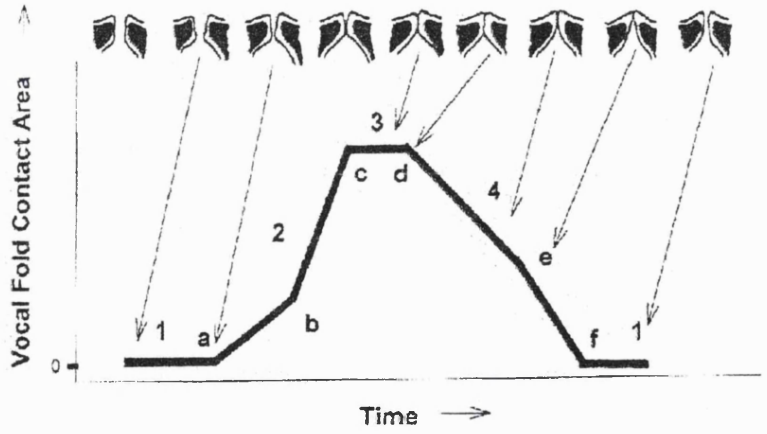


Figure IV.4 - Idealised modal Lx waveform for an adult. From: Baken & Orlikoff (2000: 419)

The Lx waveform represents the maximum separation phase of the vocal folds by a flat base line (number 1 in the figure IV.4) followed by the start of the contacting phase, from the lower vocal folds margins (letter a in the figure IV.4) to the upper margins (letter b in the figure IV.4). Baken & Orlikoff (2000) note differences in the slope of the line (a-b interval in the figure IV.4) according to the configuration of the glottis during the closing phase. The slope will be great when in some instances of closure the folds remain virtually parallel and closure occurs very rapidly along their entire horizontal length, and it will be less when the vocal folds form an acute angle horizontally, giving a divergent shape to the glottis, and causing a *'zippering'* closure.

The closing phase (number 2 in the figure IV.4) is charted by a rapid sharp rise slope implying a sudden conductance increase with the increasing vocal fold contact. Fourcin (1981) believes that vocal fold contact is often initiated by the sudden establishment of a mucus bridge between the two vocal folds, which results in a very abrupt onset of the closing phase. Also, Carlson (1993) writes that *'the speed of the vocal contact can be monitored by looking at the steepness of the rising portion of the Lx waveform'.*

At the end of the contacting phase maximum vocal fold contact is achieved (letter c in the figure IV.4) which is not necessarily synonymous with total glottal closure. For example, the maximal surface area (interval c-d in the figure IV.4) in breathy voice does not correspond to a total glottal closure.

As the subglottal air pressure increases the separating phase begins from the lower margins of the vocal folds (letter d in the figure IV.4) gradually (number 4 in the figure IV.4) until minimal vocal fold contact is reached (letter f in the figure IV.4). In some instances the Lx shape shows, in the separating phase, an inflection point (letter e in the figure IV.4) commonly referred to, as a *'knee'*. Although its exact significance is not know researchers suggests that might correspond to *'the glottal airflow onset'* (Fourcin, 1981) or *'the upward excursion of the vocal lip just before opening, at least under certain circumstances'* (Baken & Orlikoff, 2000).

In modal voice (described in detail in chapter III) there are likely to be waveforms with uniform peaks of large amplitude and distinct asymmetry (figure IV.5A lower trace). Thus, closing progresses more rapidly, (the waveform shows a high steep slope) than the separation phase (with a less steep slope). The duration of no coupling effects between the supra- and sub-glottis is represented by the Lx pulse width. The acoustic response in the speech (Sp) waveform (figure IV.5A upper trace) shows sharp small rapid variations during each cycle as a result of the non coupling effects between the supra- and sub-glottis (Fourcin & Abberton, 1971; MacCurtain & Fourcin, 1981; Fourcin, 1981; Abberton & Fourcin, 1997; Howard, 1998; Fourcin, 2000).

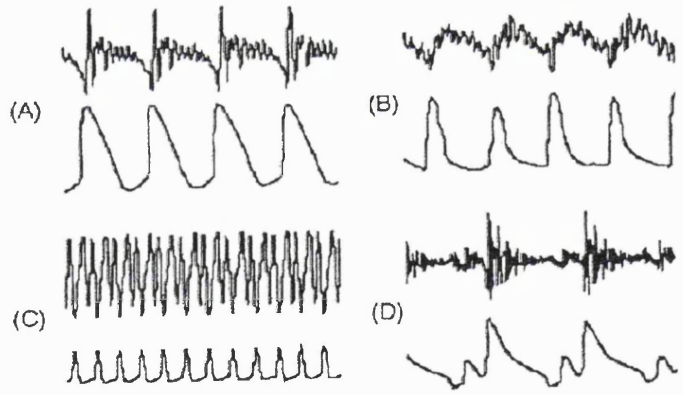


Figure IV.5 - Laryngographic (Lx) waveforms beneath speech (Sp) waveforms; (A) modal voice; (B) breathy voice; (C) falsetto voice; (D) creaky voice.
From: Howard (1998:343)

In breathy voice quality (described in detail in chapter III) the Lx waveform has a less rapid contact phase, the upward line rise is more reduced than in modal voice. The separating phase does not show a well-marked *'knee'* and the duration of the open phase

is much longer than in modal voice, shown as a prolongation of the flat bottom line of the waveform, (figure IV.5B lower trace). The Sp waveform (figure IV.5B upper trace) shows a lower and less well-defined trace in intensity and periodicity than in modal voice due to the coupling effects between the supra- and sub-glottis (Fourcin, 1974; MacCurtain & Fourcin, 1981; Fourcin, 1981; Abberton & Fourcin, 1997; Howard, 1998; Fourcin, 2000; Baken & Orlikoff, 2000).

In falsetto (loft) voice quality (described in detail in chapter III) the Lx waveform is very small in amplitude and with sinusoidal appearance (showing reduced change in contact area and that the contact and separating phases occur with equal rapidity), (figure IV.5C lower trace). The correspondingly Sp waveform (figure IV.5C upper trace) shows that the vocal tract is not so rich in high harmonics as for modal voice because the vocal folds vibration area is reduced (Fourcin, 1974; MacCurtain & Fourcin, 1981; Fourcin, 1981; Howard, 1998; Fourcin, 2000; Baken & Orlikoff, 2000).

The Lx waveform in the pulse voice quality (vocal fry or creaky voice), (described in detail in chapter III), shows a *'dicrotic'* pattern per period; large amplitude main contact peaks alternating with subsidiary small amplitude contact peaks. According to Fourcin (1974) and MacCurtain & Fourcin (1981) the onset of contact phase is always sharp and the decontacting phase is always slow, (figure IV.5D lower trace). The corresponding Sp waveform (figure IV.5D upper trace) shows unmarked traces followed by marked long excitation traces (Fourcin, 1974, 1981; MacCurtain & Fourcin, 1981; Abberton & Fourcin, 1997; Howard, 1998).
Its is worthwhile mentioning that the Lx waveform should not be over estimated as showing pathology because falsetto, creaky and breathy voice qualities or small irregularities are common characteristics seen in the 'normal' Lx waveforms depending on the voice range, the language context (social and emotional) or the speaker's gender, (described in detail in chapter III).

4.4.3.2 - Deviations from the Ideal Lx waveshape

Abnormal waveshapes and irregularities not visible in the idealised 'normal' Lx geometric configuration have been presented in a large number of studies with voice disorders suggesting that structural and functional vocal fold pathologies may affect the Lx waveform (Fourcin & Abberton, 1971; Fourcin, 1974, 1981; Fourcin & Abberton,

1976; Abberton, 1976; Weschler, 1977; Neil, Wechsler & Robison, 1977; Abberton & Fourcin, 1997).

Fourcin & Abberton (1971) illustrate Lx waveforms from speakers with three distinct voice disorders, organic movement disorder (unilateral vocal fold paralysis), non-organic disorder (ventricular vibration) and vocal fold tissue change (namely laryngitis). The authors stress that the Lx waveform shows disturbances alternating with 'normal' features.

Fourcin (1974) shows Lx waveforms features of a 72-year-old man with unilateral vocal fold paralysis, an adult female with laryngitis and a speaker (there is no reference to gender and age) with a pedunculated polyp (before and after surgery) using different speech samples (e.g. utterances and rise-fall intonations).

Wechsler (1977) studied Lx waveforms, during and after different forms of intervention for 20 speakers with a wide variety of voice disorders (e.g. nodules, polyps, cysts, chronic laryngitis, carcinoma, granuloma, unilateral paralysis and interaytenoid weakness) using different speech tasks (sustained vowels, pitch contours and reading).

Neil, Wechsler & Robinson (1977) analysed waveforms and frequency distribution analyses gathered for diagnoses and follow-up of 56 subjects (31 men and 25 women) with a wide range of voice diagnoses. Recordings were made with a standard text. They present several irregular Lx waveforms such as irregular Lx due to asynchronous vocal fold vibration in a patient with vocal fold paralysis, Lx with flattened peaks due to interference with vocal fold closure in a patient with pedunculate polyp, double Lx peaks from a patient with carcinoma of right vocal fold and Lx cycle to cycle interval and amplitude variation in a women with nodules.
It should be noted that Wechsler (1977) in her study found similar Lx waveform features to the last two described in the study cited above for similar pathologies (carcinoma of left vocal fold and bilateral nodules).

In their study, Fourcin & Abberton (1976) demonstrate Lx waveform irregularities resulting from vocal fold contact problems in a unilateral paralysis, Lx peak variability due to obstruction to complete vocal fold contact in a pedunculate polyp

and Lx amplitude variability and longer separating phase from a normal speaker who had an intramuscular injection of atropine to produce artificially dry vocal folds.

It is important to note that aside from the references cited above it has not yet been possible to establish a unique relationship between a characteristic pathological Lx waveform and a given specific pathology, which seems perfectly comprehensible considering the almost limitless forms of voice disorders and counterbalance mechanisms (Neil, Wechsler & Robinson, 1977; Abberton & Fourcin, 1997; Baken & Orlikoff, 2000).

4.4.3.3 - Voice source parameter measurements derived from the Lx waveform

Nowadays, due to advances in technology, ELG information can be not only displayed on an oscilloscope but also displayed and analysed by software programs (Kitzing, 1990). Useful fundamental frequency and contact-based measures can be derived from the detailed information contained in the Lx waveform represented in the figure IV.6.

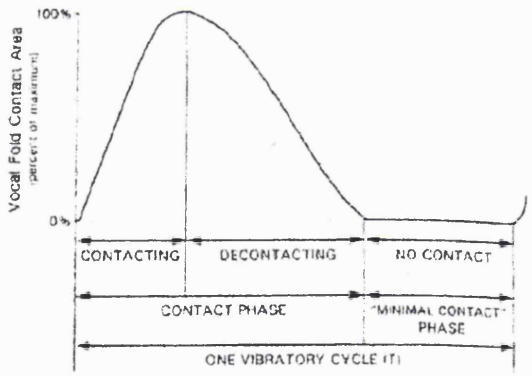


Figure IV.6 - Measurements derived from the ELG waveform
From: Baken & Orlikoff (2000: 426)

The *Fundamental frequency (F0)* of vocal fold vibration can be calculated as the number of cycles (Tx) per second (F0=1/(Tx). The *contact/closed quotient* (CQ), defined as *'the percentage of the entire cycle for which the vocal folds are in contact'* can be derived from the contact phase/vibratory cycle multiplied by 100 (Orlikoff, 1991, 1998; Howard, 1998; Garner & Howard, 1999).

In the present study the ELG data analysis has been carried out using the Speech Studio software from Laryngograph Ltd. This software is used in conjunction with the Laryngograph® processor that carries out an analogue to digital conversion of the time span (Tx) between successive vocal fold closures (Lx) from recorded, or live, data of sustained phonation and/or connected speech. Usually, with these facilities, two signals are simultaneously recorded, a speech signal (Sp) using a microphone and a DAT recorder and a laryngographic signal (Lx) using superficial electrodes (Abberton & Fourcin, 1989).

The Laryngograph (processor and software) and the measures derived from the Lx waveform have been extensively reported (Fourcin & Abberton, 1971; Fourcin, 1974; Abberton, Howard & Fourcin, 1989; Fourcin et al., 1995; Abberton & Fourcin, 1997; Howard, 1998), and also a substantial body of quantitative research has been published about Lx and voice quality types (Fourcin, 1974; MacCurtain & Fourcin, 1981, Fourcin, 1981; Abberton, Howard & Fourcin, 1989), on healthy and dysphonic subjects (Fourcin & Abberton, 1976; Abberton, 1976; Neil, Wechsler & Robinson, 1977; MacCurtain & Fourcin, 1981; Fourcin, 1981, 1986, 1989, 1993; Winstanley & Wright, 1991; Carlson, 1995a; Selby & Wilson, 1997), on children (Gleeson & Fourcin, 1983), on singers (Howard, Lindsey & Allen, 1990; Howard, 1995), on hearing impaired subjects (Abberton et al., 1985; Abberton, Fourcin & Hazan, 1986; Ball, Faulkner & Fourcin, 1990; Fourcin, Abberton & Ball, 1993), swallowing monitoring (Perlman & Liang, 1991; Firmin, Reilly & Fourcin, 2000), voice therapy (Wechsler, 1977; Carding & Horsley, 1992; Carlson, 1993, 1995b), and using simultaneous stroboscopy (Selby & Wilson, 1997; Fourcin, 2000).

Moreover, other researchers mention using the Laryngograph processor in their investigations (Haji et al., 1986; Murty, Carding & Lancaster, 1991; Orlikoff, 1991, 1995; Strik & Boves, 1992; Hertrich & Ackermann, 1995; Howard, 1995; Svec, Schutte & Miller, 1996; Bough et al., 1996; Verdonck-de Leeuw, 1998; Buekers, 1998a; Chernobelsky, 1998; Cheyne, Nuss & Hillman, 1999; Garner & Howard, 1999).

Laryngograph data may be displayed by Speech Studio software in a variety of formats and may be analysed on two distinct levels '*... a macro level, where the input speech data last at least 2 minutes, and a micro level, where short speech inputs are*

used and the outputs from individual F0 analysis can be inspected on a cycle-by-cycle basis' (Howard, 1998:345).

4.4.3.4 - Presentation of data at the micro level using Speech Studio

Analysis of a few seconds of voicing during a continous sustained vowel can be presented in a vocal profile report providing several statistical measures (figure IV.7).

A large body of voice research is based on a microanalysis over the production of a steady sustained sound (as described in the chapter III).

Start at	0.419 s	End at	4.723 s
No. periods	852	Duration	4.305 s
Minimal Fx	182.05 Hz	Maximal Fx	210.09 Hz
Average Fx	197.90 Hz	S.D. Fx	4.50 Hz
Minimal Qx	36.03 %	Maximal Qx	50.00 %
Average Qx	41.50 %	S.D. Qx	0.03 %
Jitter first	1.75 %	Jitter second	0.97 %
Shimmer +	2.74 %	Shimmer -	-2.59 %

Figure IV.7 – Vocal profile analysis report for a 'normal' female speaker during the sustained vowel [i] (From Laryngograph Ltd.)

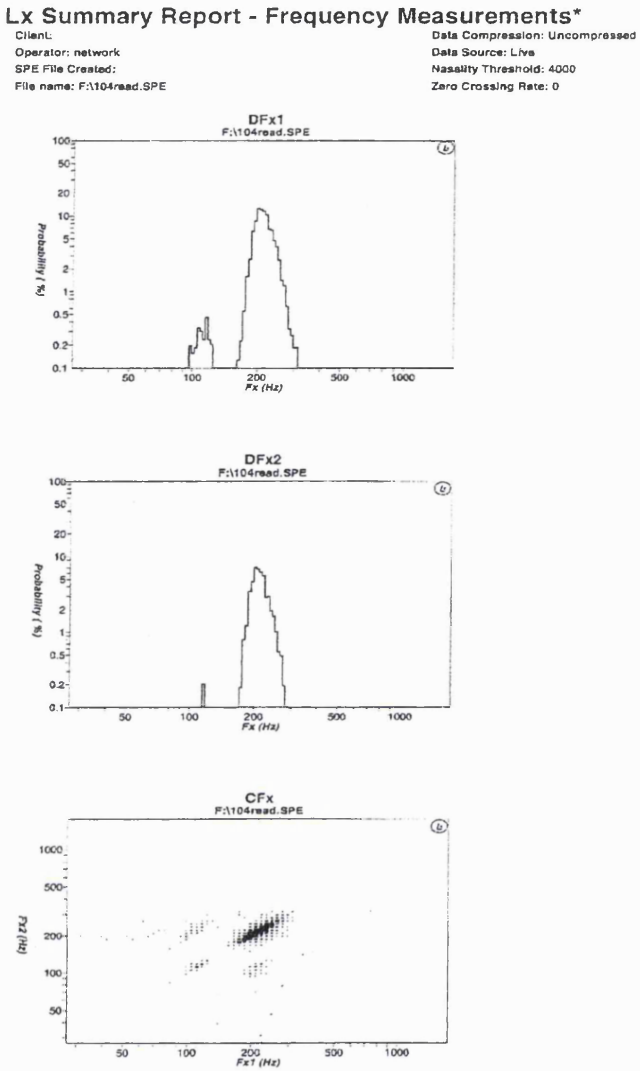
4.4.3.5 - Presentation of data at the macro level using Speech Studio

Fundamental frequency and contact-based data from large samples of speech, using Speech Studio, can be visually represented in the form of histograms and scattergrams by the use of a statistical analysis of the probability distribution of the Lx periods (figure IV.8). Additionally, associated summary descriptive statistics (central tendency, range and regularity) can be calculated from it (Abberton & Fourcin, 1997; Howard, 1998).

The *'distribution of excitation'* (DFx) plots or histograms show the probability of occurrence of a range of Fx values, based on the output from the ELG, during a recorded voice sample. The Fx values of the sample are plotted on a logarithmic scale to

correspond to perceived pitch and the frequency axis is divided into 128 logarithmically equal spaced *'bins'*, a few Hz in width, between 30.52 Hz and 1000Hz. Owing to the use of a logarithmic scale the DFx plots tend to be biased to the higher frequencies, (Fourcin, 1981; Abberton, Howard & Fourcin, 1989; Ball, Faulkner & Fourcin, 1990; Howard, 1998).

Two orders of Fx histograms (figure IV.8), or DFx plots, are displayed. The DFx1 includes all Fx values in the original total sample of Tx values while the DFx2 (second order analysis) includes only consecutive pairs of Tx values that differ by no more than 10% from their preceding neighbours. Some studies in voice report that data from the second order distribution rather than the first order is more representative of vocal fold behaviour because it eliminates extralaryngeal actions (e.g. swallowing, or excessive gross laryngeal movement during recording), that may affect the calculations of Fx parameters (Leff & Abberton, 1981; Ball, Faulkner & Fourcin, 1990; Carlson, 1995a) whilst others used data from DFx1 (Carding & Horsley, 1992) recognising the possibility of high frequency 'noise'.



Graph	Samples	< >	Mean	Mode	Median	Std. Dev.	80% Range	90% Range	Irreg.
DFx1	10152	0, 0	207.65Hz G2#	201.74Hz G2	211.30Hz G2#	0.09	188.6, 245.4Hz 0.38 Octaves	178.5, 256.8Hz 0.52 Octaves	
DFx2	4782	0, 0	207.65Hz G2#	201.74Hz G2	210.13Hz G2#	0.03	190.9, 240.6Hz 0.33 Octaves	185.8, 250.8Hz 0.43 Octaves	
CFx	9805	0, 0							11.24%

Figure IV.8 - DFx and CFx histograms in the Speech Studio analyses for a dysphonic female speaker

Visual DFx features of fundamental frequency range, and distribution of frequencies within that range, have been analysed by several researchers as significant determinants of vocal performance in 'normal' and 'pathological' voices.

Although 'normal' voice DFx distribution is well defined with prominent modes and well defined range as shown by researchers, it should be noted that there is a wide range of 'normal' distributions due to factors such as age, accents and health habits (e.g.

dehydration and prolonged smoking) (Fourcin, 1981; Abberton & Fourcin, 1997; Neil, Wechsler & Robinson, 1977; Fourcin, 2000).

Fourcin (1981) reports that typically for a young woman the low-frequency edge of the distribution is abrupt whilst for older women is less well defined. He also comments that DFx1 and DFx2 in 'normal' voice should show similar shape and similar relative sizes as an indication of 'normal' vocal fold vibration (Fourcin, 1981, 2000).

In 'pathological' voice, distinct irregular features in DFx distribution have been clearly presented. For example, Fourcin (2000) found a flat broader distribution (multimode maximum peaks) in a 35 year-old woman resulting from 'left vocal fold epithelium adherent to the underlying ligament', and in a 80 year-old woman with slight vocal fold asymmetry, while Carlson (1993) found bimodal DFx2 in a 47 year-old woman after surgery to a polyp who had 3 years progressive hoarseness. Abberton & Fourcin (1997) and Fourcin (1981) found secondary peaks at the low frequency range indicative of creaky voice or irregularity while Abberton (1976) and Fourcin (1981), found low frequency components with a broader distribution of peaks in subjects with prolonged smoking habit. It is worthwhile mentioning that Gleeson & Fourcin (1983) report (in their study with children, 9 to 14 years old) that the low frequency components are apparently normal in children.

Fourcin (1981) found restricted frequency range in inadequate vocal fold lubrication while Leff & Abberton (1981) found restricted frequency range with a high degree of *'peakedness'* of the histogram (kurtosis) related to auditory perception of a monotonous voice in depression and schizophrenia (although the depressed patients had a significantly higher mean kurtosis score than the group of schizophrenics).

Neil, Wechsler & Robinson (1977) showed similarity in Lx shape (although different for each pathology) for four men with different laryngeal pathologies (e.g. vocal fold palsy, functional dysphonia, carcinoma and polyp) and a common pattern of curve for 3 women and one man with four different pathologies (e.g. polyp, myxoedema, functional dysphonia and chronic laryngitis).

Although this form of descriptive analysis may be feasible in the clinic once the same observer does the rating, it has the disadvantage of poor correlation of results

from different observers (Wechsler, 1975). However, since 1981 additional quantitative measures of central DFx distribution (mean, mode and median), dispersion (standard deviation) and 80% and 90% Fx range begin to supplement these observations (MacCurtain & Fourcin, 1981).

The CFx scattergram (or crossplot of fundamental frequency values), in which each F0 value is plotted against the previous F0 value (in 64 x 64 matrix of 'boxes' on the logarithmic fundamental frequency axes), is another form of macro analysis particularly useful in F0 irregularity visualisation and calculation (aside from DFx) (as might be seen in figure IV.8). The degree of darkness or density of the trace reflects the frequency of occurrences of particular pairs of values at that position (Abberton et al., 1989; Howard, 1998). The length of the diagonal reflects the range of fundamental frequencies used by a speaker and the degree of scatter indicates the amount of irregularity in the sample. For a 'normal' voice the CFx should show a well-defined, long, dense, narrow plot along the diagonal (Carlson, 1993; Carlson, 1995b; Fourcin, 1981, 2000).

Different types of CFx will be noted in different voice qualities and pathologies. For example, in creaky voice quality where there is a tendency for a low frequency to be followed by a high frequency and vice-versa the CFx distribution will be three parallel lines (Fourcin, 2000) or even a widening of the plot at the low frequency end (Abberton and Fourcin, 1984), also a very wide scatter along the diagonal through the frequency range illustrates limited periodicity of vocal fold vibration (Carlson, 1993), a reduced diagonal length or with breaks may be seen in the puberphonic voice (Carlson, 1995b).

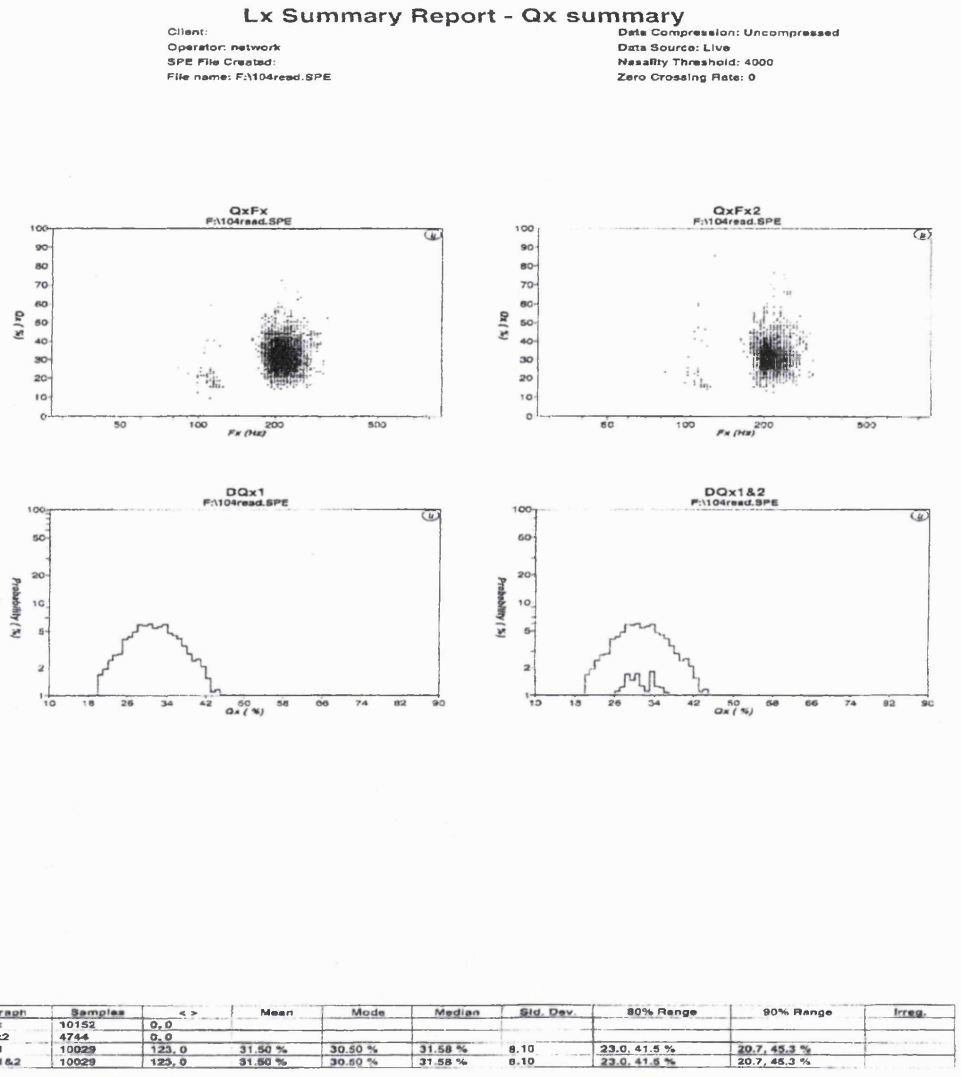
An estimate of voice regularity may be obtained through the calculation of '% Irregularity' in the CFx which reflects the proportion of Tx samples that differ from their adjacent neighbours by more than 10%.

No norms are yet available for the percentage of irregularity. Carlson (1995b) estimates based on clinical observation that a normal voice may present between 5% and 10% irregularity; however, gender, age and type of speech sample used are not mentioned. Carlson (1995a) reported the mean of 13.9% for 'normal' speakers during reading and 18.4% during conversation and notes that within 'normal' speakers percentage of irregularity ranges between 4 and 6% and between 40 and 60% were

found (showing a high intersubject variability). Bowen (1997) reported the mean percentage of irregularity of 5.83 (range, 2.36-7.41) and 5.60 (range, 3.17-10.32) for 6 'normal' males for 6 'normal' females respectively (with an age range between 35 and 55 years old) during the reading of 'Arthur the Rat'.

For puberphonia, before therapy, Carlson (1995b) showed that two male subjects (30 and 31 years old) presented 37.6% and 18.4% of irregularity during reading of the first paragraph of the Rainbow Passage (it should be noted that one subject was not an native English speaker). For laryngeal cancer male speakers (Carlson, 1995a) found that the percentage of irregularity during reading (29%) and speech (31.7%) was higher than that for 'normal' male speakers (13.9% and 18.4% respectively). However, she stresses that within the 'normal' speakers some showed low percentages of irregularity (4-6%) whereas some presented between 40 and 60%. The pointed reasons for these are the neck thickness, the tendency to laryngeal movement during speech and 'creaky voice' use.

Another graphical representation of voice quality is based on a contact-related measurement namely QxFx (contact quotient) scattergrams, where QxFx is plotted against the corresponding Fx value for every cycle in the sample (figure IV.9). Also, two plots may be obtained, first (QxFx1) and second (QxFx2) order distribution. The Qx varying extension and density shows the distribution of values, which is useful to measure the overall production of voice.



Graph	Samples	< >	Mean	Mode	Median	Std. Dev.	80% Range	90% Range	Irreg.
QxFx	10152	0, 0							
QxFx2	4744	0, 0							
DQx1	10029	123, 0	31.50 %	30.50 %	31.58 %	8.10	23.0, 41.5 %	20.7, 45.3 %	
DQx1&2	10029	123, 0	31.50 %	30.50 %	31.58 %	8.10	23.0, 41.5 %	20.7, 45.3 %	

Figure IV.9 - QxFx scattergram in the Speech Studio analyses for a dysphonic female speaker

Abberton & Fourcin (1997:132) presented a QxFx comparison between a 'normal' and a pathological speaker (Reinke's oedema), using the same speech sample (a 2-minute reading passage) showing that the pathological speaker had a more restricted QxFx range than that for the 'normal' speaker. Also, in a study with 44 young female students Abberton & Fourcin (2001) showed three different QFx patterns during connected speech. 43% of the students presented increasing breathiness with voice pitch, 25% decreasing breathiness with voice pitch and 32% breathiness independent of voice pitch.

According to Fourcin (2000:293) the contact quotient ratio values result from *'period by period ratio estimates of Lx closure width, 70% down from each positive*

peak, to the instantaneous period Tx.' Thus, the lower the QxFx the more 'breathy' perception of voice quality because of the coupling effects between the supra- and sub-glottis (Abberton & Fourcin, 1997).

4.4.3.5 – Data for vocal fold contact quotient

There is a paucity of research in the area of contact quotient derived measures comparing other measures (Orlikoff, 1991, 1998; Howard, Lindsey & Allen, 1990; Howard, 1995, Rossiter, Howard & Comins. 1995; Bowen, 1997; Verdolini et al., 1998). However, some found no evidence of gender effects (Orlikoff, 1998); others found a positive correlation with vocal intensity (Orlikoff, 1991) and with level of vocal training (Howard, Lindsey & Allen, 1990; Rossiter, Howard & Comins, 1995).

The studies of Howard & Lindsey (1988), Lindsey & Howard (1989) and Howard, Lindsey & Allen (1990) found that non-vocally trained subjects exhibit relatively lower contact quotient than that for trained subjects. However, they stress that this implies a vocal training of years (Rossiter, Howard & Comins, 1995).

Orlikoff (1991) reported a significant positive correlation between the contact quotient and vocal intensity in ten healthy men (mean age, 31,8 years) during the sustained vowel [a]. This meant that the magnitude of contact quotient increases with increasing vocal intensity. The reported mean values were: (a) .52 ± .07 (soft voice); (b) .57 ± .07 (moderate voice) and (c) .61 ± .06 (loud voice). These results are corroborated by Hirano & Bless (1993) and Dejonckere (1998) based on the fact that in 'healthy' subjects a higher subglottal pressure produces large vocal fold amplitude and an increased closed phase.

Murty, Carding & Lancaster (1991) found that during the production of the sustained vowel [i] in ten healthy men (mean age: 22.3) showed a closed glottis for approximately one-half of the total glottal cycle (the reported open quotient was .548 ± .054).

Rossiter, Howard & Comins (1995) found similar contact quotients in three men (mean age: 28.33) and six women (mean age: 48.17) before a six week course in vocal expression. Measures were obtained from the oral reading of 'The voyage by Katherine

Mansfield' and the results were 45.61% for men and 43.47% for women. Also, no evidence of improvement of the contact quotient after the course was found and the explanation presented for that was the short course duration.

Bowen (1997) reported the mean QFx of 45.17 (range, 42.5-50.5) for 6 'normal' adult males and 44.67 (range, 41.5-49.5) for 6 'normal' adult females during the reading of 'Arthur the Rat'. Gender did not exert a main effect on QFx.

Orlikoff (1998) suggests a variation of the contact quotient between 40% and 60% for normal voice and inexistence of a significant gender effect on this measurement.

Verdolini et al. (1998) found relatively lower contact quotient for 6 vocally trained subjects with nodules than that for 6 vocally trained healthy subjects and that contact quotient distinguishes resonant ('easy' phonation used in performing acts) from pressed voice but not from 'normal' and breathy voice. The reported data for the different types of voice is shown in table IV.1.

Table IV.1 – Verdolini et al.'s (1998) data

		Type of voices			
Subjects		Pressed	Resonant	Normal	Breathy
Healthy	[a]	57.2 ± 4.4	55.5 ± 7.9	53 ± 6	49.3 ± 12.3
	[i]	57.8 ± 6.3	54.7 ± 6.5	53.1 ± 7.5	45.3 ± 9.5
Nodules	[a]	57.4 ± 5.2	44.5 ± 9.5	44.6 ± 6.8	42 ± 3.4
	[i]	52.9 ± 5.4	46.7 ± 7.3	46.4 ± 4.6	43.5 ± 6.6

4.4.3.6 - Advantages of ELG

Although the Fx calculation can be based on acoustic analysis (with cepstral analysis being the most reliable) in fact Fx based on ELG measures has the advantage of being a less complex and faster technique than the cepstral techniques and also its result is associated with the effects of the source on voice (the closing and closed events in the cycle of vocal folds vibration) without the imposition of the filter (vocal tract) (Haji et al., 1986; Bough et al., 1996).

Another strength of ELG is the fact of not being invasive so it is acceptable to a wide range of speaker populations (especially those with strong gag reflexes); however, there are some subjects for whom it is not possible to gathered reliable ELG signals (Colton & Conture, 1990; Winstanley & Wright, 1991; Buekers, 1998a; Verdonck-de Leeuw, 1998).
Colton & Conture (1990) report 15% of their patients with unreliable ELG signals whilst Verdonck-de Leeuw (1998) reported 21% of their patients with unreliable ELG signals. Buekers (1998a) obtained 8 (40%) unreliable Lx waveforms in the twenty women with voice fatigue whereas for the 20 'normal' women all the ELG signals were reliable. Winstanley & Wright (1991) have gathered unreliable Lx waveforms in 29 of 100 normal subjects.

It does not interfere with the process of speaking (Fourcin & Abberton, 1971), therefore allowing long-term speech analysis, and Fo contours can be measured and displayed in real time (Fourcin, 1974).

Its acoustic noise immunity is a special advantage for data gathering in many ordinary work environments (Abberton & Fourcin, 1997; Fourcin, 2000).

It is valuable in the study of voice disorders once its result is directly related to the closing and closed phases of vocal fold vibration.

4.4.3.7 - Limitations and problems of ELG

In spite increased use of ELG, clinically and experimentally, due to its simplicity in application and interpretation, both researchers and clinicians argue about its reliability because of the inherent limitations and problems in data collection (devices and procedures) (Childers et al., 1990; Colton & Conture, 1990; Watson, 1995; Bough et al., 1996; Baken & Orlikoff, 2000). Limitations of the technique have been fully described by Colton & Conture (1990) according to three potential sources of difficulty; *'instrumental'*, *'procedural'* and *'subjects'concern'*.

Before the ELG data can be considered reliable and useful it is important to consider that although the main contribution to the detailed Lx waveform comes from

the fast changes in vocal fold contact it also partially results from the slow movement of the other structures of the vocal tract (e.g. the tongue, epiglottis, ventricular folds and the laryngopharynx), the oesophagus, the extra laryngeal musculature of the neck, superficial fascia, and the skin (Baken & Orlikoff, 2000). For this reason, Fourcin (1981) in his ELG processor uses a high pass filter to extract an Lx waveform, where the Gx movements (term used by Fourcin (1974) for the output resulting from low , 'gross', frequency variations) are not much in evidence, and there is an automatic gain control (AGC) to adjust the instrument sensitivity to compensate for baseline impedance differences (e.g. different neck resistance among individuals and differences in electrode placement).

Researchers (Colton & Conture, 1990; Howard, 1998; Baken & Orkiloff, 2000) have noted, however, that both the high-pass filtering and the automatic gain control function used in ELG devices may introduce artefacts into the shape of the open phase of the Lx waveform which instead of being flat, if there were no filtering, increases in slope as the high-pass cut-off frequency rises. However, according to several researchers (Fourcin, 1974, 1981; Baken & Orlikoff, 2000) the usefulness of Lx is based on its representation of the closure events.
Aside from the comments above, filtering can also invalidate the ELG *'contact quotient'* measurement that is derived by voice analysis software. Baken & Orlikoff (2000) point out that the implications of filtering on the waveform geometry are *'a sharper peak and the decontacting phase to become less 'rounded' and more linear'*.

Although it is known that the automatic gain control accounts for potential distortion because it increases the signal-to-noise ratio and the amplitude of the Lx waveform (which for that reason should be interpreted with caution) other effects are not yet predicted (Howard, 1998; Baken & Orlikoff, 2000).

A final instrumental problem focused on in the literature is the low phase distortion that the analogue tape recorder imposes (Fourcin & Abberton, 1971; Fourcin, 1981; Abberton, Howard & Fourcin, 1989). Thus it is possible that the Lx waveform appearance shows the presence of low frequency irregularity leading to erroneous conclusions about the real vocal fold motion (for example as might be expected from 'creaky' voice quality). This can be avoided with the use of digital machines such as digital audiotape (DAT) recorders for data gathering.

Recording procedures for ELG related to the correct placement and maintenance of the electrode skin contact is another potential source of bias because they are crucial factors to ensure a reliable Lx waveform output (Abberton, Howard & Fourcin, 1989; Colton & Conture, 1990; Baken & Orlikoff, 2000). Although there is no standard rule for the exact location of the electrodes, owing to subjects' neck variations, this may be monitored on an oscilloscope during trial phonation. Incorrect electrode positioning is indicated by low signal amplitude, which can result in changes in the shape of the waveform: an important factor since extraction of Tx values relies on the shape of the waveform.

Maintenance of adequate electrode placement may be difficult because of the vertical larynx height changes resultant from speech, swallowing and head movement. This can produce an irregular shape and density variability of the current field within the electrodes adding uninterpretable features to the Lx output. Although electrolaryngographs are designed with high-pass filters to eliminate much of the baseline shifts, some artefacts may remain. Consequently, the electrodes must be firmly fixed at a proper place to ensure constant neck-electrode contact (Abberton, Howard & Fourcin, 1989; Colton & Conture, 1990; Howard, 1998; Baken & Orlikoff, 2000).

An interesting finding reported by MacCurtain & Fourcin (1981) is the existence of a double signal distorting the Lx waveform caused by rising blood pressure with a pulse response by the carotid artery (either because of the tightness of the electrode neckband or the heat of the recording chamber).

Colton & Conture (1990) reported that the individual's laryngeal status and functional differences (described in detail in chapter II) might be responsible for the problems in ELG data gathering, especially in adult females. They state that, on the whole, adult females produce poorer ELG signals than adult males. The possible explanation is the variability in the larynx dimension (females have smaller vocal folds and shorter thyroid laminae area than men, thereby small area for the current to pass through), the angle of the thyroid cartilage (women have wider angle of the thyroid cartilage than men causing the electrodes to be not so close to parallel as possible, thereby creating a lower signal-to-noise ratio) and adipose tissue (women usually have

more adipose neck tissue than men, thereby greater resistance to the current flow, and occasionally it is impossible to obtain a robust Lx signal).

Another constraint relates to the risk of artefacts in the ELG waveform due to skin-surface conductance due for example to sweating, (Fourcin & Abberton, 1971; Colton & Conture, 1990; Baken & Orlikoff, 2000). In order to avoid this the Laryngograph processor (used in this study) is equipped with a guard ring at earth potential around the electrodes (Abberton, Howard & Fourcin, 1989; Fourcin & Abberton, 1971).

Finally, Abberton, Howard & Fourcin (1989) mention the possible Lx waveform distortions due to the electrical connection of the equipment to the mains supply. The suggested solution to solve this problem is to use the equipment battery powered.

4.5 - Voice-related quality of life assessment

Since Hipocrates quality and its evaluation has been constantly in clinicians' minds in order to get the best benefit for the patient (Winding, Clamp & Ohmann, 1993) and nowadays references to quality of life are numerous across a range of diverse medical conditions (Smith et al., 1996; Morton & Tandon, 2000) but still few regarding the impact of voice. Nevertheless, a review of the published literature, within the past years, reveals that many authors stress the importance of voice disorder consequences on the subject's life and propose that the information should be gathered during the interview (Prater & Swift, 1985; Verdolini, 1994; Morrison et al., 1994). However, it is a personal, subjective assessment that needs to be self-reported by using a standardised instrument (usually questionnaires) because observers cannot rate it accurately.

What is included in the category of 'outcomes', defined by Donabedian (1985 cited by Fratalli, 1998b) as either favourable or adverse changes in the actual or potential health status of persons, depends on how narrowly or broadly one defines 'health' and the operationalisation of what is to be measured.

A reasonably comprehensive measure should include domains that are generally accepted as major contributors to health-related quality of life such as physical, social,

psychological functioning, family and social support, and disease- and treatment-related symptoms based on the conceptual framework of health defined by the World Health Organization (1958) as *'not merely the absence of disease or infirmity, but a state of complete physical, mental and social well-being'* (Berzon, 1998; Morton & Tandon, 2000). Nevertheless, within the clinical trial research on the health-related quality of life according to Cella and Tulsky (1990 cited by Berzon, 1998) '...it refers to patients' appraisals of their current level of functioning and satisfaction with it compared to what they perceive to be ideal'. Therefore, those instruments, commonly referred to as functional assessment were designed in accordance with one or the three concepts of the ICIDH (WHO, 1980) classification of impairment, disability and handicap (defined in detail at the end in table IV.5).

Although domains such as occupational fuctioning, economic status, body image and sexuality are considered to contribute to quality of life they are not usually included in quality of life assessments (Morton & Tandon, 2000). For example, in a large group study of 174 dysphonic adults and 173 healthy adults Smith et al. (1996) found significantly higher adverse voice-quality of life effects in the dysphonic subjects in work, social, psychological, physical and communication domains. Their self-adminstered lifestyle questionnaire includes information about the functional impact of voice in domains such as, work, social, current occupation, work related changes resulting from the symptoms, as well as psychological and emotional aspects such as self-esteem.

Scott et al. (1997) found 467 self-reported problems (categorised according to the ICIDH (WHO, 1980)) in a group of 133 dysphonics (43 men and 90 women, mean age 54 years) were 60% impairments, 26% disabilities and 14% handicap related. The handicaps encompass social, psychological and employment problems. The authors do not present any data on the patients' professional activity (the age range is between 18 and 80 years old) and the aetiology of the dysphonia.

Unfortunately, there is only a modest number of studies dealing with the development of standardised voice-related quality of life (VQoL) instruments (Smith et al., 1996; Jacobson et al., 1997; Hogikyan and Sethuraman, 1999) and their application (Benninger et al., 1998; Stewart et al., 1998; Rosen and Murry, 2000; Hogikyan et al, 2000; Rosen et al., 2000). A possible reason for this may be due to the multifactorial

nature of the conceptual basis of quality of life and consequently the complexity of measuring it. Another possible limitation may be due to the fact that, within the medical sciences, investigation of the quality of life is focused initially on the quality of survival in life threatening conditions such as cancer and, as Verdolini et al. (1998) stated, voice disorders (as other communication disorders) may be viewed as less serious conditions with less visible constraining effects. In fact one of the first attempts to quantify VQoL self-assessment was developed by Llewellyn-Thomas et al. (1984 cited by Carlson, 1995a and Verdonck-de Leeuw, 1998) for patients with laryngeal cancer. Based on this disease-specific instrument that rated the voice symptoms and function on a linear analogue scale other authors (Carlson, 1995a and Verdonck-de Leeuw, 1998) created similar scales also for patients with glottic cancer treated with radiotherapy.
Surprisingly, Smith et al. (1996) found that patients' voice-related quality of life reports were in a similar or worst range of severity than those reported by patients with medical conditions generally considered more serious (e.g. cancer, rheumatoid arthritis, asthma).

Meanwhile, maybe because the concept of voice problems affecting the subjects' well-being is deeply embedded in the clinician's consciousness and as Verdolini (1994) estimates, 50%-60% of clinical patients with voice disorders report social, communicative, physical and psychological problems as a result of the voice disorder, several studies started to show the importance of self multi-item questionnaires and/ or scales as an outcome measure (Smith et al, 1996; Jacobson et al., 1997; Enderby, 1997; Hogikyan and Sethuraman, 1999). Their application to the general range of voice problems (Smith et al., 1996; Jacobson et al., 1997; Benninger et al., 1998), to specific voice professionals (Sapir, Keidar & Mathers-Schmidt, 1993; Sapir, Mathers-Schimdt & Larson, 1996), to specific laryngeal diseases (Carlson, 1995a; Verdonck-de Leeuw, 1998; Epstein, 1999; Hill et al., 2000), and to the measurement of the effectiveness of voice therapy and surgery techniques (Carding, Horsley and Docherty, 1998; Verdolini, Ramig & Jacobson, 1998; Hajioff et al., 2000; Hogikyan et al., 2000) started to grow in the health field.

Both abovementioned generic or specific voice-related health quality of life instruments have advantages and limitations and must be selected within the context of a particular study. Nevertheless, generic instruments allow cross comparisons but may be not sensitive to a specific condition or intervention whereas the specific instruments are the reverse (Benzon, 1998; Hill et al., 2000).

Verdonck-de Leeuw (1998) created a self-rating scale of vocal performance (in her study with laryngeal cancer patients) with thirteen questions covering vocal abilities (e.g. can you make a telephone call regarding your voice?) and social situations (e.g. have your social contacts changed because of your voice?) using a 7-point rating scale analogue (e.g. 1-never and 7-always). Results showed that patients vocal performance improved after radiotherapy treatment but remained worse than the 'normal' control vocal performance. Predictability of vocal performance (multiple regression analysis) showed that continuing smoking after treatment, radiation dose and vocal fold stripping influenced vocal performance following radiotherapy. Nevertheless, the author stresses that '...no attempt was made to determine psychometric aspects, like validity, reliability and feasibility of the vocal performance scales.'

Similarly, Carding & Horsley (1992), within the use of a battery of assessment techniques to test speech therapy effectiveness in the treatment of 30 non-organic dysphonics, developed a self-rating questionnaire of vocal performance. Three possible treatment groups were created (direct and direct therapy and no treatment) for a period of 8 weeks. The questionnaire, comprising 12 questions, was a rating scale answer format with scores varying from 12 (normal) to 60 (severely). The authors reported that the questionnaire was previously trialed in a sample of 10 subjects (without reference to their laryngeal or voice status), used on several occasions, and response reliability was statistically tested (using Cronbach's α). Results showed no statistically significant difference between the three groups in the pre-treatment assessments but a significant difference was found between the groups in the review assessment (one month after 8 weeks of treatment). The biggest difference was between the group without treatment and the group with direct treatment. The questionnaire was neither in this study or in the subsequent study (Carding, Horsley & Docherty, 1998) used in comparison to a control group.

An important contribution to this field was the development of a VQoL self-assessment ordinal scale, the VHI (voice handicap index) by Jacobson et al. (1997) (in appendix E, and described in detail in chapter V). The VHI produces a non-standardised index of the subject's self-rating degree of his/her voice-related problems in three domains (emotional, functional and physical). According to the authors the items comprising the questionnaire were selected from case history interviews with patients

over the past 7 years to ensure that the scale had both content and face validity. In its development the questionnaire was used in a wide range of voice disorders (e.g. mass lesions, neurogenic, laryngectomy, musculoskeletal tension, inflammatory and atypical) because the purpose was to create a scale that could have a widespread application. The VHI proved to have a strong internal consistency, good test-retest reliability and correlation with patients' judgment of voice disorder severity. Domain scores have been shown to correlate between 0.70 and 0.93. Although scores with reference to a control group have not yet been reported it has been reported for self-rated voice severity (table IV.2).

Table IV.2 - VHI scores (Jacobson et al., 1997)

	N	Functional M±SD*	Physical M±SD*	Emotional M±SD*	Total M±SD*
Dysphonia severity					
Mild	13	10.07±1.99	15.54±1.97	8.08±2.31	33.69±5.60
Moderate	27	12.41±1.38	18.63±1.37	13.33±1.61	44.37±3.88
Severe	23	18.30±1.50	22.78±1.48	20.30±1.74	61.39±4.21

* Mean±Standard deviation

A study by Benninger et al. (1998) used the VHI tool with 260 patients with a variety of laryngeal conditions but they only reported the degree of VHI correlation with the Medical Outcomes Trust Short Form 36-Item (SF-36). They found that mental health, general health and role functioning SF-36 domains correlate significantly with the VHI domains and total score. Moreover, they stress that patients with vocal fold oedema show a significantly higher VHI total score and functional and physical subscores (meaning worse voice impact) than patients with vocal fold masses and paralysis (although data is not reported) and no significant difference in the emotional subscore. No significant differences were found between the patients with vocal fold masses and vocal fold paralysis. This study also shows that the voice disorders impact is perceived by the subject as a significant disability in comparison to other medical conditions (e.g. angina pectoris, sciatica and chronic sinusitis).

Stewart, Chen & Stach (1998) also used the same tools as Benninger et al. (1998) in a retrospective study with 80 laryngeal cancer patients who had been treated with laryngectomy or radiotherapy or both. All patients were men and the mean age was

65.3 years (range, 44-86 yrs). They found significant correlation between the SF-36 items that were represented on the VHI such as social functioning (VHI emotional) and role functioning (VHI physical). They also found that laryngectomy patients presented higher VHI scores than the radiotherapy patients. Demographic factors (eg. age, lack of employment and lower level of education) did not predict significantly worse VHI scores. However, patients living alone presented significantly lower VHI functional subscale scores. Data (shown in table IV.3) demonstrate that voice handicap is not the most important or the only dimension affecting the quality of life in these patients which confirms the findings of other studies (Terrell, Fisher & Wolf, 1998).

Table IV.3 - VHI scores (Stewart, Chen & Stach, 1998)

	N	Functional M±SD*	Physical M±SD*	Emotional M±SD*
Cancer patients who				
Did not undergo Laryngectomy	24	8.2±8.3	11.0±9.1	5.8±7.1
Who underwent Laryngectomy	56	14.9±7.3	11.2±6.8	11.0±8.3

* Mean±Standard deviation

A more recent study with the VHI tool in 106 singers (professionals and recreationals) and 369 non-singers with voice complaints by Rosen and Murry (2000) showed that singers score significantly lower (less severe) when compared to non-singers, and professional singers also show lower significant scores when compared to recreational singers (see table IV.4). The authors explain that this difference may be related to the nature of the VHI tool and/or the specificity of singers' voice problems. As has already been mentioned the VHI purpose is to be relevant to a wide spread of laryngeal pathologies and non specific voice problems.

Table IV.4 - Mean VHI scores (Rosen & Murry, 2000)

	N	Functional Mean	Physical Mean	Emotional Mean	Total Mean
Professional singers	73	7.0	14.7	9.3	31.0
Recreational singers	33	11.0	19.3	12.9	43.2
Nonsingers	336	16.8	21.3	15.2	53.2

Rosen et al. (2000) indicate that VHI is sensitive to subject's perception of voice change following different types of intervention (surgery, medical and voice therapy) therefore a useful instrument to measure treatment efficacy. This study included patients with unilateral vocal fold paralysis (UVFP), muscle tension dysphonia (MTD), vocal fold polyp and vocal fold cyst grouped in VFP/C.

Wyuts et al. (2000) found a significantly high correlation (r=-0.79) between VHI and the DSI (dysphonia severity index), an objective measure of vocal quality based on a multiparameter approach. The study involved 68 'normal' speakers and 319 dysphonics. Unfortunately VHI data for the total and subscales scores is not reported.

As the same time as the VHI publication (Jacobson et al., 1997) Enderby (1997) developed a standardised assessment to assist Speech and language therapists (SLT) in diagnosing patients and monitoring changes gathering information related to the ICIDH (1980) typology (impairment, disability and handicap) and well-being named Therapy Outcome Measures (TOM). The study involved eight districts (in England), more than 1,000 patients (with several communication disorders such as fluency disorders, dysphonia, aphasia and dysphagia) and 132 SLT. The TOM was shown to be reliable, valid, sensitive to change and may be used in comparison of services. For dysphonia TOM has four domains (impairment, disability, handicap and well-being) and a five point scale from 0 (the worst, for example, in impairment, severe persistent dysphonia, unable to phonate) to 5 (the best, for example, in impairment, no dysphonia, appropriate/modal consistently used). Neverthless no published article using this tool has been found.

Another voice-related quality of life (V-RQOL) tool of 10 items and a rating scale of five points (0-none, no problem and 5- problem is as 'bad as it can be') that assess social-emotional and physical functioning was developed by Hogikyan & Sethuraman (1999). Equations are used to calculate a standard score from the raw scores. Although the authors recognised that items are similar to those of VHI they criticised the VHI for being focused on the handicap rather than on the quality of life, that no standardised score is calculated for VHI and also pointed out that the burden (time of responsiveness) is lower in the V-RQOL. Neverthless, the authors felt that

although the brevity of V-RQOL made it easy to administer it was less comprehensive or diagnostically discriminating than a longer one.

The only reported data (Hogikyan et al., 2000) of the V-RQOL after the validation study (Hogikyan & Sethuraman, 1999) is on patients with vocal fold paralysis who had undergone to a type I thyroplasty in comparison to untreated patients and healthy speakers. A significantly higher V-RQOL score was found for the treated patients in comparison to the nontreated patients although lower than healthy speakers.

From the abovementioned it can be noted that the ICIDH (WHO, 1980) encouraged the development of some voice-related quality of life assessment tools within the context of specific studies where that was not the main aim in the study (e.g. Carding & Horsley, 1992; Carlson, 1995a and Verdonck-de Leeuw, 1998) or the specific aim was to evaluate the voice-related quality of life of speakers or to study the psychometric properties of the developed questionnaire (e.g. Smith et al.'s (1996) Quality of life questionnaire; Enderby's (1997) Therapy outcome measures; Jacobson et al.'s (1997) Voice handicap index; Benninger et al. (1998); Hogikyan and Sethuraman's (1999) Voice-related quality of life; Epstein (1999); Hill et al. (2000); Rosen et al. (2000), Rosen & Murry (2000), Hogikyan et al. (2000)). However the ICIDH (WHO, 1980) typology was recently modified to ICIDH-2 Beta 2 (WHO, 1999) and the 'disability' level is replaced by 'limitation of activities', and 'handicap' is replaced by 'restriction in participation' (table IV.5) and consequently new voice quality of life assessment tools have been proposed (eg. VAPP- voice activity and participation profile by Ma & Yiu, 2001).

Table IV.5 – WHO (1980; 1997; 1999) conceptual frameworks

ICIDH (WHO, 1980)	ICIDH-2 Beta 1 (WHO, 1997) and Beta 2 (WHO, 1999)
Disease or disorder/pathology	*Disease or disorder/pathology*
Etiology/signs or symptoms.	Etiology/signs or symptoms
Impairment	*Impairment*
Any loss or abnormality of psychological, physiological, or anatomic structure or function.	Represents a deviation from the accepted standards (as judged by a specialist) in the biomedical status of the body and its function (WHO, 1999).
Disability	*Activities/ Activity limitation*
Any restriction or lack of ability to perform an activity in a manner or within a range considered normal for a human being. The functional consequence of an impairment mabnifested in integrated activities represented by tasks and skills.	The nature and extent of functioning at the level of the person. Refers to a person's integrated activities associated with everyday life (Fratalli, 1998a). Limitation is the discordance between the observed and the expected performance on the basis of a norm for the persons without the specific health condition (WHO, 1999). In the voice field can be analysed as constraints in voice activities (Ma & Yiu, 2001).
Handicap	*Participation/ participation restriction*
A disadvantage resulting from an impairment and/or disability that limits or prevents the fulfilment of a role that is normal (depending on age, sex, social, cultural factors). The social consequence of an impairment or disability, defined by the attitudes and responses of others.	The nature and extent of a person's involvement in life situations in relation to impairments, activities, health conditions and contextual factors. Refers to societal phenomena (Fratalli, 1998a). Restriction is assessed against an accepted population standard (WHO, 1999). In the voice field can be interpreted as a reduction or avoidance of voice activities by the individual (Ma & Yiu, 2001).

Verdolini, Ramig & Jacobson (1998) provide an excellent review of the published studies on voice therapy outcome from the 1940s to the 1990s. The authors mention that although the number of published studies on the results of voice therapy is modest there is a clear qualitative (from an 'expert opinion' to quasi-experimental and experimental designs) and quantitative (large N studies) improvement.

Chapter V - Methodology

This dissertation is a prospective exploratory study based on a cross-sectional research design, which examined 52 patients with dysphonia who attended the ENT unit at the S. José Hospital, Lisbon. In addition, a sample of 57 volunteers was recruited for control purposes. The study was carried out during a period of nine months (from January to September 2000).

The primary objectives of the present study were (a) normative, to establish tentative normative data for Portuguese speakers. Baseline measurements were obtained for each subject in different voice conditions; (b) comparative, to investigate whether statistically significant differences exist between two groups with different voice conditions, on performance of different speech tasks and self-rating voice impact; and (c) predictive, to test for the variables (subjects' case history, ENT examination, SLT assessment and subjects' self-rated stress and psychosocial voice impact) that best predict aspects of the voice data (fundamental frequency (F0) and contact closed quotient derived measures) and the subjects' self-rated psychosocial voice impact data.

The major questions to be answered were:

(1) What are the quantifiable voice characteristics of Portuguese speakers? (a) are there any statistically significant differences between two groups with different voice conditions on performance of different speech tasks? (b) which variables contribute most importantly to predicting variance in the quantifiable voice measures?

(2) What is the dysphonic Portuguese speakers' self-rated voice psychosocial impact? (a) are there any statistically significant differences compared to a group of healthy Portuguese speakers? (b) which variables contribute most importantly to predicting variance in the quantifiable data?

A previous formal acceptance from the administrative board and ethical committee from the São José Hospital in Lisbon was required. Also, ethical approval by the University College London/University College Hospital committees on the ethics of human research was required before the experiment could take place.

An assessment protocol had been developed and tested in a pilot study (between June to October 1999). Minor corrections were made to this data form after evaluation of the pilot study results (described in appendix A)

The sample of volunteers was recruited from subjects who were attending a normal examination at the ENT unit of the hospital and others were unpaid healthy volunteers (selected from the peer group of the researcher and from the staff working at the hospital). All the patients who came to this unit (except the healthy volunteers who came on purpose for the study) had had a previous ENT examination and they were referred in order to obtain a more precise diagnosis by using laryngoscopy and/or stroboscopy. From the total number of persons at the out-patient ENT unit the researcher randomly selected those who were sent with a voice complaint. Several difficulties arose in the subjects' recruitment either because they simply refused to participate, were illiterate, or needed glasses to read and did not have them at the time of assessment. Among potential healthy volunteers the major reason for refusal was the need to submit to a stroboscopy examination.

Prior to the assessment each subject had a verbal explanation about the purposes and the procedures of the experiment that was in progress. If he/she agreed to participate a consent form had to be signed (appendix B).

The assessment protocol included stroboscopic examination by two ENT surgeons (appendix C), individual history questionnaire (appendix D), acoustic and laryngographic recording and auditory voice assessment by the same speech and language therapist (the researcher), self-rating of psychosocial consequences of voice disorder (appendix E) and self-rating degree of psychological stress (appendix F) by each subject.

All subjects were assessed, only once, with all the sources above explained, always during the morning (from 9 to one) to avoid, as much as possible, within-day variations in F0 (Brown, Murry & Hughes, 1976; Garrett & Healey, 1987; Higgins, Netsell & Schulte, 1994).

5.1 - Subjects

One hundred and thirteen subjects participated in this study but the result was a set of 109 subjects (or 96.5% of all subjects recruited) whose data could be used. Only four subjects were eliminated based on the lack of or unreliable quality of the Lx signal (two were 61 year old men with leukoplakia, a 19 year old woman with psychogenic aphonia and a 43 year old woman without voice and laryngeal problems).

Subjects had to meet the following criteria for inclusion in the study: (1) age over 18 years old and under 65. This lower limit of 18 years was set based on the criteria that subjects had completed voice change and the upper limit of 65 years was selected to avoid the effects associated with the ageing process. That is, the same criteria used by several other researchers (Fitch, Holbrook & Tallahassee, 1970). However, a 67 year old female, subject 84, without voice problems was accepted; (2) Portuguese speakers with fluent and intelligible speech, as rated by the researcher, (3) no history of an upper respiratory tract infection (URTI) on the day of the evaluation and no auditory handicap and, (4) reasonable reading skills. Although two subjects with literacy difficulties were accepted (subjects 86 and 100) all their data from oral reading and subjects' self-rating Voice Handicap Index (VHI) and Social Readjustment Rating Scale (SRRS) were not accepted for analyses (as will be mentioned in the data analyses section).

From the total 109 subjects, 82 were female subjects with a mean age of 42 years old (ranging from 19 to 67 years old) and 27 were male subjects with a mean age of 39 years old (ranging from 20 to 63 years old).

The subjects were classified into two groups on the basis of their voice quality. For the purposes of this study, a subject was included in the 'dysphonic' group when he/she presented with a voice problem on the day of the assessment or history of voice complaints and for which the SLT and the ENT surgeon found corroborative evidence. As history of voice complaints meant only permanent or frequent episodic voice problems not related to upper respiratory tract infection or allergic situations. Subjects were included in the 'control' group when they did not report a voice problem on the day of the assessment or history of frequent voice problems not related to upper

respiratory tract infection or allergic situations and for which the SLT and the ENT surgeon did not find corroborative evidence.

The demographic information about the groups is presented in table V.1. The dysphonic group consisted of 52 subjects (39 females and 13 males). The mean age was 44 years and the mean level of education was eleven years. The control group consisted of 57 subjects (43 females and 14 males). The mean age was 39 years and the mean level of education was thirteen years. In both groups the majority of subjects were caucasian.

Table V.1 - Subjects' demographic information

	Dysphonics (n=52)	Controls (n=57)
Age	44.06 ± 12.31*	38.53 ± 14.94*
Age range	19-64 years	20-67 years
Females' age	43.62 ± 12.99*	40.12 ± 15.13*
Males' age	45.38 ± 10.38*	33.64 ± 13.69*
% Females	75%	75.4%
% Males	25%	24.6%
Years of education	10.69 ± 5.43*	12.5 ± 4.57*
% Caucasian	96.2%	94.7%

* M ± SD (Mean ± Standard deviation)

Independent T-tests were performed to compare the mean age according to dysphonic and control groups and the results showed that there was a significant difference (t=2.116, df=105, p=0.037). Nevertheless, when the mean age was compared according to gender between the groups the results show that there was no significant difference in age between the females' groups (t=-1.118, df=80, p=0.267) but there is a significant difference in age between the males' groups (t=-2.495, df=25, p=0.020) with the dysphonic males being older than control males (as might be seen in figure V.1).

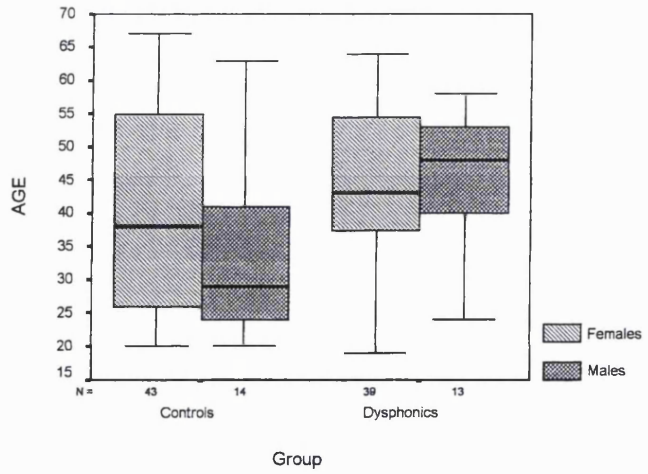


Figure V.1 – Age distribution according to gender and groups.

Also, no significant differences were found between the groups in the mean years of education (t=1.878, df=100, p=0.063), in gender frequencies ($\chi 2$=0.003, df=1, p=0.958) and in the frequencies of caucasians ($\chi 2$=0.125, df=1, p=0.724).

The social and voice habits history data (frequencies, means and standard deviation values) gathered during the interview is shown in the table below (table V.2) for both groups.

Table V.2 - Dysphonics and controls' social and voice habits history

	Dysphonics (n=52) M±SD*	Controls (n=57) M±SD*
% Lisbon natives	55.8%	61.4%
Number of years living in Lisbon	29.8 ± 15.3	26.4 ± 14.7
% In employment	61.4%	75%
% Living alone	15.8%	15.4%
% Professional voice use	25%	22.8%
% Voice hobby	12.3%	7.7%

* M ± SD (Mean ± Standard deviation)

No significant differences were found in the percentage of Lisbon natives ($\chi 2$=0.768, df=1, p=0.381), number of years living in Lisbon (t=1.176, df=107, p=0.242), in employment ($\chi 2$=2.306, df=1, p=0.129), in living status ($\chi 2$=0.003, df=1,

p=0.954), in professional voice use ($\chi2$=0.72, df=1, p=0.788) and voice hobby ($\chi2$=0.631, df=1, p=0.427) between the dysphonics and controls.

Table V.3 (below) shows the health history data gathered during the interview (the means and standard deviation values and frequencies) and from the ENT's examination (nose, hypopharynx and rhinopharynx) and from the self-rated scale of stress (SRRS).

Table V.3 - Dysphonic and control groups' health history

	Dysphonics (n=52) M ± SD*	Controls (n=57) M ± SD*
Height (cm)	1.62 ± .093	1.63 ± .087
Weight (Kg)	65.71 ± 13.38	63.23 ± 11.40
Medication consumption	0.79 ± 1.37	1.16 ± 2.52
Daily coffee intake	1.92 ± 1.64	1.35 ± 1.28
Daily water intake	1.08 ± 0.66	0.84 ± 0.51
Cigarettes a day	6.5 ± 11.64	1.64 ± 4.05
% smokers	28.8%	22.8%
Surgery over last 3 years	11.5%	10.5%
Respiratory diseases	15.4%	8.8%
Gastroesophageal disease	26.9%	12.3%
Minor nose abnormalities	57.7%	52.6%
Minor rinopharynx abnormalities	21.2%	15.8%
Minor hypopharynx abnormalities	50%	42.1%
Stress (SRRS)		
Total	117.8 ± 88.9	87.5 ± 63.9
Family	13.9 ± 25.6	16.16 ± 29.2
Personal	52.9 ± 56.6	30.11 ± 36.8
Work	14.4 ± 19.7	8.3 ± 14.4
Financial	36.7 ± 36.2	32.8 ± 36.7

*M ± SD (Mean ± Standard deviation)

No significant differences were found in the mean heights (t=0.511, df=105 p=0.611), mean weights (t=-1.033, df=105, p=0.304), percentage of smokers ($\chi2$=0.247, df=1, p=0.619), mean consumption of prescribed medication per day (t=0.930, df=107, p=0.354), reported frequency of use of birth control pills ($\chi2$=2.022, df=1, p=0.155),

frequency of reported respiratory diseases ($\chi2$=1.132, df=1, p=0.287), frequency of surgery over last 3 years ($\chi2$=0.028, df=1, p=0.866), frequency of minor nose abnormalities ($\chi2$=2.819, df=2, p=0.244), frequency of minor rhinopharynx abnormalities ($\chi2$=2.908, df=2, p=0.234), and in the frequency of minor hypopharynx abnormalities ($\chi2$=3.279, df=2, p=0.194) between the dysphonics and controls.

Significant differences were found in the mean number of cigarettes smoked a day (t=-2.850, df=62 p=0.006), mean quantity of daily water intake (t=-2.143, df=107, p=0.034), mean number of coffees per day (t=-2.032, df=107, p=0.045), the frequency of reported gastroesophageal disease ($\chi2$=3.748, df=1, p=0.045) and the total level of stress (t=-2, df=90, p<.05) between the groups.

In summary, dysphonics are older than controls (just for the male subjects), present higher daily consumption habits (cigarettes, coffee and water) than controls and present more health related problems (gastroesophageal disease and stress) than controls.

In order to determine the effects of group, age and gender on the results of the SRRS scale a multiple repeated analysis was carried out. Results revealed that the main effect of group was statistically significant (F(1,103)=3.85, p<.05) and the within subjects' effect also showed a statistically significant difference between the sub-scales scores (F(2,103)=94.4, p<.0001). A two way significant interaction was also found for SRRS-group interaction (F(2,103)=3.16, p<.05).

Post hoc comparisons were done so that significant differences between the SRRS sub-scales and group combinations could be assessed. Results showed significant differences for personal stress (t=-2.4, df=84, p<.05) and work stress (t=-1.8, df=89, p<.05) between the groups. Results shown in table V.3 confirm that dysphonics show higher scores in the subscales of personal and work stress than those of controls.

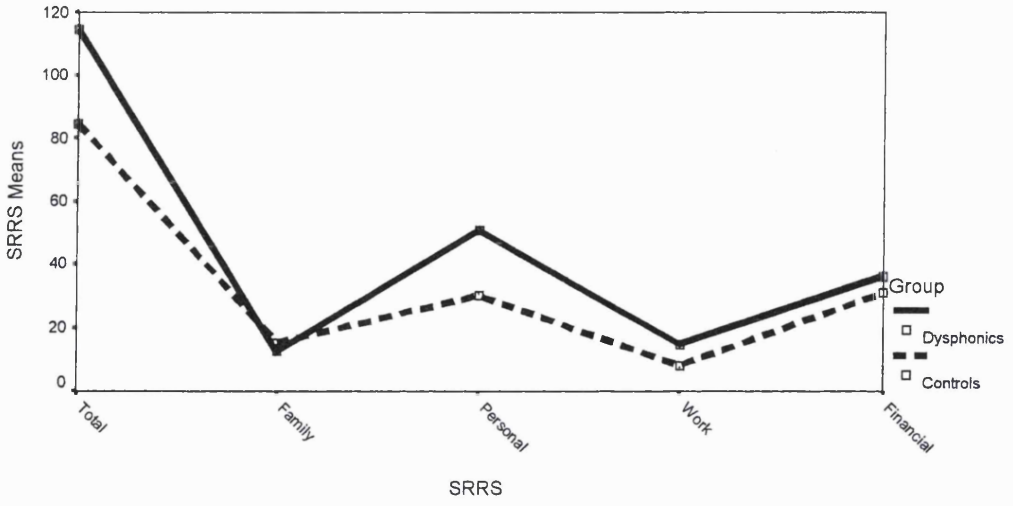


Figure V.2 – SRRS means according to groups

Figure V.2 (above) shows the significant main effect of group (by the gradient of the traces) and the significant interaction between SRRS subscale (family) and group (by the intersection of the traces).

According to the ENT results (figure V.3) in the dysphonics, 3 subjects (6%) present healthy larynxes and 1 subject (2%) shows minor larynx variation. The most frequent perturbation is the functional disorder (24 subjects) followed by the mass lesion disorder (14 subjects) and by the tissue change disorder (10 subjects). In the control group, 24 subjects (42%) presented minor larynx variations that do not interfere in the vocal fold vibration (e.g. 13 females presented slight posterior glottal chink).

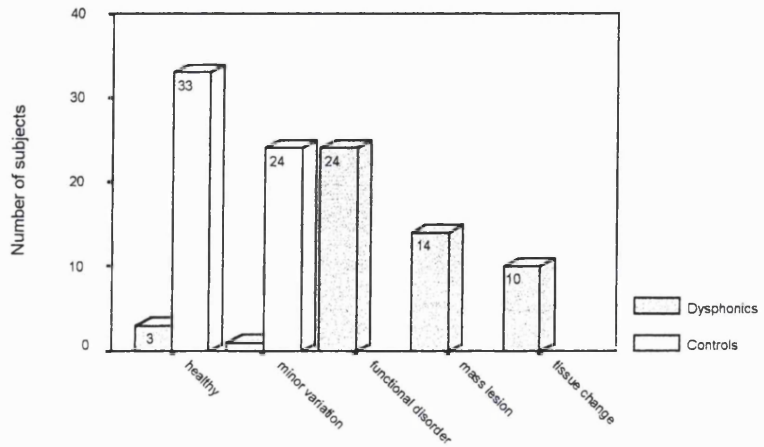


Figure V.3 - ENT larynx diagnosis

The SLT voice diagnosis (figure V.4) shows that in the dysphonics' group 4 subjects (7.6%) were diagnosed with 'normal' voice and the hoarse voice is the most frequent dysphonic pattern (19 subjects) followed by harsh voice (15 subjects) and breathy voice (14 subjects).

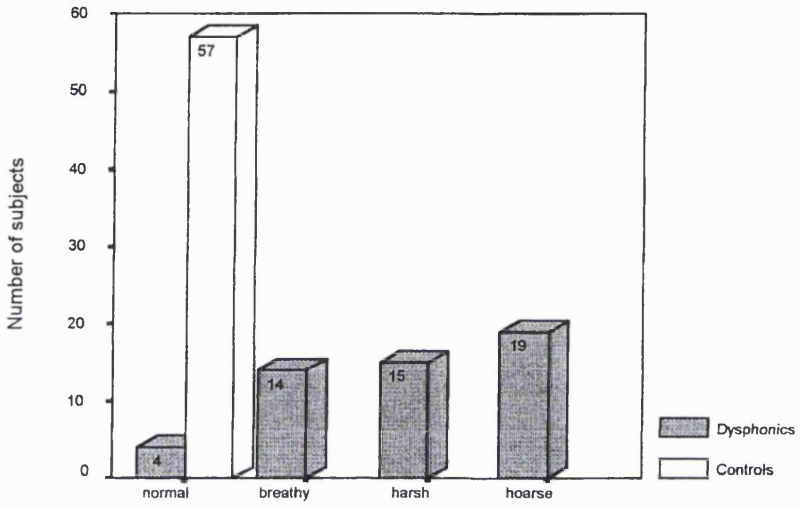


Figure V.4 - SLT voice diagnosis

According to the SLT dysphonia severity diagnosis among the 48 diagnosed dysphonic subjects, 35.4% show a severe degree of dysphonia, 35.4% a moderate degree of dysphonia whilst 29% show a slight degree of dysphonia.

Table V.4 – Dysphonia history

Dysphonics N= 52			
Cause	Onset	Symptoms	Duration
Work conditions – 25%	Gradual – 79%	Permanent – 73%	Less than one year – 21%
Disease – 21%	Sudden – 19%	Intermittent – 27%	One to five years – 44%
Psychological – 13.5%	Unknown – 2%		More than 10 years – 35%
Tobacco – 13.5%			
Several others – 12%			
Unknown – 12%			

The subjects' reported characteristics of dysphonia (table V.4) shows that the most frequent reported cause of dysphonia was work conditions followed by disease, psychological factors and tobacco, several causes, and unknown cause. The majority of subjects reported a gradual onset of the dysphonia and permanent symptoms of dysphonia. 44% of the subjects reported dysphonia from one year to five years, 35% had had dysphonia for more than 10 years and 21% for less than one year.

5.2 - Materials

The material used in this study was developed in or translated into Portuguese for the project and was based on materials used in previous studies investigating voice quality. The reliability or validity of the materials used was not formally tested. The translations into Portuguese were made by the researcher, back-translated by a Portuguese English teacher at secondary school in Portugal (who completed a Master's degree in the University of London in 1998/99) and revised by a bilingual (Portuguese/ English) speech and language therapist. Some of the investigator's relatives and friends read the translated materials in order to check for the quality of the translation. All the materials were previously tested in a pilot study (appendix A).

There were different types of materials: for screening (an ENT assessment form, a standard screening questionnaire, an auditory voice assessment and a self-rating scale of life events occurrence) and measurement (sustained vowels, reading, conversation and a self-rating questionnaire of the psychosocial consequences of voice). Screening materials were developed to gathered information and classify the subjects into the different groups (dysphonic and control). Although they constitute an important form of assessment for the purpose of the present study only part of the information was analysed.

The ENT assessment form is a modified version of the existent nasolaryngostroboscopy assessments (Hirano, 1981; Hirano and Bless, 1993), developed by the Portuguese ENT surgeons, to register their findings. It consists of personal details, clinical history, personal background, and family clinical history, laryngoscopic information about the nasal cavity, rhino- and hypopharynx and laryngeal anatomy. The items for the larynx observation are organised as: vibration (symmetrical or asymmetrical), glottal closure (complete, or incomplete), regularity of the vocal fold

edge (regular or irregular), the amplitude of vibration and mucosal wave ranging from normal to no visible movement, and non-vibrating portion ranging from none to total. The English and Portuguese version of the ENT assessment form is shown in appendix C.

A global auditory method of voice assessment in accordance with some of the existing perceptual methods and commonly used in perceptual voice studies (Hirano, 1981; Haji et al., 1986; Eskenazi, Childers & Hicks, 1990; Laver, 1991) was organised by the researcher. It consists of 4 overall voice qualities: normal, harsh, hoarse and breathy (described in detail in chapter III) and a four-point severity scale ranging from normal voice to severe dysphonia with the following criteria: (1) 'normal' voice - auditory impression of a voice characteristic of the speaker's age and gender and in accordance with the speaking situation. The auditory impression of non-dysphonic voice with peculiarities that are accepted socially and culturally was also classified as 'normal' voice because it is characteristic either of laryngeal physiology (e.g. slightly breathy voice associated with posterior glottal chink) or personal attitude (e.g. too low in loudness without giving the auditory impression of hypokinetic voicing), or the result of habits (e.g. habitual smoking is connected with too low pitch; and primary teachers' vocal load usually results in voices that are too loud); (2) slight dysphonia– auditory impression of very slight, intermittent voice abnormalities or isolated pitch or loudness perturbations without difficulties in intelligibility; (3) moderate dysphonia - auditory impression of frequent voice abnormalities of pitch or loudness with intermittent difficulties in intelligibility; (4) severe dysphonia - auditory impression of constant voice abnormalities of pitch or loudness with constant difficulties in intelligibility.

The screening questionnaire (appendix D) was developed according to the principal investigator's professional experience and with reference to some published 'history questionnaires', (Koschkee & Rammage, 1997; Dworkin & Meleca, 1997; Sapir, Mathers-Schmidt, Larson, 1996; Sapir, Keidar & Mathers-Schmidt, 1993; Sataloff et al., 1994a) and following similar procedures of some voice studies (Fitch, 1990). The questionnaire gathers information about a subject's personal details, social history, general health history, ENT history, voice history, smoking history, drinking history and food history. The purpose was to identify all the factors that may contribute individually, or in combination, to dysphonia. It was decided to use open-ended questions without pre-coded responses as subjects' answers may be complex,

unpredictable or too numerous to pre-code. Also, closed questions could carry the risk that subjects' answers may be forced into inappropriate categories.

The Social Readjustment Rating Scale (SRRS), Holmes & Rahe (1967), consists of a 43-item scale of major life events either indicative of the life style of the individual (personal) or of occurrences involving the individual (family, personal, work and financial). The emphasis is on major life events that could force the organism into illness or maladaptive behaviour, and not on psychological meaning, emotion or social desirability. The SRRS assigns points (through the use of the magnitude estimation technique) to stressful events on the basis of standards from the average individual. The scores on SRRS became well known as the 'resultant magnitude estimates', later termed life change units (LCU's) that reflect the amount of change in life that one will have to adjust to. Holmes & Rahe (1967) showed a high degree of consensus (correlation coefficients between 0.82 and 0.90) concerning the relative order and magnitude of the items in the scale for a sample of convenience composed of 394 American adult subjects, as did Ruch & Holmes (1971) in a replication study using an adolescent sample (correlation coefficient of 0.97).
Subjects are instructed to cross those items which they have experienced in a given period of time (usually 6 month, one year or the lifetime). A quantitative amount of change in ongoing life adjustment required for each subject in any given period of time was determined by Holmes & Rahe (Holmes & David, 1989) to define categories known as life crises with different magnitudes.
Since 1967, the original SRRS (used in the pilot study, appendix A) has undergone only two slight modifications. Firstly, dollar inflation makes items 20 and 37 inadequate ('mortgage over $10,000' and 'mortgage or loan less than $10,000' respectively) and they now read ('mortgage or loan for a major purchase' and 'mortgage or a loan for a lesser purchase' respectively), (used in the main study, appendix F). Secondly, the authors realised that the original SRRS with abbreviated wording for the life event items (decided by them to save space) was not as adequate as the complete wording used nowadays as a more accurate and helpful form (Holmes & David, 1989).

The SRRS form (appendix F) was used here because it is one of the most commonly used measures of general stress in the literature (Holmes & David, 1989; Miller, 1989; Goldman et al., 1996; Rosen & Sataloff, 1997). Moreover, it does not place high demands on the subjects' interpretation and time to complete it, and the

researcher needs no specific training. The stressors of the preceding six months were self-marked by each subject. Using six months as the time frame for subjects' ratings reduces the problems associated with subjects' recall of life events.

The speech materials used for the laryngographic and acoustic measurement were: (1) sustained vowels (the Portuguese vowels [a], [i] and [u]), (2) reading aloud- 'The Story of Arthur the Rat' developed by Abercrombie (1967) (appendix G) and, (3) conversation.

The rationale behind the choice of the sustained vowels was: (1) sustained vowels provide a relative stable condition of the phonatory system, therefore relevant to the evaluation of the larynx stability (Horii, 1979; Heiberger & Horii, 1982; Askenfelt & Hammarberg, 1986; Titze, Horii & Scherer, 1987; Fitch, 1990; Murry, Brown & Morris, 1995; Colton & Casper, 1996; Baken & Orlikoff, 2000); (2) these three vowels represent the range of physiological production (even in Portuguese as shown in table V.5) and include the two dimensions that have been mostly studied (vowel height and front/back) (Whalen & Levitt, 1995) (described in detail in chapter III) and (3) they have been used in other voice studies to which the data from this study can be compared (shown in chapter III).

Table V.5 – Portuguese vowels used in the study

Degree of mouth opening	Anterior	Central	Posterior	Tongue height
Relatively closed (lips spread)	[i]		[u]	High (lips rounded)
Open		[a]		Low

The rationale behind the choice of reading was: (1) it is a feasible procedure in order to obtain uniform samples (with the same linguistic content) of connected speech, (2) it provides repeatable and relatively consistent subject matter for cross-comparison (Ball, Faulkner & Fourcin, 1990; Orlikoff, 1991).
'The Story of Arthur the Rat' (Abercombie, 1967) was specifically chosen because: (1) its translation into Portuguese is interesting and easy to read (as mentioned by some Portuguese speakers not included in this study), (2) it has examples of direct speech as

well as narrative, which may help to gain a better idea of the speaker's fundamental frequency range. Direct speech also sounds more like natural conversation, (3) it gives a sample length appropriate to the analytical requirements of the measurement system (at least 2 minutes) (Horii, 1979; Ball, Faulkner & Fourcin, 1990; Howard, 1998). Although reading aloud is to a large extent governed by literacy skills and reading fluency, it may also be influenced by vision and by the difficulty of the passage read.
In the Portuguese version of 'The Story of Arthur the Rat' there are 26 sentences and 290 words. Sentence length ranges from 2 to 24 words with an average length of 12.42 words. The Portuguese version has the typical sentence structure for direct speech as used in other Portuguese traditional stories. The reading passage was presented, to each subject, in A4 format written in Arial font, size 14 (appendix G).

The rationale behind the choice of conversation is the fact that connected speech is considered to yield more representative results of habitual pitch than those obtained during sustained phonation (Britto & Doyle, 1990; Laver, Hiller & Beck, 1992; Yiu et al., 2000) and is more natural, in as opposed to more formal activity such as reading. Conversation was elicited by asking common questions that required extended answers ('What did you do yesterday?' and 'How did you come to the hospital?') as used in other studies (Hollien, Hollien & De Jong, 1997).

The Voice Handicap Index (VHI), Jacobson et al. (1997), is a questionnaire that measures the psychosocial handicapping effects of voice disorders as perceived by the subject. It consists of 30 self-report items where subjects rate their responses on a five point scale. The items are grouped in three domains (each domain has 10 items randomly ordered) representing functional (the effects of voice problem on daily activities), emotional (the subject's emotional response to the problem of voice), and physical effects (the subject's perception of his/her laryngeal discomfort and voice output characteristics) (appendix E).
The VHI provides a total score and three ratings for the subscales that will reflect the subject's current self-rating degree of her/his voice difficulties. No standardised score is calculated for the VHI; the higher scores indicate greater voice-related problems. The range of scores for each subscale is 0 to 40 and for the total score is 0 to 120. Although further studies are going on with the VHI to test its validity this instrument was chosen because: (1) at the data of the research design of the present study this was the only published tool which has been shown to be a reliable measure in the published data

(described in detail in chapter IV); (2) its construction is based on patients' statements and not on professionals assumptions which consequently might be a more sensitive outcome; (3) it allows the self-assessment of the general range of problems experienced by dysphonics rather than specific patterns of outcome by diagnostic categories; (4) it is short and easibly understandable to respond to and to score. Neverthless, published data on 'healthy' voices is still not available nor correlation with other forms of clinical evaluation.

5.3 - Equipment

The experiment used equipment widely used clinically and in voice research (as described in detail in chapter IV).

Direct observation of the vocal tract was achieved using a flexible nasoendoscope (Karl Storz endoscope 20211020) and a Bruel & Kjaer Rhino-larynx stroboscope light source, type 4914). The flexible endoscope has the advantage of permitting more natural voice production for the patient, and the examiner can observe the supraglottal region. However, the flexible endoscope is more invasive than the rigid endoscope; consequently it is a common procedure at the ENT unit in Lisbon to use an anaesthetic spray (Xilonibsa Spray from Inibsa Laboratories Lda). Peppard and Bless (1991 cited by Stemple, Glaze & Gerdeman, 1995) have demonstrated that the use of topical anesthesia does not alter the resulting visual image of the larynx. However, as Sataloff (1998:98) states, severe mucosal irritation, in a small number of persons, may result from the use of topical anaesthetic. Nasolaryngostrobosocopy gives complementary information to voice assessment because it allows a significant knowledge about the larynx structure and vibratory patterns in 'normal' and 'pathological' subjects.

During stroboscopy an electret condenser omnidirectional microphone, Pro. 2 EM-605, placed externally on the left thyroid ala, was used for simultaneous audio recording of the phonatory sample.

Simultaneous laryngographic and acoustic recordings were carried out using a portable Laryngograph Processor (Laryngograph Ltd), and a portable Digital Audio Tape (DAT) recorder (Sony TCD – D8) battery powered (see figure V.5). The

laryngograph processor is an electrolaryngographic (ELG) device (also called EGG outside UK), which allows non-invasive monitoring of vocal fold compliance during phonation (Fourcin, 2000).

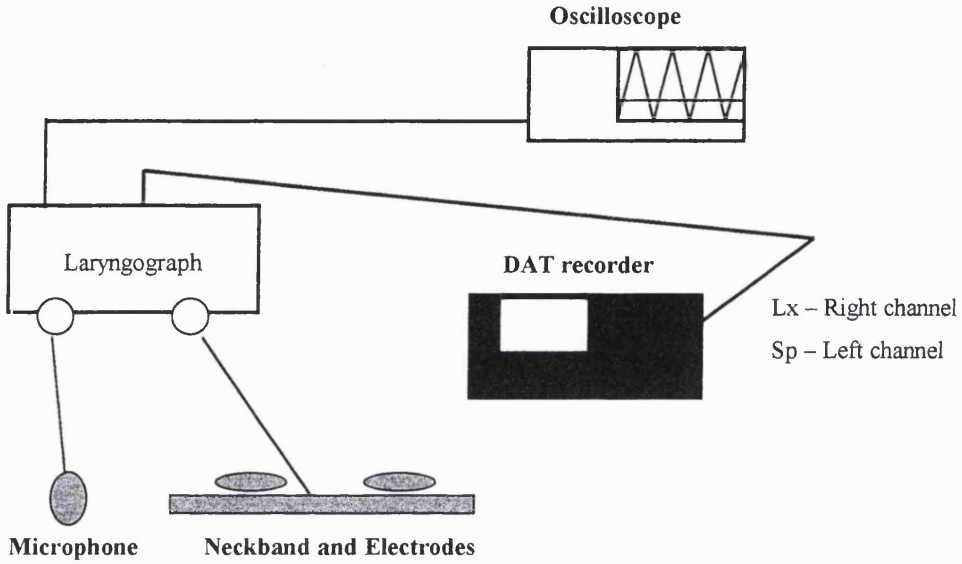


Figure V.5 - Diagram of equipment circuitry for data collection

The speech pressure waveform (Sp) was recorded on the left channel of the DAT recorder using a headmounted type electret condenser omnidirectional microphone (Knowles EK-3132), which allows for a greater standardisation of mouth to microphone distance (Titze & Winholtz, 1993; Gould & Korovin, 1994) and the upper frequencies will not be so reduced in amplitude with respect to the lower frequencies as in the microphone clipped to the clothing of the speaker (Howard, 1998). The larynx excitation waveform (Lx), recorded on the right channel, reflects the amount of translaryngeal current passage between bilateral surface electrodes (placed externally on either side of the thyroid cartilage at the level of the vocal folds), during phonatory tasks. The Laryngograph signal accuracy was monitored online using a Thandar, portable oscilloscope (type SC 110A). An external microphone was used to avoid picking up noise from the tape recorder and the digital tape recorder was used to avoid the phase disturbances often incurred using analogue recording and to obtain more accurate signal recordings (16-bit resolution at a sample rate of 48KHz) (Titze, Horii & Scherer, 1987; Doherty & Shipp, 1988; Laver, Hill & Beck, 1992; Gelfer & Fendel, 1995; Perry, Ingrisano & Blair, 1996).

Speech and Laryngograph analyses required a DAT recorder connected to a Laryngograph processor (the same equipment used to gather the data), which, in turn, was connected on to a PC computer with Lx Speech Studio software (Laryngograph Ltd.).

5.4 - Experimental environment

The laryngographic recordings were carried out in the same audiometric cubicle in the Audiology Department, used in the pilot study. Despite certain limitations (e.g. the limited dimensions of the cubicle) found during the pilot study (described in appendix A) it was considered the best possible place to do the recordings.

The DAT recorder, the laryngograph processor and the oscilloscope were set up on a table inside the audiometric cubicle (not connected to the main electricity supply) at the left side of the subject and the researcher was in front of the table.

5.5 - Experimental procedure

The data collection procedure involved two different phases, a pilot study (described in appendix A) and a main study. It used qualitative and quantitative assessment techniques that took on average more than one hour for each subject. The order for administering assessment measures is outlined below.

The primary assessment involved the use of stroboscopy. The examination protocol is briefly summarised as follows. Each subject was instructed to sit down next to the laryngoscopy equipment, as this is a normal procedure at the ENT unit. Firstly, the subject was anaesthetised via one nostril and the ENT surgeon collected basic information and filled in the rating form (appendix C). An ENT surgeon conducted this examination in conjunction with a nurse and the researcher. Simultaneously, the principal investigator (the SLT) took note of the subject's answers to avoid repeating questions or in order to pose other exploratory questions afterwards during the interview and did the auditory voice assessment. After approximately five minutes the subject was asked to sit up straight, and lean forward slightly whilst extending the head forward. A flexible scope was inserted into the anaesthetised nostril and down through the pharynx, and positioned so that a broadly symmetrical and complete image of the vocal folds and

gross laryngeal structures was obtained with continuous light during breathing and during phonation on [i]. The high tongue characteristics of this vowel causes the epiglottis to forward movement and consequently the lens-to-glottis distance is shortened and considered as the best anatomical setting for obtaining a high quality image by several researchers (Peppard, Bless & Milenkovic, 1988; Colton & Casper, 1996).

After this procedure the subject was asked to hold a contact microphone against the skin in the larynx region, providing input source for setting the flash rate of the stroboscope. The ENT surgeon changed to stroboscopy flashlight and asked the subject to phonate again on [i]. Trials were repeated until adequate samples were obtained, as judged by observation on the videomonitor. None of the subjects were excluded from the study because all successfully tolerated the examination. Each subject's examination was recorded on videotape and the ENT's observations were recorded on the appropriate rating form.

After the ENT examination, the assessment continued with the researcher engaging the subject in conversation eliciting demographic information, and identifying variables that could bring about changes in the voice. During the period of this research one or two interviews a day were carried out. All interviews were conducted in Portuguese and recorded on the English questionnaire form in Portuguese (appendix D). The interviews were conducted in the same office for each subject, but in different locations for the various subjects according to the available locations at any one time. Both researcher and subject were seated face-to-face and the interview was initiated with a brief explanation stressing confidentiality and an explanation of how the information would be used. Each interview took approximately 45 minutes for each subject. According to each subject's attitude different exploration strategies were used to create interactivity, to encourage response and clarify questions and answers.

At the end of the interview subjects were informed about the next location and about the following procedure. In the Audiology Department a self-rating sheet of the Portuguese version of the VHI was given to each subject (appendix E). Although this sheet contained written instructions, the principal investigator also gave the instructions orally. The subject was left alone, seated, while the researcher present in the room was setting up the recording equipment (as shown in figure V.5) in the audiometric cubicle.

At the end, each form was checked to determine whether it had been completed fully and the subject was asked to fill in any missing responses.

Following the self-rating of the VHI, tape recordings of each subject were made in a standardised manner in the audiometry cubicle. The subject was seated to minimise variability due to postural changes, and to create a more natural enviroment for reading and speaking. The equipment was on his/her left side and the SLT on his/her left front facing the equipment table. The electrodes (attached to an elastic neckband with 'velcro') were placed externally, one on either side of the thyroid notch, at the level of the vocal folds and secured at the back with 'velcro'. The larynx location was found using palpation. The head-worn type microphone was at a constant distance of a few centimetres from the right corner of the speaker's mouth. The researcher explained the importance of the subject not moving the head or the arms too much because of the electrodes placed on the neck. A short unrecorded conversation took place to allow the subject to become accustomed to wearing the laryngograph electrodes and to relax in the test situation. During this period, the subject was also instructed to produce a sustained vowel at a comfortable level. The levels on the DAT recorder were set to avoid very low recorded signals or distortion and clipping and the position of the electrodes was confirmed by the waveform on the oscilloscope screen. Once all these procedures had been carried out and a brief explanation as to the order the recordings would follow was given, the door was shut. The subject and the researcher stayed inside and recording began.

Simultaneous speech and Lx waveform recordings, for each subject, were made on a DAT recorder under the following conditions: First, each subject was asked to sustain the vowel [a] at a comfortable conversational pitch and loudness level. This was done by providing a previously demonstration on sustained [a]. The following vowels [i] and [u] used the same procedure and demonstration was provided if necessary. The subject was asked to repeat the productions if they were considered by the researcher to be too short in duration or too irregular. Second, each subject was required to read aloud (with comfortable pitch, loudness and speaking rate) an unrehearsed Portuguese version of 'The Story of Arthur the Rat' (appendix G). Finally, the subject was engaged in a a conversation (to elicit spontaneous speech) started by the researcher asking the question 'What did you do yesterday?'. If the subject was not talkative, a second question 'How

did you come to the hospital?' was used in order to obtain a minimum three minutes duration.

Subjects were instructed to maintain comfortable conversational pitch and loudness level because the researcher wanted to prevent any unnatural laryngeal variation that might affect subject's typical F0. This procedure is similar to investigations by Horii (1975); Koike, Takahashi & Calcaterra (1977); Linville (1988); Peppard, Bless & Milenkovic (1988); Orlikoff (1991); Kotby et al. (1993); Sussman & Sapienza (1994); Orlikoff et al. (1997); Chae et al. (2001).

After leaving the audiometry cubicle, the subject was presented with another self-rating scale (SRRS) shown in appendix F. A brief explanation was given and it was pointed out that he/she should mark the changes that may have happened in his/her life within the last six months. Because this direction might not have been immediately comprehensible the specific period was made clear. For example, subjects assessed on June should answer for the period between January and June 2000.

5.6 - Data analyses

5.6.1 - Screening data analyses

Data from the screening examinations (questionnaire, stroboscopy, auditory voice assessment and the self-rating scale of life events occurrence) were analysed subject by subject and edited in the Statistical Package for the Social Sciences (SPSS), version 10.

Data from the questionnaire was tabled according to demographic information (age, years of education, race and place of birth), social and voice habits history (economical activity, living status, professional voice use and voice hobby), health history (height, weight, body mass index, respiratory and gastroesophageal reported diseases, reported surgeries within 3 years, smoking habits, prescribed medication consumption, birth or hormonal compensation pill, alcohol consumption, type of drinks) and dysphonia history (subject's self-reported possible cause of the dysphonia, type of onset, type of symptoms and duration of the dysphonia). Some of the data from the questionnaire was interval (e.g. age; years of education; number of years of living in Lisbon; height, weigth, body mass index; number of pills consumed, number of coffees

per day, daily water intake, number of cigarettes a day; smoking habit duration; voice hobby duration). The criterion used for categorical data analyses is presented in the table V.6.

Table V.6 – Questionnaire categorical data analyses

Questionnaire data	Criterion used
Demographic information	
Gender	% of females and males;
Race	% of Caucasian and African;
Place of birth	Portugal – North (Minho; Douro litoral and Trás-os-Montes); Centre (Beira Litoral, Beira Alta e Beira Baixa); Lisbon and sourroundings (Estremadura and Ribatejo); South (Alentejo and Algarve); Islands (Madeira and Azores); Other Portuguese speaking countries (Angola, Mozambique, Cabo Verde).
Social and voice habits history	
Live alone	Yes or no;
In employment	No (e.g. housewife; unemployed, retired and students) or Yes.
Categories of occupation	1 - unemployed and retired; 2- students; 3- working in the health system; 4- working in the education system; 5- others;
Professional voice use	Koufmann & Isaacson's classification (described in chapter III).;
Voice hobby	Use of voice in singing or theatre.
Health history	
Smoking habits	Non smoker (never smoked or stopped smoking) or smoker;
Alcohol consumption	No (never or rarely), Yes (daily or frequent); Type (eg. beer or wine).
Surgery	Yes or no surgery in last 3 years;
Respiratory diseases	Asthma; bronchitis and emphysema;
Gastroesophageal disease	Yes or no.
Hormonal cycle	No (men); Menstrual cycle (premenstrual, menstrual, ovulation) menopause.
Hormonal pill	Yes or no.
Type of medication	The most prevalent (antihypertensives, antidepressants and hormonal compensation, others).
Dysphonia history	
Subject'self-reported:	
cause of dysphonia	1-Psycological factors; 2-Work; 3-Disease; 4-Tobacco; 5-Several; 6-Did not know.
type of onset	Gradual, sudden and unknown;
type of symptoms	Permanent or intermittent;
duration of the dysphonia	1-less than 1 year; 2 –one to 5 years; 3-more than 10 years.

Overall (within the limitations of the sample) place of birth was analysed according to the Portuguese phonetic variances (Cunha & Cintra, 1984) judged to be important for the voice study.

It is also clear from table V.6, that the dysphonia history (possible causes, onset of dysphonia, type of symptoms and duration) was also analysed. The dysphonic subjects' reported causes were analysed using the following criteria: (1) work, if dysphonia onset was related to vocal loading, (e.g. teaching) associated with noisy

ambience (e.g. machinery or people), chemical or dusty ambience and psychological stress (e.g. tiredness); (2) psychological, if the dysphonia onset was related to personality (e.g. anxiety) and/or traumatic period (e.g. death of a relative); (3) disease, if dysphonia onset was related to a upper respiratory infection (e.g. cold, allergies) and associated with non stopping vocal use; (4) tobacco, if dysphonia onset was related to smoking habits; (5) several causes, if dysphonia onset was related to iatrogenic causes (e.g. post-thyroidectomy or steroid injection) or trauma (e.g. accident) or sporadic voice abuse (e.g. singing) and (6) did not know the possible cause.

The larynx ENT diagnoses were organised in five different groups (according to the reported classification systems described in detail in chapter III): (1) healthy (if there was no structural or functional larynx variation or pathology); (2) minor variation (if there was a minor structural or functional larynx variation although not interfering with phonation e.g. posterior glottal chink, slight arytenoid asymmetry, slightly reduced anterior-posterior larynx dimension); (3) functional disorders (e.g. ventricular phonation; supraglottic tension and longitudinal gap); (4) mass lesions (e.g. nodules, polyps and cysts); (5) tissue changes (e.g. Reinke's oedema and laryngitis). All the larynx aetiology categories are described in detail in chapter II and III. Moreover, the ENT examination involved the diagnosis of minor nose abnormalities (e.g. deviated nasal septum, polyps, increased mucus production and congested mucosal membrane and paranasal sinuses hypertrophy), minor rhinopharynx abnormalities (e.g. cyst in the cavum, adenoidal hypertrophy, increased mucosal production) and minor hypopharynx abnormalities (lingual and palatine tonsils hypertrophy, tongue root hypertrophy, increased pharyngeal mucus production).

The SLT's auditory voice assessment results were organised according to type of voice and severity (as already described in the materials section). Data from the first perceptual evaluation (done at the hospital setting during the interview) was compared with a second perceptual analysis done 6 months after through listening to the taped conversation from eleven randomly chosen subjects. An agreement of approximately 82% between the first and second analysis was found for the voice type and severity respectively, which was considered a good perceptual evaluation.

The total data from the self-rating scale of life events occurrence (SRRS) was edited subject-by-subject as the sub-total data on the sub-scales (family, personal, work

and financial). Only data from two African subjects (a female and a male) was not considered for analysis because the contents had to be read by the researcher due to subjects' literacy problems. The tabled data according to the total and subtotal scores was 43 items versus 107 subjects.

5.6.2 – Laryngographic (Lx) data analyses

The two signals, speech (Sp) and laryngographic (Lx) were simultaneously displayed in real time on the computer monitor (Speech Studio program from Laryngograph Ltd.) subject-by-subject for each task (sustained vowels, reading and conversation). However, only Lx signals were used for analyses for the reasons expressed in the pilot study (described in detail in appendix A). The gain controls (in the Laryngograph Processor) had been adjusted, subjectively, subject-by-subject and task-by-task.

Before implementing automatic analyses, in order to limit the analyses error, several procedures were performed: (1) the Lx waveforms were bandpass filtered (Kemo type VBF8 filter) between 10 Hz and 5 kHz in order to remove low frequency variations (due to for example to larynx movement) as well as high frequency noise. This a common procedure used in similar studies (Higgins & Saxman, 1989a; Garner & Howard, 1999); (2) Lx waveforms with discontinuities were not accepted for data analysis. This judgment was based on visual inspection of the computer screen for continuity of the signal according to the published criteria (Ingrisano, Perry & Jepson, 1998); (3) oral reading samples which contained a substantial number of mispronunciations, hesitations, part-word segmentations and dysrhythmic flow of reading, were rejected for analyses (a similar procedure was used in McGlone & Hollien's, 1963 and Linke's, 1973 studies). Table V.7 presents the total speech tasks samples used for analyses according to groups.

Table V.7 – Speech samples used for the Lx analyses

	Dysphonics N=52	Controls N=57
Vowel [a]	40	54
Vowel [i]	46	57
Vowel [u]	46	56
Reading	49	53
Conversation	52	57

The voice samples (sustained vowels, reading and conversation) were digitised stored and analysed using the following criteria: (1) the 22050 Hz sampling rate was employed at it is critical for perturbation measures (described in detail in chapter III) and for practical reasons (a higher sampling rate would have result in a larger amount of data and consequently a higher cost of computer processing); (2) maximum possible periods of mid-segments of an extracted vowel portion of stable sound source to eliminate extra irregularity in onsets and terminations (Milenkovic, 1987; Horii, 1979; Kitajima, 1981). If no portion of relatively stable segments could be found no perturbation values were computed for that vowel production. (3) all the data from the selected oral readings; (4) three minutes of the selected spontaneous speech sample;

A total of 565 samples (5 speech tasks x 113 subjects) were analysed with the computer program.

Fundamental frequency and contact-based data (central tendency, range and regularity measures) was calculated using all values in the original voice sample (e.g. DFx1 and jitter first order analyses). These measures were selected because they reportedly have clinical utility (as described in chapter III) and are more efficient for discrimination of voice quality (Kingholtz & Martin, 1985).

The rationale behind the choice of DFx1 measures was because: (1) all the samples were previously filtered (as already mentioned) therefore extraneous noise was removed from the measurements; (2) DFx2 samples were fewer than half of the DFx1 samples both in 'controls' and dysphonics (e.g. DFx1 samples for the 'control' subject 78 were 5,239.00 whereas DFx2 samples were 1,711.00); (3) samples for the sustained vowels are gathered from all the samples (DFx1) so comparision between sustained phonation and connected speech would be more appropriate between similar sampling analyses.

The rationale behind the choice of the mean as a measure of central tendency (instead of the mode or median) was because: (1) it is a common measure for sustained vowels and connected speech; (2) the distribution of the values during connected speech approached 'normality' as shown in figure V.6 (Baken & Orlikoff, 2000); (3) the majority of studies report the mean (described in detail in chapter III). The 90% F0 range was chosen as it excludes the outlying 5% of the distribution at both ends of the frequency range and it is the closest approximation to the 'functional' range of the speaker (Ball, Faulkner & Fourcin, 1990).

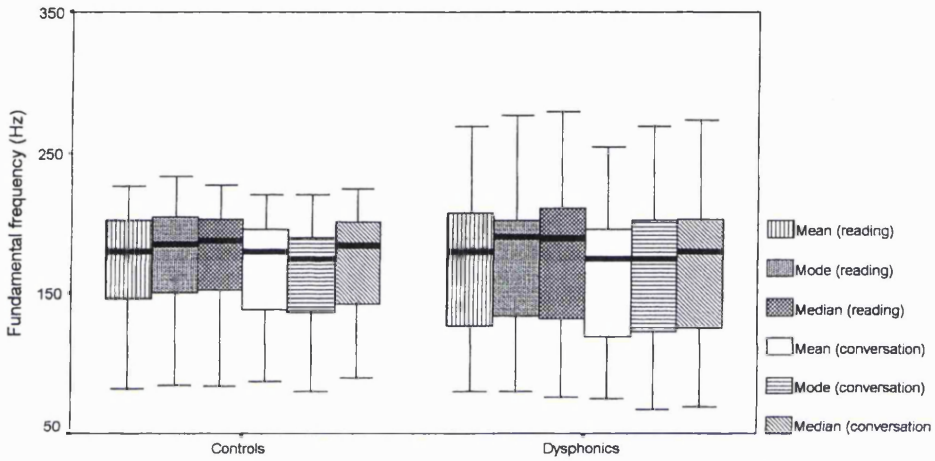


Figure V.6 – F0 central tendency measures

In order to ensure that Speech Studio is reliable in reproducing the acoustic measurements, voice samples from all the speech tasks from eleven randomly chosen subjects were reanalysed (two weeks after the first data analyses) by the same researcher. Mean absolute differences between the two sets of measurements were less than 1 Hz for F0, less than 1% for contact quotient, jitter and percentage of irregularity and less than .01 octaves for F0 range. These small differences were considered to represent a good measurement accuracy.

5.6.3 - Subjects' self-reported data analyses

The total data of psychosocial consequences of voice disorder (VHI) was edited subject-by-subject as the sub-total data on the sub-scales (emotional, functional and physical). Only two subjects (the same ones as in the SRRS scale) were excluded from data analysis because they were the only ones to whom the researcher read the contents

of the questionnaire due to their literacy problems. The tabled data according to the total and subtotal scores was 30 items versus 107 subjects.

5.7 - Statistical analyses

The design of this study included three independent or grouping variables: voice group (dysphonic and control), speech tasks (sustained vowels, reading and spontaneous speech), and subjects' self-rating (VHI).

The data were entered onto a personal computer and statistical analyses were performed using the SPSS (Statistical Package for Social Sciences), version 10.

Firstly, an exploratory analysis of all the edited screening data (questionnaire, stroboscopic examination, auditory voice assessment and the self-rating scale of life events occurrence) was performed to improve the internal validity of the study. Also an exploratory analysis of the speech samples duration, number of periods (from the sustained vowels) and number of samples (from reading and spontaneous speech), was done to find out if there were significant differences between the groups.

Table V.8 – Periods/samples used for the Lx analyses

	N=52	Dysphonics	N=57	Controls	Statistical test	
Vowel [a]	40	493±166.47	54	553.57±163.19	t=-1.764	**n.s.**
Vowel [i]	46	519.52±196.38	57	570.10±196.75	t=1.298	**n.s**
Vowel [u]	46	523.82±205.45	56	587.80±191.56	t=1.624	**n.s**
Reading	49	11139.87±4485.26	53	10079.01±3525.63	t=-1.333	**n.s**
Conversation	52	10683.38±3775.89	57	11632.018±3477.26	t=1.366	**n.s**

n.s. (p>.05)

From the data presented above (table V.8) it was concluded that there were no significant differences in the speech samples duration (used for analysis) between the groups. Moreover, the amount of Lx data included in the analyses (table V.8) was significantly high (between 86 to 100% of the total samples) compared to the data obtained in the pilot study and that reported on the published research (Colton & Conture, 1990; Winstanley & Wright, 1991; Verdonck-de Leeuw, 1998; Buekers, 1998a) (described in detail in chapter IV).

Non-parametric, Chi square, analysis was used to examine the difference between independent samples using categorical data. Parametric testing, T-test, was used whenever the data was ratio or interval and the assumptions of normal distribution (Shapiro's test) and homogeneity of variance (Levene's test) were met. The T-test was used to examine the differences between means of two independent groups (independent t-test) as well as in the same group in different conditions (related t-test).

A repeated measures analysis (ANOVA) was carried out with group (dysphonics and controls), age and gender as a between factor, and speech tasks (vowels, reading and conversation) as within subjects factors. For the VHI the between factors were also (group, gender and age) and the within-subjects factors were the total and sub-total scores (emotional, functional and physical).

The independent variable age in ANOVA analysis was dichotomised in subjects aged between 19 and 40 years old and subjects aged between 41 and 67 years old. This partition was carried out with Answer tree (chaid) from SPSS (1999), a statistical multi-way tree algorithm that allows examination of the data and quickly identifies critical segments and factors.

Whenever there was, for the within-subjects effects, heterogeneity of variance-covariance matrix known as sphericity (i.e. Mauchly Test is significant) the p-value for F in the Greenhouse-Geisser Test was assumed. Howell (1997) notes that the Greenhouse-Geisser Test is an adequate conservative test for correcting the problems with the sphericity assumption when testing for overall main effects or interactions.
Simple effect follow up t-tests were performed for the significant effects and interactions.

The probability level accepted as significant is $p<.05$ for all Chi-square, t-test and ANOVA.

Multiple linear regression was used to find which variable combinations are associated most with the key dependent measures: (a) F0 mean (b) F0 variability measures (standard deviation and 90% range); (c) F0 perturbation measures (jitter and percentage of irregularity); (d) contact quotient derived tendency measures (mean and

mode), (e) contact quotient derived variability measures (standard deviation) for all speech tasks and (f) VHI scores.

The predictor variables included in the model to ascertain the voice dependent measures were: case history data, ENT examination, SLT perceptual diagnosis, SRRS scores and VHI scores. Also, the same predictor variables (except VHI scores) were included in the model for the VHI dependent scores.

The selection of the predictor variables to be included in the analysis was; firstly, guided empirically where possible, or if not by what seemed to be appropriate. Secondly, a stepwise regression was run, whereby the included variables that may have contributed to the prediction of the dependent Lx and VHI measures were those that had a correlation probability level of significance of .05 or better with the dependent measure. This purging process was successful because each predictor variable contributes to the model independently of the other variables.

A final multiple linear regression equation for each speech task according to each speech task according to each voice parameter was obtained, $y=b0 + b1x1 + b2x2 +...bpxp$, where y is the dependent measure, the b0 is the regression constant and the parameters b1, b2 , ...bp are the partial regression coefficients.

In summary, this study comprised:

(1) a significant sampling of 'normal' and dysphonic speakers to allow the use of powerful statistical methods such as multiple regression analyses;
(2) acoustic and ELG recordings of vowels and connected speech;
(3) a complex evaluation that included, speaker's history case, ENT examination, SLT perceptual evaluation, speaker's self-rating of stress and speaker's self-rating of psychosocial voice impact;
(4) data analyses (using Speech Studio from Laryngograph Ltd.) derived from DFx1 measurements (for the reasons already mentioned in page 157).

Chapter VI – Results

6.1 – Introduction

The results will be presented according to the questions in study (1) what are the quantifiable voice characteristics of Portuguese speakers? (a) are there any statistically significant differences between two groups of different voice conditions on performance of different speech tasks? (b) which variables contribute most importantly to predicting variance in the quantifiable voice measures? (2) what is the dysphonic Portuguese speakers' self-rated voice psychosocial impact? (a) are there any statistically significant differences compared to a group of healthy Portuguese speakers? (b) which variables contribute most importantly to predicting variance in the quantifiable data?

The Lx data results will be presented according to the dependent variables: (a) mean fundamental frequency (F0); (b) mean F0 standard deviation; (c) 90% F0 range; (d) Jitter (mean); (e) percentage of irregularity (mean); (f) mean contact closed quotient (QFx); and (g) mean QFx standard deviation. In order to determine the effects of group, gender, age and speech on the above referred dependent variables repeated measures ANOVA was carried out.

6.2 – Voice characteristics

6.2.1 – Mean Fundamental frequency (F0)

Mean fundamental frequency values of all speech tasks according to the groups, and gender are displayed in table VI.1.

Table VI.1 – Means of F0 for all the speech tasks according to groups

		Vowel [a] Hz	Vowel [i] Hz	Vowel [u] Hz	Reading Hz	Conversation Hz
Dysphonics						
	Females	199.49	212.67	213.98	183.38	179.74
	Males	113.02	130.18	128.09	111.20	108.26
Controls						
	Females	210.51	220.96	225.06	190.33	186.61
	Males	118.35	127.58	122.99	109.59	109.21

It is clear from table VI.1, firstly, that females present higher means than male speakers. Secondly, the mean F0 values for the controls are systematically higher than the corresponding values for the dysphonics in the females and for the vowel [a] and conversation in the male speakers. However, in the male speakers controls present lower F0 means for the vowels [i],[u] and reading. Thirdly, within each group the mean F0 of the vowels is consistently higher than that of connected speech (reading and conversation). Fourthly, the mean F0 of reading is higher than that of conversation for all the speakers in the sample.

Intersubject variability within each group (controls and dysphonics) according to gender is presented in figure VI.1 and VI.2.

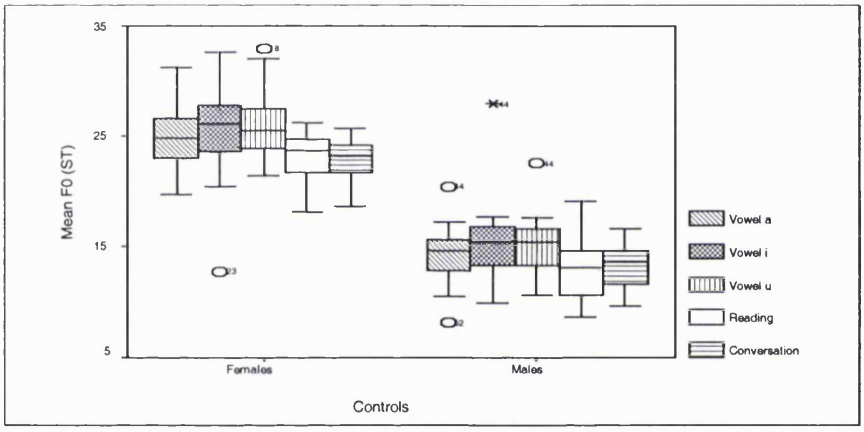


Figure VI.1 – Mean F0 (in semitones) for the control speakers according to gender

It is clear from figure VI.1 that in both groups the distribution of the mean F0 values approaches 'normality' (seen by the thick black bar across the interior of the box that represents the median). The vertical lines outside the box ('whiskers') show that the largest and smallest F0 values (that are not considered an outlier or an extreme score) are slightly wider for the females than for the males; however, there is an overlap between the lowest values during reading and conversation in the females with the highest values from reading in the male speakers. Also, both gender groups show outlier values (values that are more than 1.5 box-lengths away from the box and are represented with the symbol 'o') only during the sustained vowels. The male speakers

present an extreme value (more than 3 box-lenghts away from the box, represented by the symbol '*') in the sustained vowel [i] (subject 44).

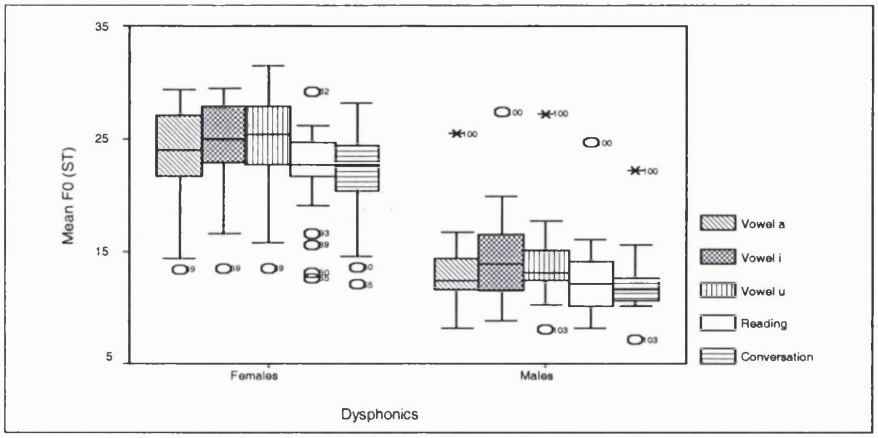


Figure VI.2 – Mean F0 (in semitones) for the dysphonics according to gender

It is clear from figure VI.2 that in the dysphonics: (1) females show a wider range of F0 values than the males; (2) there is an overlap of the F0 values between the female and male speakers; (3) there are outlier values for all speech tasks in the females and males; (4) there are extreme values for all speech tasks in the males (nevertheless, corresponding to the same speaker, number 100).

Comparison of the mean F0 values distribution, between controls and dysphonics, shows that: (1) the extension (smallest and largest) of the mean F0 values distribution (seen by the vertical lines outside the box) are similar between controls and dysphonics; (2) there is an overlap of all the mean F0 values between females and males (within the dysphonics) whereas in the controls is only between the reading of males and reading and conversation of females.

Statistical testing using multiple repeated analyses between the groups revealed that neither the main factor of group ($F(1,79)=.032$, $p>.05$) nor age ($F(1,79)=2.14$, $p>.05$) exerted a main significant statistical effect on the mean F0. However, gender has a highly significant statistical effect on the mean ($F(1,79)=123.29$, $p<.0001$). All the interactions between the factors were not statistically significant as might be seen in table VI.2.

Table VI.2 – Between subjects' effects for the mean F0

Source	df	Sum of Squares	Mean square	F-test	p (Sig. of F)
Between main effects					
Group	1	133.062	133.062	.032	.859
Gender	1	513262.294	513262.294	123.298	**.000**
Age	1	8890.547	8890.547	2.136	.148
Interactions					
Group* Gender	1	2069.464	2069.464	.497	.483
Group*Age	1	611.578	611.578	.147	.703
Gender* Age	1	13516.153	13516.153	3.247	.075
Group* Gender* Age	1	10502.503	10502.503	2.523	.116
Error	79	328858.431	4162.765		

Statistical testing using independent t-tests revealed that gender differences were highly significant for all speech tasks ($p<.001$) confirming that females present significantly higher F0 than the male subjects.

Multiple comparison analyses within the groups revealed that there was a highly significant speech effect ($F(2,179)=35.27$, $p<.001$) as was a significant speech by gender interaction ($F(2,179)=4.3$, $p<.05$). Data for within subjects' effects is presented in table VI.3.

Table VI.3 – Within subjects' effects for the mean F0

Source	df	Sum of Squares	Mean square	F-test	p (Sig. of F)
Within main effect					
Speech	2	35819.878	15791.416	35.268	**.000**
Between subjects interactions					
Speech* Group	2	480.828	211.976	.473	.648
Speech* Gender	2	4398.551	1939.129	4.331	**.011**
Speech* Age	2	1224.135	539.667	1.205	.305
Speech* Group * Gender	2	45.155	19.907	.044	.969
Speech* Group * Age	2	430.927	189.977	.424	.680
Speech* Gender * Age	2	201.328	88.757	.198	.846
Speech* Group * Gender* Age	2	64.978	28.646	.064	.954
Error	179	80236.233	447.755		

The significant interaction (speech-gender) resulting from this analysis was followed up by post hoc paired t-tests comparisons (shown in table VI.4) so that significant differences between various speech-gender combinations could be assessed.

Table VI.4 – Related t-tests for the mean F0 speech tasks pairs according to gender

	Females			Males		
Mean fundamental frequency	t	df	p	t	df	p
Vowel [a] versus vowel [i]	-4.21	68	**.000**	-2.82	23	**.010**
Vowel [a] versus vowel [u]	-6.32	67	**.000**	-3.56	23	**.002**
Vowel [I] versus vowel [u]	-1.35	74	.181	1.14	26	.265
Vowel [a] versus Reading	6.21	65	**.000**	3.19	22	**.004**
Vowel [a] versus Conversation	7.37	69	**.000**	4.14	23	**.000**
Vowel [i] versus Reading	9.29	71	**.000**	3.83	23	**.001**
Vowel [I] versus Conversation	10.28	75	**.000**	4.02	26	**.000**
Vowel [u] versus Reading	10.38	70	**.000**	4.68	23	**.000**
Vowel [u] versus Conversation	10.73	74	**.000**	4.72	26	**.000**
Reading versus Conversation	2.78	77	**.007**	1.13	23	.273

Results revealed that this combination produced significant differences between almost all the speech pairs. Firstly, there is a significant difference between all the vowels except for the pair vowel [i] versus vowel [u] for both genders. Secondly, it is also clear from table VI.4 that there are significant differences between all the vowels and connected speech (reading and conversation) for all subjects. Thirdly, differences between reading and conversation were only significant for females.

The results of the multiple regression analysis, indicating the contribution of variables, which accounted for the data variance in the fundamental frequency of all the speech tasks (vowels, reading and conversation) is presented in table VI.5.

Results shown in table VI.5 indicate that the overall model (combining gender, number of cigarettes smoked a day, sudden dysphonia onset and minor nose abnormalities) is a highly significant predictor ($p<.001$) model of the mean F0 for vowel [a], accounting for 70% of its variance (R squared=.699). Moreover, t and p-values within the model show that all variables are significant independent predictors. Beta weight coefficients show that the best predictor in the model is gender (-.692) followed

by the number of cigarettes smoked a day (-.265), sudden dysphonia onset (.206) and minor nose abnormalities (-.186).

Table VI.5 – Multiple regression analysis for variance in the mean F0

Vowel [a]				
ANOVA (regression) F(4,85)= 52.75, p=.000 Adjusted R Square = .699				
Predictor variables	Coefficient B	Beta Weight	t-value	p-value
(Constant)	218.52		44.92	.000
Gender (male)	-80.58	-.692	-11.01	.000
Number of cigarettes a day	-1.49	-.265	-4.25	.000
Dysphonia onset (sudden)	39.14	.206	3.52	.001
Minor nose abnormalities	-18.98	-.186	-3.00	.004
Vowel [i]				
ANOVA (regression) F(5,92)=31.27, p=.000 Adjusted R Square = .609				
Predictor variables	Coefficient B	Beta weight	t-value	p-value
(Constant)	214.87		38.92	.000
Gender (male)	-81.84	-.653	-9.39	.000
Number of cigarettes a day	-2.04	-.310	-4.33	.000
Number of coffees a day	5.94	.164	2.38	.020
Cause of dysphonia (disease)	-25.34	-.147	-2.23	.029
SLT diagnosis (hoarse)	22.95	.143	2.15	.034
Vowel [u]				
ANOVA (regression) F(5,91)=37.47, p=.000 Adjusted R Square = .655				
Predictor variables	Coefficient B	Beta Weight	t-value	p-value
(Constant)	223.403		35.01	.000
Gender (male)	-91.06	-.702	-10.79	.000
Cause of dysphonia (tobacco)	-50.06	-.212	-3.23	.002
Smoker	-21.99	-.172	-2.68	.009
SLT severity diagnosis (slight)	-35.68	-.207	-3.16	.002
Stress (SRRS total score)	.104	.142	2.15	.034
Reading				
ANOVA (regression) F(5,93)=54.94, p=.000 Adjusted R Square = .733				
Predictor variables	Coefficient B	Beta weight	t-value	p-value
(Constant)	217.22		30.30	.000
Gender (male)	-70.38	-.707	-12.72	.000
Number of cigarettes a day	-1.25	-.262	-4.78	.000
Age	-.759	-.249	-4.57	.000
Cause of dysphonia (psychological)	27.65	.166	3.11	.002
Dysphonia onset (sudden)	21.41	.151	2.84	.006
Conversation				
ANOVA (regression) F(4,98)=62.72, p=.000 Adjusted R Square = .708				
Predictor variables	Coefficient B	Beta weight	t-value	p-value
(Constant)	185.55		67.50	.000
Gender (male)	-62.11	-.647	-11.42	.000
Number of cigarettes a day	-1.09	-.236	-3.785	.000
Cause of dysphonia (tobacco)	-26.16	-.160	-2.64	.010
Cause of dysphonia (psychological)	20.34	.124	2.29	.024

Table VI.5 also shows that the combining predictor variables (gender, number of cigarettes smoked a day, number of coffees a day, subjects' reported disease as a cause of dysphonia and hoarse voice) is a highly significant model ($p<0.001$) of the mean F0 for vowel [i], accounting for 61% of its variance (R square=.609). All the variables in the model are significant independent predictors (as might be seen in the t and p-values). The best predictor in the model is gender (beta weight=-.653) followed by the number of cigarettes smoked a day (-.310), the number of coffees a day (.164), disease as a cause of dysphonia (-.147) and hoarse voice (.143).

As can be seen in table VI.5, the predictor model for vowel [u] is also highly significant ($p<.001$) with five independent significant variables accounting for 66% of the [u] mean F0 variance (R square=.655). The model has four negative independent predictors and a positive independent predictor variable. The best predictor in the model is gender (beta weight=-.707) followed by tobacco as a cause of dysphonia (-.212), slight dysphonia severity (-.207), smoker (-.172), and total stress (.142).

As can be seen in table VI.5, 71% of the mean F0 conversation variance results from a 4 predictor highly significant model ($p<.001$). Within the model four predictor variables are significantly negatively independent variables and one is a significantly positive predictor variable. The best predictor variable is gender (-.647) followed by the number of smoked cigarettes a day (-.236), tobacco as a cause of dysphonia (-.160); and psychological factors as a cause of dysphonia (.124).

In summary, although there are differences in the mean F0 between the dysphonics and controls only gender ($F(1,79)=123.29$, $p<.001$) and speech tasks ($F(2,179)=35.27$, $p<.001$) were shown to exert a main effect on the mean F0. The interaction between the two factors (speech-gender) was also significant ($F(2,179)=4.33$, $p<.05$).

The comparative analyses show that the mean F0 is: (1) significantly higher for the females than for the male subjects; (2) significantly higher for all vowels than connected speech (reading and conversation); (3) significantly different for all the vowels except for the pair vowel [i] versus vowel [u]; (4) significantly higher during reading than conversation (only for the female subjects).

The multiple regression analyses showed that a high prediction (between 61 and 73 %) for the mean F0 variance is explained by a combination of four or five variables for all speech tasks. The variables associated with the mean F0 were demographic (gender and age), health (smoking and coffee habits), dysphonia history (aetiology), ENT diagnosis (e.g. minor nose abnormalities) and the SLT diagnosis (e.g. hoarse voice and dysphonia severity).

Gender (male) was the best predictor variable for all the speech tasks (as shown in table VI.5) and it is associated with a lower mean F0. Other predictors that are associated with a lower mean F0 are: age, number of cigarettes a day, smoker, disease and tobacco as a cause of dysphonia, minor nose abnormalities and slight dysphonia severity. The variables (number of coffee a day, sudden dysphonia onset, psychological factors as a cause of dysphonia, hoarse voice and stress) are associated with a higher mean F0. The resulting regression equations that predict the mean F0 variance for all speech tasks are respectively:

- Vowel [a] = 218.52 – 80.58 (male) – 1.49 (number of cigarettes smoked a day) + 39.14 (sudden dysphonia onset) – 18.98 (minor nose abnormalities);

- Vowel [i] = 214.87 – 81.84 (male) – 2.04 (number of cigarettes smoked a day) + 5.94 (number of coffees a day) – 25.34 (subject's reported disease as a cause of dysphonia) + 22.95 (hoarse voice);

- Vowel [u] = 223.40 – 91.06 (male) – 50.06 (tobacco as a cause of dysphonia) – 35.68 (slight dysphonia severity) - 21.99 (smoker) +.104 (total stress);

- Reading = 217.22 – 70.38 (male) - 1.25 (number of cigarettes smoked a day) - .759 (age) + 27.65 (psychological factors as a cause of dysphonia) + 21.41 (sudden dysphonia onset);

- Conversation = 185.55 – 62.11 (male) - 1.09 (number of smoked cigarettes a day) – 26.16 (tobacco as a cause of dysphonia) + 20.34 (psychological factors as a cause of dysphonia).

6.2.2 – F0 standard deviation

The mean F0 standard deviation values of all speech tasks for all the subjects according to groups, gender and age are shown in table VI.6.

Table VI.6 – Means of F0 standard deviation according to the groups and age

	Vowel [a] Hz	Vowel [i] Hz	Vowel [u] Hz	Reading Hz	Conversation Hz
Dysphonics					
Females	4.8	4.9	4.9	.09	.11
Males	7.2	6.8	5.2	.11	.14
Controls					
Females	5.6	3.6	3.6	.12	.12
Males	1.04	1.56	7.04	.09	.12
Age					
Young	2.61	2.84	3.77	.09	.11
Old	7.69	5.57	5.72	.11	.13

Several trends can be noted from table VI.6. Firstly, ther is no systematic difference between dysphonics and controls. The dysphonics present higher values than that of the controls for the vowels [i] and [u] for the females whereas the male speakers is for the vowel [i], reading and conversation. Secondly, older speakers present higher F0 standard deviations. Thirdly, the mean F0 standard deviation of all vowels is higher than that of connected speech (reading and conversation) for both groups. Fourthly, although the mean F0 standard deviations for all the vowels have different trends according to the groups for connected speech (reading and conversation) are similar between the groups.

Intersubject variability within each group (controls and dysphonics) according to gender is presented in figure VI.3 and VI.4.

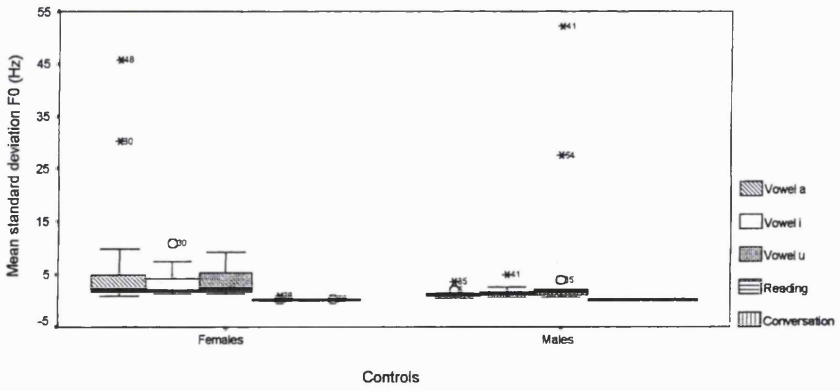


Figure VI.3 – Mean F0 standard deviation for the control speakers

It is clear from figure VI.3 that in both groups the distribution of the standard deviation F0 values is (1) more variable for the sustained vowels than for connected

speech; (2) slightly skewed (seen by the location of the thick black bar across the interior of the box that represents the median) but much more skewed for the female than the male speakers; (3) the vertical lines outside the box ('whiskers') show that the largest and smallest standard deviation F0 values (that are not considered an outlier or an extreme score) are wider for the females than for the males; (4) Also, both gender groups show outlier values (values that are more than 1.5 box-lengths away from the box and are represented with the symbol 'o') and extreme values (more than 3 box-lengths away from the box, represented by the symbol '*') during the sustained vowels and the connected speech (only for the females).

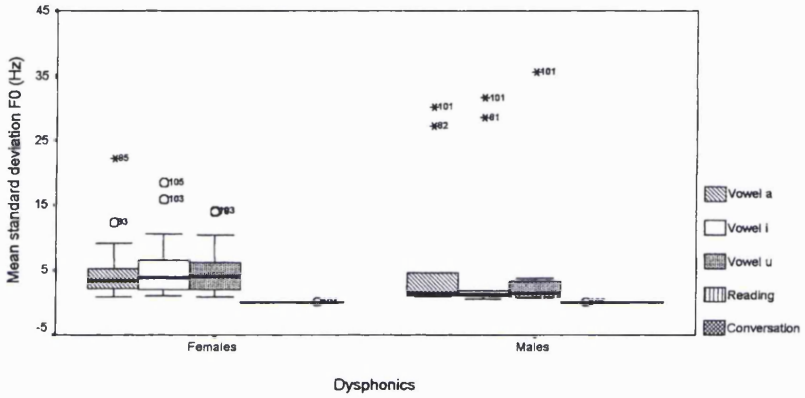


Figure VI.4 – Mean standard deviation F0 for the dysphonic speakers

It is clear from figure VI.4 that in the dysphonics there (1) is a more skewed distribution of the values in the males; (2) are outliers and extreme values for all speech tasks in the females (except reading); (3) are extreme values for all vowels and a outlier for reading in the males.

Overall, mean F0 standard deviation are wider (the smallest and largest values) for all vowels than for connected speech in both groups (controls and dysphonics), however, dysphonics show wider mean F0 standard deviation values for the vowels than controls.

Statistical testing using multiple repeated analyses (shown in table VI.7) revealed that neither group ($F(1,79)=1.36$, $p>.05$) nor gender ($F(1,79)=.18$, $p>.05$) had a significant statistical effect on the F0 standard deviation. However the main effect of

age is statistically significant ($F(1,79)=4.14$, $p<.05$). The interactions between all the factors were not statistically significant.

Table VI.7 – Between subjects' effects for the F0 standard deviation

Source	df	Sum of Squares	Mean square	F-test	P (Sig. of F)
Main effects					
Group	1	55.109	55.109	1.359	.247
Gender	1	7.416	7.416	.183	.670
Age	1	167.698	167.698	4.136	**.045**
Interactions					
Group* Gender	1	27.107	27.107	.669	.416
Group*Age	1	4.505	4.505	.111	.740
Gender* Age	1	.599	.599	.015	.904
Group* Gender* Age	1	30.899	30.899	.762	.385
Error	79	3203.071	40.545		

The significant age effect resulting from this analysis was followed up by post hoc paired t-tests comparisons (shown in table VI.8) so that significant differences between various speech-age combinations could be assessed. Results revealed that this combination produced significant differences between all vowels and connected speech for both age groups and between reading and conversation for the young speakers.

Table VI.8 – Related t-tests for the mean F0 speech tasks pairs according to age

	Young			Old		
Mean fundamental frequency	t	df	p	t	df	p
Vowel [a] versus vowel [i]	-.495	50	.623	1.32	41	.195
Vowel [a] versus vowel [u]	-1.06	50	.296	1.37	40	.177
Vowel [i] versus vowel [u]	-.895	52	.375	-.695	48	.490
Vowel [a] versus Reading	9.09	50	**.000**	4.27	37	**.000**
Vowel [a] versus Conversation	9.03	50	**.000**	4.63	42	**.000**
Vowel [i] versus Reading	5.02	52	**.000**	5.54	42	**.000**
Vowel [i] versus Conversation	4.99	52	**.000**	6.36	49	**.000**
Vowel [u] versus Reading	3.73	52	**.000**	5.48	41	**.000**
Vowel [u] versus Conversation	3.71	52	**.001**	6.22	48	**.000**
Reading versus Conversation	-4.85	52	**.000**	-.876	48	.385

The multiple comparison analyses (present in the table VI.9) revealed that within the groups there was a highly significant speech effect ($F(3,203)=16.75$, $p<.001$).

Table VI.9 – Within subjects' effects for the F0 standard deviation

Source	df	Sum of Squares	Mean square	F-test	p (Sig. of F)
Within main effect					
Speech	3	1403.490	546.584	16.745	**.000**
Between effects interactions					
Speech* Group	3	237.785	92.605	2.837	**.047**
Speech* Gender	3	124.661	48.549	1.487	.224
Speech* Age	3	204.443	79.620	2.439	.075
Speech* Group * Gender	3	232.992	90.738	2.780	.051
Speech* Group * Age	3	14.427	5.619	.172	.890
Speech* Gender * Age	3	74.815	29.136	.893	.433
Speech* Group * Gender* Age	3	136.040	52.980	1.623	.192
Error	203	6621.317	32.641		

Post hoc comparisons (shown in table VI.10) carried out to the results on the significant speech-group interaction confirmed that it is restricted to differences between the sustained vowels and connected speech for both groups and between reading and conversation for the dysphonic speakers.

Table VI.10 – Related t-tests for the F0 standard deviation speech tasks according to groups

	Dysphonics			Controls		
Mean fundamental frequency	t	df	p	t	df	p
Vowel [a] versus vowel [i]	.084	38	.933	1.25	53	.217
Vowel [a] versus vowel [u]	1.02	38	.314	.066	52	.948
Vowel [i] versus vowel [u]	.654	45	.517	-1.71	55	.093
Vowel [a] versus Reading	4.99	38	**.000**	3.55	49	**.001**
Vowel [a] versus Conversation	5.12	39	**.000**	3.78	53	**.000**
Vowel [i] versus Reading	5.17	42	**.000**	5.82	52	**.000**
Vowel [i] versus Conversation	5.49	45	**.000**	6.36	56	**.000**
Vowel [u] versus Reading	5.41	42	**.000**	3.96	51	**.000**
Vowel [u] versus Conversation	5.79	45	**.000**	4.27	55	**.000**
Reading versus Conversation	-5.81	48	**.000**	-.669	52	.507

The overall significance ANOVA regression on the mean F0 standard deviation for all the speech tasks is presented in table VI.11.

Table VI.11 - Multiple regression analysis in the F0 standard deviation

Vowel [a]				
ANOVA (regression) F(5.84)=7.87, p=.000 Adjusted R Square = .279				
Predictor variables	Coefficient B	Beta weight	t-value	p-value
(Constant)	-1.57		-.82	.417
Age	.131	.260	2.86	.005
Cause of dysphonia (tobacco)	8.82	.312	3.02	.003
Being born in Alentejo or the Algarve	11.40	.334	3.58	.001
ENT diagnosis (tissue change)	-9.82	-.397	-3.32	.001
Number of cigarettes a day	.203	.260	2.13	.036
Vowel [i]				
ANOVA (regression) F(8,89)=7.52, p=.000 Adjusted R Square =.350				
Predictor variables	Coefficient B	Beta weight	t-value	p-value
(Constant)	5.81		4.27	.000
Dysphonia onset (sudden)	5.37	.278	2.99	.004
SRRS (financial stress)	.047	-.259	-3.03	.003
Number of years living in Lisbon	.084	.240	2.89	.005
Cause of dysphonia (tobacco)	6.93	.314	3.36	.001
Alcohol consumption	-3.69	-.202	-2.39	.019
SRRS (personal stress)	.028	-.260	-2.79	.007
ENT diagnosis (tissue change)	-5.07	-.263	-2.91	.005
Control (normal voice)	-2.22	-.209	-2.11	.037
Vowel [u]				
ANOVA (regression) F(2,94)=5.88, p=.000 Adjusted R Square = .092				
Predictor variables	Coefficient B	Beta weight	t-value	p-value
(Constant)	6.64		6.27	.000
SRRS (financial stress)	-.059	-.248	-2.55	.012
Born in the islands (Azores or Madeira)	10.71	.220	2.26	.026
Reading				
ANOVA (regression) F(1,97)=42.16, p=.000 Adjusted R Square = .296				
Predictor variables	Coefficient B	Beta weight	t-value	p-value
(Constant)	.087		12.49	.000
Consume of medications per day	.019	.550	6.49	.000
Conversation				
ANOVA (regression) F(3,99)=5.55, p=.000 Adjusted R Square = .118				
Predictor variables	Coefficient B	Beta weight	t-value	p-value
(Constant)	.096		9.99	.000
Age	.096	.265	2.82	.006
Gender (male)	.00059	.273	2.83	.006
Alcohol consumption	-.015	-.214	-2.24	.027

28% of variance in the mean F0 standard deviation of the vowel [a] is accounted for by the combined effects of age, subject's reported tobacco use as a cause of dysphonia, being born in Alentejo or the Algarve, larynx pathology (tissue change) and

number of cigarettes smoked a day (as shown in table VI.11). All predictor variables are significant independent predictors. The beta weight coefficients show that the best predictor is larynx pathology (tissue change) being a negative predictor (-.397) followed by being born in Alentejo or the Algarve (.334), tobacco as a cause of dysphonia (.312), age (.260) and number of cigarettes smoked a day (.260).

Results shown in table VI.11 indicate that 35% of the variation in the overall quality of the standard deviation F0 for vowel [i] can be predicted on the basis of five positive predictor variables and three negative predictor variables. The best predictor is tobacco use as a cause of dysphonia (.314) followed by sudden dysphonia onset (.278), larynx pathology (distributed tissue change) (-.263), SRRS (personal stress) (-.260), SRRS (financial stress) (-.259), number of years living in Lisbon (.240), control subjects (normal voice) (-.209) and alcohol consumption (-.202).

It is clear from table VI.11 that the overall model (combining SRRS-financial stress and born in the islands) is a highly significant predictor model ($p<.001$) of the mean F0 standard deviation for vowel [u], accounting for 9% of its variance (R squared=.092). The best predictor is SRRS (financial stress) a negative independent variable ($t=-2.55$, $p<.05$) followed by being born in the islands (Azores or Madeira) a significant positive independent variable ($t=2.26$, $p<.05$).

Number of pills consumed a day is the only predictor variable that accounts for 30% of the mean F0 standard deviation variance during reading as can be seen in table VI.11. The variable is a significant independent predictor variable ($t=6.49$, $p<.001$).

As can be seen in table VI.11 the combining significant independent predictor variables, age ($t=2.82$, $p<.01$), gender ($t=2.83$, $p<.01$) and drinking habit ($t=-2.24$, $p<.05$) is a highly significant model ($p<.001$) of the mean F0 standard deviation during conversation, accounting for 12% of its variance (R squared=.118). The best predictor is gender followed by age and alcohol consumption.

In summary, although there are differences in the mean F0 standard deviation between the dysphonics and controls only age ($F(1,79)=4.14$, $p<.05$) and speech tasks ($F(3,203)=16.745$, $p<.001$) were shown to exert a main effect on the mean F0 standard

deviation. The interaction between the two factors (speech-group) was also significant ($F(3,203)=2.84$, $p<.05$).

The comparative analyses showed that the mean F0 standard deviation is: (1) not significantly different between all vowels for all speakers; (2) significantly different between all vowels and connected speech for all speakers; (3) there is no significant F0 standard deviation difference between reading and conversation (except for the older control speakers).

The multiple regression analyses showed that between a low (9% and 12%) and moderate (28% and 35 %) prediction of the variance in the F0 standard deviation (table VI.11) is explained by a combination of one to eight variables produced highly significant models ($p<.001$). Some demographic variables (age, gender-male, birth place and number of years living in Lisbon), social habit (smoking habit) and health related variables (consumption of medication, cause of dysphonia, dysphonia onset and personal and financial stress) are associated with a higher mean F0 standard deviation whereas others like 'controls', 'alcohol consumption', ENT diagnosis (larynx tissue change) and 'financial stress' are associated with a lower mean F0 standard deviation. Nevertheless it should be mentioned that 'financial stress' may also contribute to a lower mean F0 standard deviation (depending on the combined model variables as shown in table VI.11). The resulting regression equations that predict the mean F0 standard deviation for all speech tasks are respectively:

-Vowel [a] = -1.57 – 9.82 (larynx pathology – distributed tissue change) + 11.40 (being born in Alentejo or the Algarve) + 8.82 (tobacco use as a cause of dysphonia) + .131 (age) + .203 (number of cigarettes smoked a day);

-Vowel [i] = 5.81 + 6.93 (tobacco use as a cause of dysphonia) + 5.37 (sudden dysphonia onset) – 5.07 (larynx pathology – distributed tissue change) + .028 (SRRS-personal stress) + .047 (SRRS-financial stress) + .084 (number of years living in Lisbon) - 2.22 (normal voice).- 3.69 (alcohol consumption);

-Vowel [u] = 6.64 -.059 (SRRS-financial stress) + 10.71 (being born in the islands);

- Reading = .087 + .0 19 (number of pills consumed a day);

- Conversation = .096 + .0006 (male) + .096 (age) - .015 (alcohol consumption).

6.2.3 - F0 range

The 90% mean F0 range values of connected speech (reading and conversation) according to the groups and age are shown in table VI.12.

Table VI.12 – 90% means of F0 range (in octaves) according to groups

	Reading	Conversation
Dysphonics		
Females	.85 ± .33	1.12 ± .41
Males	1.0 ± .47	1.39 ± .39
Young adults	.99 ± .42	1.11 ± .40
Older adults	.80 ± .31	1.23 ± .42
Controls		
Females	.86 ± .22	1.12 ± .48
Males	.75 ± .18	1.18 ± .49
Young adults	.79 ± .19	.95 ± .39
Older adults	.90 ± .24	1.41 ± .47

M (mean) ± SD (standard deviation)

Several trends can be noted from table VI.12. Firstly, the 90% mean F0 range of all the groups is wider during conversation than reading. Secondly, within the females, dysphonic and controls present similar ranges. Thirdly, overall, the dysphonics, males and older adults present slightly wider range values in both reading and conversation tasks than that of the controls, females and young adults (except for the young adults during reading).

Intersubject variability within each group (controls and dysphonics) according to gender is presented in figure VI.5 and VI.6.

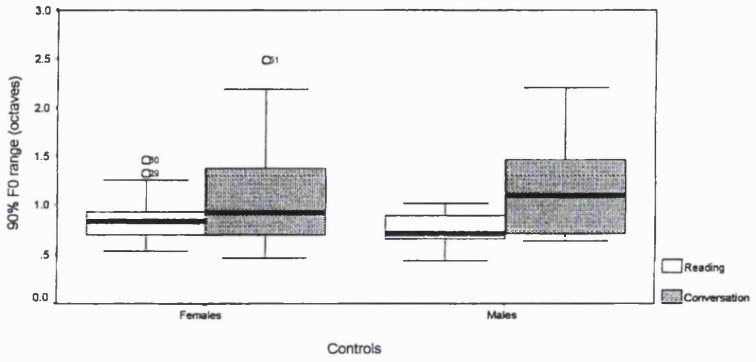


Figure VI.5 – 90% means of F0 range for the control speakers

It is clear from figure VI.5 that in both control groups the distribution of the 90% F0 range values is (1) wider (the vertical lines outside the box - 'whiskers' are longer) for the conversation than for the reading; (2) slightly more skewed during reading (seen by the location of the thick black bar across the interior of the box that represents the median) for both genders (3) only females show outlier values (values that are more than 1.5 box-lengths away from the box and are represented with the symbol 'o') during reading (subject 29 and 30) and conversation (subject 51).

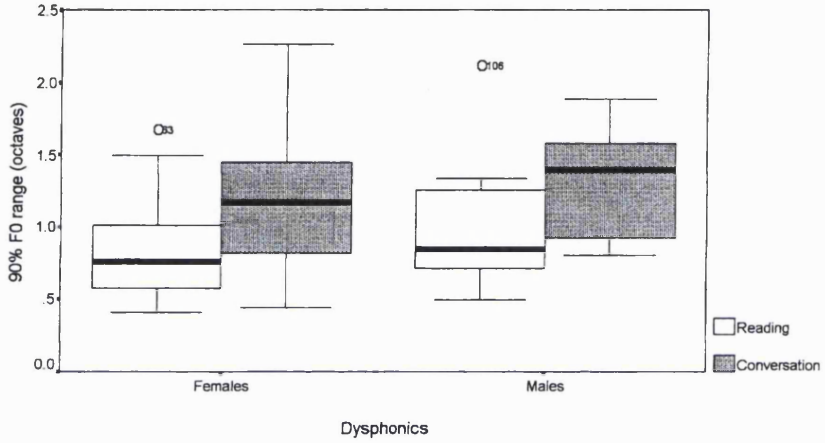


Figure VI.6 – 90% means of F0 range for the dysphonic speakers

It is clear from figure VI.6 that in the dysphonics the distribution of the 90% F0 range values is (1) wider (the vertical lines outside the box - 'whiskers' are longer) for the conversation than for the reading; (2) more skewed in the males; (3) only reading shows outlier values for both female (subject 63) and male (subject 106) speakers.

Statistical testing using multiple repeated measures analyses (shown in table VI.13) revealed that none of the factors (group, gender and age) had a significant main effect on F0 range between the groups. Nevertheless, a highly statistically significant interaction was found between the group and age ($F(1,94)=7.9$, $p<.01$).

Table VI.13 – Between subjects' effects for 90% F0 range

Source	df	Sum of Squares	Mean square	F-test	P (Sig. of F)
Main effects					
Group	1	.155	.155	.871	.353
Gender	1	.458	.458	2.568	.112
Age	1	.196	.196	1.096	.298
Interactions					
Group* Gender	1	.197	.197	1.102	.297
Group*Age	1	1.416	1.416	7.934	**.006**
Gender* Age	1	.0755	.0755	.423	.517
Group* Gender* Age	1	.197	.197	1.104	.296
Error	94	16.778	.178		

Inspection of means according to age and group presented in table VI.12 indicate that the young subjects present narrower ranges (during reading and conversation) as well as the young dysphonic speakers (during conversation) than that of the older speakers. Moreover, young dysphonic speakers show a wider range (during reading) than that of the older ones. Independent t-tests comparisons revealed that the effect of age is significant for the young controls during reading ($t=-2.251$, $df=33$, $p<.05$) and conversation ($t=-3.851$, $df=18$, $p<.05$) as well as for the dysphonic speakers during conversation ($t=-6.074$, $df=28$, $p<.001$) but not for the young dysphonic speakers during reading ($t=-1.742$, $df=18$, $p>.05$).

The multiple comparison within the groups revealed that speech has a highly statistically significant speech effect ($F(1,94)=46.1$, $p<.001$) and a significant interaction between the speech and age ($F(1,94)=13.4$, $p<.001$) as shown in table VI.14. Inspection of means shows that the F0 range for conversation is wider than that for reading.

Table VI.14 – Within subjects' effects for 90% mean F0 range

Source	df	Sum of Squares	Mean square	F-test	p (Sig. of F)
Within main effect					
Speech (90% range)	1	3.612		46.091	**.000**
Interactions					
Speech* Group	1	.170	.170	2.168	.144
Speech* Gender	1	.267	.267	3.401	.068
Speech* Age	1	1.048	1.048	13.366	**.000**
Speech* Group * Gender	1	.180	.180	2.294	.133
Speech* Group * Age	1	.065	.065	.838	.362
Speech* Gender * Age	1	.113	.113	1.445	.232
Speech* Group * Gender* Age	1	.094	.094	1.207	.275
Error	94	7.367	.078		

The results of the multiple regression analysis, indicating the contribution of variables, which accounted for the data variance in the fundamental frequency range of reading and conversation and also, t and p-values and the overall significance ANOVA regression are presented in table VI.15.

Table VI.15 - Multiple regression analysis in F0 range

90% range for reading				
ANOVA (regression) F(4,94)=7.25, p=.000 Adjusted R Square = .203				
Predictor variables	Coefficient B	Beta weight	t-value	p-value
(Constant)	.802		25.08	.000
ENT diagnosis (tissue change)	.254	.248	2.61	.010
SLT diagnosis (harsh)	.216	.248	2.61	.011
Menstruation	.212	.217	2.39	.019
Alcohol consumption	-.184	-.180	-1.99	.050
90% range for conversation				
ANOVA (regression) F(6,96)=7.97, p=.000 Adjusted R Square = .291				
Predictor variables	Coefficient B	Beta weight	t-value	p-value
(Constant)	.771		5.98	.000
Age	.0071	.226	2.64	.010
Minor nose abnormalities	.297	.333	3.82	.000
SLT diagnosis (breathy)	-.369	-.286	-3.133	.000
Number of cigarettes a day	.0133	.266	2.99	.004
Number of coffees a day	-.07	-.240	-2.77	.007
Cause of dysphonia (psychological)	.398	.226	2.45	.016

Results shown in table VI.15 indicate that 20% of the variance in the 90% F0 range during reading results from a 4 predictor highly significant model ($p<.001$). Within the model three predictor variables are significant independent positive variables whilst one is a significant independent negative variable. The best predictors are larynx pathology (distributed tissue change) (.248) and harsh voice (.248) followed by menstruation (.217) and alcohol consumption (-.180).

29% of variance in 90% F0 range during conversation is accounted for by age, nose alterations, breathy voice, number of cigarettes a day, number of coffees a day and psychological factors as a cause of dysphonia (table VI.15). All the variables are independent predictors with four positive and two negative. The best predictor is minor nose abnormalities (.333) followed by breathy voice (-.286), number of cigarettes a day (.266), number of coffees a day (-.240), age (.226) and psychological factors as a cause of dysphonia (.226)

In summary, only speech tasks ($F(1,94)=46.09$, $p<.001$) were shown to exert a main effect on the 90% mean F0 range and a interaction between the two factors (speech-age) was also significant ($F(1,94)=13.36$, $p<.001$).

The comparative analyses showed that the 90% mean F0 range is: (1) significantly wider during conversation than reading; (2) significantly narrower for the young control speakers than that for the older ones during reading and conversation; (3) significantly narrower for the young dysphonic speakers than that for the older ones during conversation.

The multiple regression analyses showed that a moderate prediction (between 20% and 29%) for the 90% mean F0 range variance is explained by a combination of four and six variables for reading and conversation respectively. The demographic variable 'age', health related (menstruation, psychological factors as a cause of dysphonia), social habit (number of cigarettes), ENT diagnosis (larynx tissue change and minor nose abnormalities), and STL diagnosis (harsh voice) are associated with a wider 90% F0 range whereas the social habit (alcohol consumption and coffee intake) and the STL diagnosis (breathy voice) are associated with more restricted 90% F0 range. The final multiple regression equation that predict the 90% mean F0 variance are:

- Reading = .802 + .254 (larynx pathology- distributed tissue change) +.216 (harsh voice) + .21 (menstruation) - .18 (alcohol consumption);

- Conversation = .77 + .297 (minor nose abnormalities) – .369 (breathy voice) + .013 (number of cigarettes per day) - .07 (number of coffees a day) + .007 (age) + .398 (psychological factors as a cause of dysphonia).

6.2.4 – Percentage of jitter

Table VI.16 (below) presents means and standard deviations of jitter for all the vowels according to group, gender and age.

Table VI.16 – Means of jitter percentage for all the sustained vowels

		Vowel [a]	Vowel [i]	Vowel [u]
Dysphonics				
	Females	1.48 ± 1.54	1.76 ± 2.48	1.49 ± 3.29
	Males	1.06 ± .83	2.00 ± 2.46	1.31 ± 1.52
Controls				
	Females	1.28 ± 2.62	.73 ± .88	.64 ± .57
	Males	.61 ± .36	.69 ± 1.03	.81 ± .93

M (mean) ± SD (standard deviation)

It is clear from table VI.16 that, firstly, the jitter percentage of the dysphonics is higher than that of the controls for all the vowels.

Intersubject variability within each group (controls and dysphonics) according to gender is presented in figure VI.7 and VI.8.

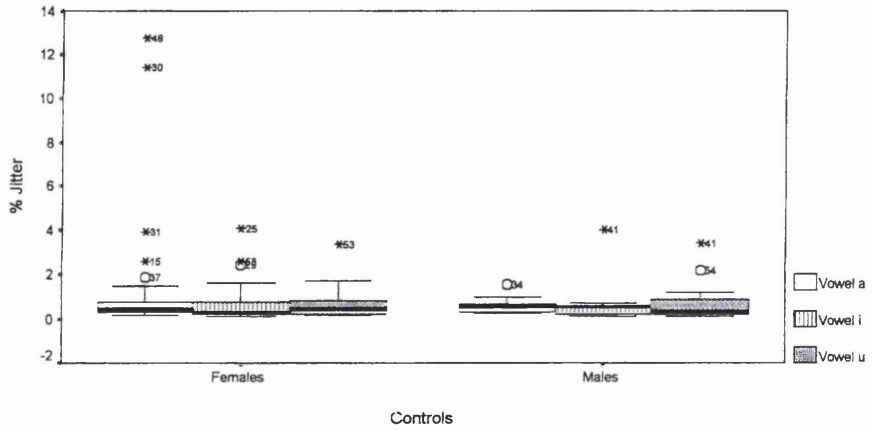


Figure VI.7 – % jitter for the control speakers

It is clear from figure VI.7 that in both control groups the distribution of the percentage jitter values is (1) skewed for all the vowels in both genders; (2) outlier values (values that are more than 1.5 box-lengths away from the box and are represented with the symbol 'o') and extreme values (more than 3 box-lengths away from the box, represented by the symbol *) are seen for all vowels in both genders (but more frequent in females).

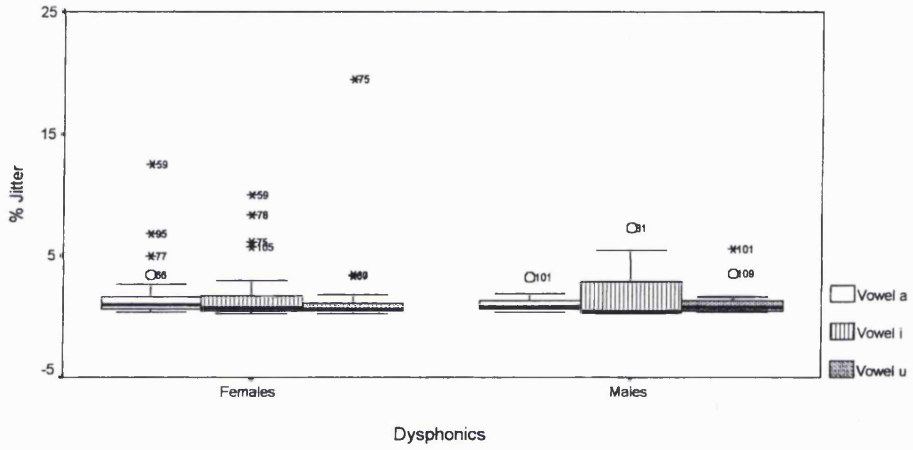


Figure VI.8 – % jitter for the dysphonic speakers

It is clear from figure VI.8 that in the dysphonics the distribution of the % jitter values (1) is skewed for all vowels; (2) is wider (the vertical lines outside the box - 'whiskers' are longer) for the vowel [i] in both genders; (3) shows outlier values (values that are more than 1.5 box-lenghts away from the box and are represented with the symbol 'o') and extreme values (more than 3 box-lenghts away from the box, represented by the symbol *) for all vowels in both genders.

Comparison between the two groups (controls and dysphonics) shows that the jitter values distribution is: (1) skewed in both groups; (2) wider in the dysphonics.

The multiple repeated analyses performed to examine whether statistical differences exist according to the main effects in study (group, gender and age) indicated no statistically significant differences between subjects and no significant interactions (table V.17).

Table VI.17 – Between subjects' effects for jitter

Source	df	Sum of squares	Mean square	F-test	p (Sig. of F)
Main effects					
Group	1	9.902	9.902	3.247	.075
Gender	1	.130	.130	.043	.837
Age	1	6.132	6.132	2.011	.160
Interactions					
Group* Gender	1	1.946	1.946	.638	.427
Group*Age	1	.276	.276	.090	.764
Gender* Age	1	.422	.422	.139	.711
Group* Gender* Age	1	8.465	8.465	2.776	.100
Error	82	250.050	3.049		

The multiple comparison analyses within the groups (table VI.18) did not reveal any statistically significant vowel effect nor any statistically significant interaction with the between subjects' factors.

Table VI.18 – Within subjects' effects for jitter

Source	df	Sum of squares	Mean square	F-test	p (Sig. of F)
Main effect					
Speech	1.5	1.805	1.162	.481	.572
Interactions					
Speech* Group	1.5	5.989	3.854	1.595	.211
Speech* Gender	1.5	9.131	5.877	2.433	.105
Speech* Age	1.5	3.630	2.336	.967	.364
Speech* Group * Gender	1.5	1.338	.861	.357	.647
Speech* Group * Age	1.5	1.205	.775	.321	.671
Speech* Gender * Age	1.5	2.941	1.893	.784	.430
Speech* Group * Gender* Age	1.5	1.096	.706	.292	.691
Error	127.4	307.819	2.416		

The results of the multiple regression analysis, indicating the contribution of variables, which accounted for the data variance in jitter and also, t and p-value and the overall significance ANOVA regression are shown in table VI.19.

Table VI.19 - Multiple regression analysis for jitter variance

Jitter percentage for vowel [a]				
ANOVA (regression) F(3.84)=8.9, p=.000 Adjusted R Square =.214				
Predictor variables	Coefficient B	Beta weight	t-value	p-value
(Constant)	2.79		5.09	.000
Born in Alentejo or the Algarve	3.52	.372	3.89	.000
Years of education	-.123	-.287	-3.01	.003
ENT diagnosis (minor variation)	-.93	-.208	-2.18	.032
Jitter percentage for vowel [i]				
ANOVA (regression) F(3,90)=10.14, p=.000 Adjusted R Square =.228				
Predictor variables	Coefficient B	Beta weight	t-value	p-value
(Constant)	1.44		3.14	.002
Group (control)	-1.59	-.416	-4.16	.000
SLT diagnosis severity (slight)	-1.95	-.329	-3.29	.001
Number of years living in Lisbon	.033	.257	2.81	.006
Jitter percentage for vowel [u]				
ANOVA (regression) F(4,91)= 8.09, p=.000 Adjusted R Squared = .230				
Predictor variables	Coefficient B	Beta weight	t-value	p-value
(Constant)	.855		2.83	.006
Being born in the North	1.92	.316	3.50	.001
Ovulation phase	1.71	.216	2.86	.005
ENT diagnosis (functional disorder)	1.31	.250	2.68	.009
SRRS (financial stress)	-.015	-.235	-2.55	.012

Results shown in table VI.19 indicate that 21% of the variance in jitter for vowel [a] is accounted for by the combined effects of being born in Alentejo or the Algarve years of education and ENT diagnosis (minor larynx variation). All the variables are independent predictors. The best predictor is being born in Alentejo or the Algarve (.372) followed by years of education (-.287) and minor larynx variation (-.208).

It is clear from table VI.19 that 23% of the variance in jitter for vowel [i] is accounted for by the combined effects of normal voice, slight dysphonia severity and number of years living in Lisbon. Being from the control group ($t=-4.16$, $p<.001$) and slight dysphonia severity ($t=-3.29$, $p<.001$) are negative predictor variables whereas number of years living in Lisbon ($t=2.8$, $p<.01$) is a positive predictor variable. The beta weight coefficients show that the best predictor in the model is 'normal' voice followed by slight dysphonia severity and number of years living in Lisbon.

The overall combining model (being born in the North, menstrual cycle, functional disorder and financial stress) is a significant predictor ($p<.01$) of jitter variance in vowel [u], accounting 23% of its variance (table VI.19). The variables,

being born in the North (t=3.5, p<.001), menstrual cycle (t=2.9, p<.01), and functional larynx disorder (t=2.7, p<.01) are significant positive independent variables whereas financial stress is a negative independent variable (t=-2.6, p<.05). Beta weight coefficients show that the best predictor is being born in the North followed by larynx functional disorder, financial stress and menstrual cycle.

In summary, although there are systematic jitter differences between the dysphonics and controls as well as between the females and males neither group, gender or age exerted a main effect on the jitter percentage.

The multiple regression analyses showed that a moderate prediction (between 21 and 23%) of the jitter variance is explained by a combination of three or four variables for the speech tasks (table VI.19). Demographic variables (number of years living in Lisbon, born in the Alentejo or in the Algarve and in the North), health (ovulation phase) and ENT diagnosis (functional disorders) are associated with higher jitter values whereas being from the 'control group', 'years of education', ENT diagnosis (minor larynx variation), SLT diagnosis (slight dysphonia severity) and 'financial stress' are associated with lower jitter values. The final multiple regression equations that predict F0 jitter variance are:

-Vowel [a] = 2.8 + 3.5 (being born in Alentejo or the Algarve) - .12 (years of education) – 93 (minor larynx variation);

-Vowel [i] = 1.44 - 1.59 (control group) - 1.95 (slight dysphonia severity) + .033 (number of years living in Lisbon);

-Vowel [u]) = .86 + 1.92 (being born in the North) + 1.31 (larynx functional disorder) - .015 (financial stress) + 1.71 (ovulation phase).

6.2.5 - Percentage of Irregularity

The values for means and standard deviations of percentage of irregularity for both speech tasks and according to group, gender and age are presented in table VI.20.

It is clear from table VI.20 that the mean percentage of irregularity of the dysphonics during reading and conversation is higher than that of the controls. Moreover, conversation presents higher values of irregularity than reading.

Table VI.20 – Means and standard deviations of percentage of irregularity

		Reading	Conversation
Dysphonics			
	Females	14.50 ± 10.6	17.3 ± 10.1
	Males	19.12 ± 12.6	28.4 ± 15.1
Controls			
	Females	7.6 ± 5.2	12.3 ± 7.4
	Males	12.5 ± 7.3	23.5 ± 15.6

M (mean) ± SD (standard deviation)

Intersubject variability within each group (controls and dysphonics) according to gender is presented in figure VI.9 and VI.10.

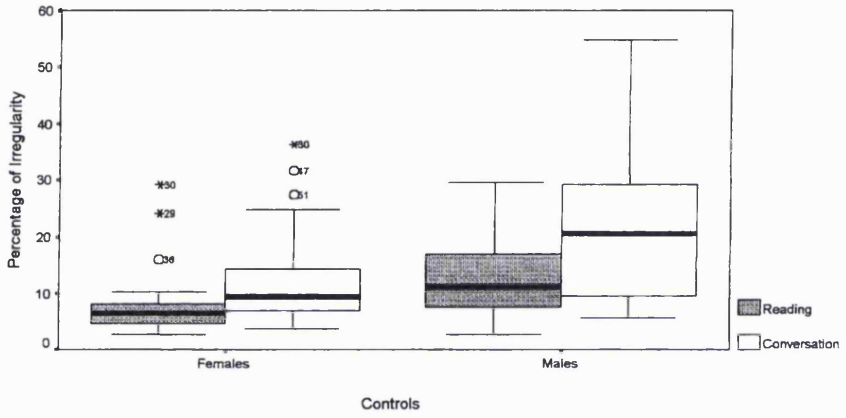


Figure VI.9 – % of irregularity for the control speakers

It is clear from figure VI.9 that in both genders the distribution of the percentage irregularity values is (1) more skewed and wider (the vertical lines outside the box - 'whiskers' are longer) for conversation in both genders; (2) outlier values (values that are more than 1.5 box-lengths away from the box and are represented with the symbol 'o') and extreme values (more than 3 box-lengths away from the box, represented by the symbol *) are seen just in the females.

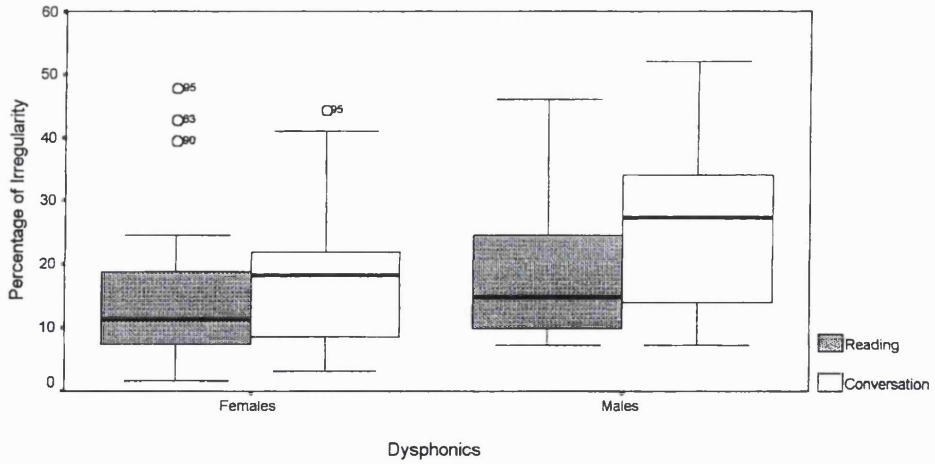


Figure VI.10 – % of irregularity for the dysphonic speakers

It is clear from figure VI.10 that in the dysphonics the distribution of the percentage of irregularity values is (1) more skewed and wider (the vertical lines outside the box - 'whiskers' are longer) for conversation in both genders; (2) outlier values (values that are more than 1.5 box-lengths away from the box and are represented with the symbol '○') are seen in the females only.

Statistical testing using repeated measures analyses (shown in table VI.21) confirmed that neither group ($F(1,94)=3.59$, $p>.05$) nor age ($F(1,94)=1.54$, $p>.05$) had a significant effect on the percentage of irregularity between the groups. Only gender did have a highly significant effect ($F(1,94)=13.71$, $p<.001$). Also, no statistically significant interactions were found.

Table VI.21 – Between subjects' effects for percentage of irregularity

Source	df	Sum of Squares	Mean square	F-test	p (Sig. of F)
Main effects					
Group	1	597.245	597.245	3.589	.61
Gender	1	2280.726	2280.726	13.707	**.000**
Age	1	256.756	256.756	1.543	.217
Interactions					
Group* Gender	1	51.601	51.601	.310	.579
Group*Age	1	296.308	296.308	1.781	.185
Gender* Age	1	4.690	4.690	.028	.867
Group* Gender* Age	1	145.942	145.942	.877	.351
Error	94	15640.569	166.389		

Pair-wise comparisons confirmed that the gender effect is restricted to conversation (t=-3.584, df=32, p=.001) with men presenting higher percentage of irregularity than women.

The multiple comparison analyses within the groups, presented in table VI.22, revealed a statistically significant speech effect (F(1,94)=71.65, p<.001) as was speech-group interaction (F(1,94)=5.16, p<.05), speech-gender interaction (F(1,94)=15.76, p<.001) and speech-gender-age interaction (F(1,94)=1.836, p<.05).

Table VI.22 – Within subjects' effects for irregularity

Source	df	Sum of squares	Mean square	F-test	p (Sig. of F)
Within main effect					
Speech	1	1501.893	1501.89	71.654	**.000**
Interactions					
Speech* Group	1	108.050	180.050	5.155	**.025**
Speech* Gender	1	330.225	330.225	15.755	**.000**
Speech* Age	1	39.810	39.810	1.899	.171
Speech* Group * Gender	1	49.920	49.920	2.382	.126
Speech* Group * Age	1	42.455	42.455	2.025	.158
Speech* Gender * Age	1	92.615	92.615	4.419	**.038**
Speech* Group * Gender* Age	1	38.488	38.488	1.836	.179
Error	94	1970.283	20.960		

Post hoc comparisons confirmed that the significant interaction effect of group is for both reading (t=3.609, df=74, p=.001) and conversation (t=2.252, df=107, p=.026). Observation of the means (table VI.20) confirms that the dysphonics present a higher percentage of irregularity for both speech tasks than that from the controls.

The significant interaction speech-gender-age was followed up by paired-wise comparisons so that significant differences could be assessed. Results confirm that the effect is restricted to females during reading (t=-2.23, df=76, p<.05).

Table VI. 23 presents means and standard deviations for the F0 irregularity during reading and conversation. Although older subjects present higher percentages of

irregularity than the younger one only reading for the females was statistically significant.

Table VI.23 – Percentage of irregularity according to gender and age

		Reading	Conversation
Females			
	Young	8.83 ± 6.4*	13.14 ± 7.9*
	Older	13.26 ± 10.6	16.12 ± 9.9
Males			
	Young	14.2 ± 10.9	21.74 ± 14.4
	Older	16.84 ± 9.4	30.25 ± 15.5

* M(mean) ± SD (standard deviation)

The results of the multiple regression analysis in F0 irregularity (during reading and conversation), t and p-values and also the overall significance ANOVA regression is shown in table VI.24.

Table VI.24 - Multiple regression analysis in F0 irregularity

Reading				
ANOVA (regression) F(4,94)=12.09, p=.000 Adjusted R Square = .312				
Predictor variables	Coefficient B	Beta weight	t-value	p-value
(Constant)	8.09		7.72	.000
Cause of dysphonia (tobacco)	12.74	.346	3.99	.000
SLT diagnosis (hoarse)	8.80	.351	4.15	.000
Being born in the North	5.39	.193	2.28	.025
Gender (male)	3.89	.176	2.04	.044
Conversation				
ANOVA (regression) F(6,96)=11.83, p=.000 Adjusted R Square = .389				
Predictor variables	Coefficient B	Beta weight	t-value	p-value
(Constant)	19.18		10.49	.000
Gender (male)	7.85	.295	3.53	.001
Cause of dysphonia (tobacco)	17.46	.385	4.57	.000
Being born in the islands	25.19	.305	3.79	.000
Number of coffees a day	-2.09	-.275	-3.42	.001
SLT diagnosis (hoarse)	7.17	.233	2.91	.005
SRRS (total stress)	-.034	-.231	-2.81	.006

31% of the variance in F0 irregularity during reading is accounted for by tobacco as a cause of dysphonia, hoarse voice, being born in the North and gender as shown in table VI.24. The independent significance of each predictor is: tobacco as a cause of

dysphonia (t=3.99, p<.001), hoarse voice (t=4.15, p<.001), being born in the North (t=2.28, p<.05), and gender (t=2.04, p<.05). The best predictor is hoarse voice followed by tobacco as a cause of dysphonia, being born in the North and gender.

As can be seen in table VI. 24 the overall combining model (gender, tobacco use as a cause of dysphonia, being born in the islands, number of coffees a day, hoarse voice and total stress) is a highly significant (p<.001) predictor model of F0 irregularity in conversation, accounting for 39% of its variance (R squared= .389). Each predictor in the model is an independent significance variable: gender (t=3.53, p<.001), tobacco as a cause of dysphonia (t=4.57, p<.001), being born in the islands (t=3.79, p<.001), number of coffees a day (t=-3.42, p<.001), hoarse voice (t=-2.91, p<.01) and total stress (t=-2,81, p<.01) The best predictor is tobacco use as a cause of dysphonia followed by being born in the islands, gender, number of coffees a day, hoarse voice and total stress.

In summary, gender (F(1,94)=13.71, p<.001) and speech (F(1,94)=71.65, p<.001) have a highly statistically significant effect in the percentage of irregularity. Moreover statistically significant interactions between speech-group (F(1,94)=5.2, p<.05), speech-gender (F(1,94)=15.75, p<.001) and speech-gender-age (F(1,94)=4.42, p<.05) were also found.

The comparative analyses show that the percentage of irregularity is: (1) significantly higher for the dysphonics than for the 'controls' in both speech tasks; (2) significantly higher for the males than for the females during conversation; (3) significantly higher for the older females than for the young ones during reading.

The multiple regression analyses showed that a moderate prediction (between 31 and 39%) of F0 irregularity variance during reading and conversation is explained by a combination of four and six variables respectively. Within the combining models the demographic variables (place born, gender-male), SLT diagnosis (hoarse voice) and health related variables (tobacco as a cause of dysphonia) are associated with a higher percentage of irregularity whilst 'coffee intake' and 'total stress' are associated with a lower percentage of irregularity. The final multiple regression equations that predict the F0 percentage of irregularity are:

-Reading = 8.09 + 8.8 (hoarse voice) + 12.74 (tobacco as a cause of dysphonia) + 5.39 (being born in the North) + 3.9 (male);

-Conversation = 19.18– 17.46 (tobacco use as a cause of dysphonia) + 25.19 (being born in the islands) + 7.85 (male) – 2.09 (number of coffees a day) + 7.17 (hoarse voice) – .034 (total stress).

6.2.6 – Mean contact closed quotient (QFx)

Means of contact closed quotient for all the speech tasks according to group, gender and age are presented in table VI.25.

Table VI.25 – Means of QFx (%)

	Vowel [a]	Vowel [i]	Vowel [u]	Reading	Conversation
Group					
Dysphonics	44.80	46.47	46.01	44.17	44.59
Controls	45.34	46.05	46.44	44.89	44.27
Gender					
Females	45.24	45.94	45.69	44.12	43.82
Males	44.74	47.09	47.79	45.91	46.27
Age					
Young adults	45.19	46.05	46.17	44.99	44.50
Older adults	45.01	46.45	46.33	44.07	44.36

It is clear from table VI.25 that the contact quotient means are very similar between groups and only male speakers present slightly higher values than females for all speech tasks except for the vowel [a].

Intersubject variability within each group (controls and dysphonics) according to gender is presented in figure VI.11 and VI.12.

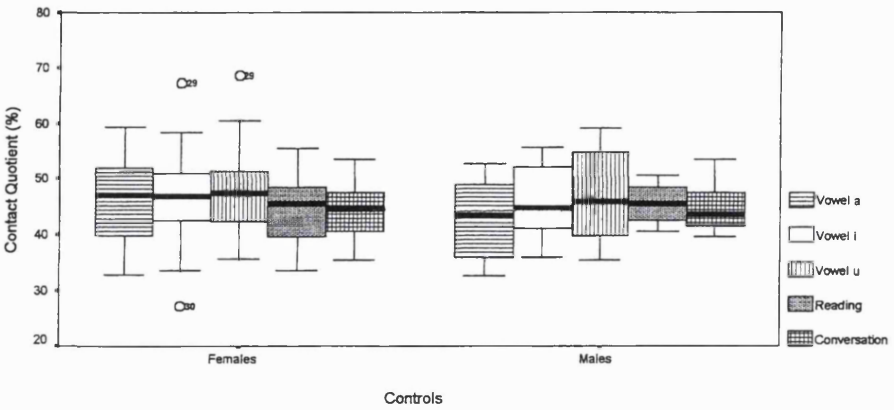


Figure VI.11 – % of contact quotient for the control speakers

It is clear from figure VI.11 that in both genders the distribution of the percentage of contact quotient values is (1) wider (the vertical lines outside the box - 'whiskers' are longer) for vowels than connected speech in both genders; (2) outlier values (values that are more than 1.5 box-lengths away from the box and are represented with the symbol 'o') are seen just in the females (for vowel [i] and [u]).

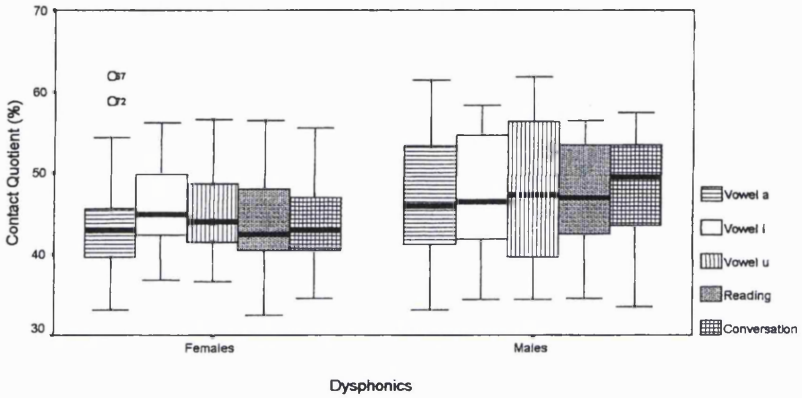


Figure VI.12 – % of contact quotient for the dysphonic speakers

It is clear from figure VI.12 that in the dysphonics the distribution of the percentage of contact quotient values is (1) wider (the vertical lines outside the box - 'whiskers' are longer) for vowels than connected speech in the males; (2) outlier values (values that are more than 1.5 box-lengths away from the box and are represented with the symbol 'o') are seen in the females only for vowel [a].

A multiple repeated analysis was performed to examine whether statistical differences between the subjects exist according to the main effects in the study (group, gender and age) and interactions. Results presented in the table VI.26 indicated no statistical significant differences or interactions.

Table VI.26 – Between subjects' effects for QFx

Source	df	Sum of Squares	Mean square	F-test	p (Sig. of F)
Main effects					
Group	1	29.384	29.384	.174	.677
Gender	1	148.961	148.961	.884	.350
Age	1	38.917	38.917	.231	.632
Interactions					
Group* Gender	1	209.411	209.411	1.243	.268
Group*Age	1	80.110	80.110	.475	.493
Gender* Age	1	15.301	15.301	.091	.764
Group* Gender* Age	1	12.082	12.082	.072	.790
Error	79	13312.410	168.512		

The multiple comparison analyses revealed that within the groups there is a significant speech effect ($F(3,195)=4$, $p<.05$). No statistically significant interactions were found. Results are presented in table VI.27.

Table VI.27 – Within subjects' effects for QFx

Source	df	Sum of Squares	Mean square	F-test	p (Sig. of F)
Within main effect					
Speech	3	198.351	80.528	4.001	**.013**
Interactions					
Speech* Group	3	53.938	21.898	1.088	.348
Speech* Gender	3	69.685	28.291	1.406	.246
Speech* Age	3	19.824	8.048	.400	.714
Speech* Group * Gender	3	85.252	34.611	1.720	.174
Speech* Group * Age	3	35.449	14.392	.715	.518
Speech* Gender * Age	3	31.813	12.916	.642	.559
Speech* Group * Gender* Age	3	18.123	7.358	.366	.738
Error	195	3916.046	20.125		

Results of the paired-wise comparisons for the significant effect of speech shown in table VI.28 indicate that for the dysphonics there is a statistically significant difference between the vowel [a] and the vowel [i] ($t=-2.18$, $df=38$, $p<.05$) and between the vowel [i] and reading ($t=2.3$, $df=42$, $p<.05$) and between the

vowel [i] and conversation (t=2.3, df=45, p<.05). For the control subjects there is also a statistically significant difference between the vowel [i] and conversation (t=2.4, df=56, p<.05) between the vowel [u] and reading (t=2.2, df=51, p<.05) and between the vowel [u] and conversation (t=3.2, df=55, p<.01).

Table VI.28 – Related t-test for the QFx speech tasks pairs

	Dysphonics			Controls		
F0 standard deviation mean	t	df	p	t	df	p
Vowel [a] versus vowel [i]	-2.178	38	**.036**	-1.044	53	.301
Vowel [a] versus vowel [u]	-1.259	38	.216	-1.655	52	.104
Vowel [i] versus vowel [u]	.907	45	.369	-.855	55	.396
Vowel [a] versus Reading	.182	38	.857	.592	49	.557
Vowel [i] versus Reading	2.264	42	**.029**	1.376	52	.175
Vowel [u] versus Reading	1.612	42	.115	2.234	51	**.030**
Vowel [a] versus Conversation	.107	39	.915	1.680	53	.099
Vowel [i] versus Conversation	2.266	45	**.028**	2.378	56	**.021**
Vowel [u] versus Conversation	1.786	45	.081	3.156	55	**.003**
Reading versus Conversation	-.503	48	.617	1.880	52	.066

The results of the multiple regression analysis in mean QFx during all speech tasks (vowels, reading and conversation), t and p-values and also the overall significance ANOVA regression are presented in table VI.29.

Table VI.29 (below) shows that 7% of the variance in the mean contact quotient for vowel [a] is accounted for by mass larynx lesion and economic factors and that the first variable is the best and positive independent predictor (t=2.38, p<.05) followed by the second that is a significant negative independent predictor (t= - 2.1, p<.05).

As can be seen in table VI.29, 24% of the variation in the mean QFx for vowel [i] can be predicted on the basis of work stress, drinking habit, in employment, number of coffees a day and distributed tissue larynx change. Within the model there are positive and negative significant independent predictors (as shown by the t and p-values). The beta weight coefficients show that the best predictor is alcohol

consumption (-.311) followed by number of coffees a day (.299), in employment (.271), work stress (.264) and distributed tissue larynx change pathology (-.197).

Table VI.29 - Multiple regression analysis in the mean QFx

Vowel [a]				
ANOVA (regression) F(2,87)=4.58, p=.013 Adjusted R Square =.074				
Predictor variables	Coefficient B	Beta weight	t-value	p-value
(Constant)	46.67		36.38	.000
ENT diagnosis (mass lesion)	5.85	.244	2.38	.019
In employment	-3.20	-.209	-2.05	.044
Vowel [i]				
ANOVA (regression) F(5,92)=7.03, p=.000 Adjusted R Square =.237				
Predictor variables	Coefficient B	Beta weight	t-value	p-value
(Constant)	46.64		39.6	.000
SRRS (work stress)	.103	.264	2.97	.004
Alcohol consumption	-7.10	-.311	-3.97	.001
In employment	-3.81	.271	-2.94	.004
Number of coffees a day	1.31	.299	3.15	.002
ENT diagnosis (tissue change)	-4.76	-.197	-2.16	.033
Vowel [u]				
ANOVA (regression) F(5,91)=8,72, p=.000 Adjusted R Square = .287				
Predictor variables	Coefficient B	Beta weight	t-value	p-value
(Constant)	47.67		38.97	.000
SRRS (work stress)	.108	.262	3.03	.003
In employment	-5.10	-.343	-3.86	.000
Alcohol consumption	-7.73	-.321	-3.63	.000
Number of coffees a day	1.51	.322	3.44	.001
Gastroesophageal disease	-3.89	-.216	-2.43	.017
Reading				
ANOVA (regression) F(4,94)=5.58, p=.000 Adjusted R Square = .158				
Predictor variables	Coefficient B	Beta weight	t-value	p-value
(Constant)	47.59		36.88	.000
Number of years living in Lisbon	.093	-.255	-2.72	.008
SLT diagnosis (harsh)	3.53	.219	2.32	.023
In employment	-2.8	-.241	-2.49	.015
Number of coffees a day	.759	.209	2.16	.034
Conversation				
ANOVA (regression) F(6.96)=7.85, p=.000 Adjusted R Square =.287				
Predictor variables	Coefficient B	Beta weight	t-value	p-value
(Constant)	38.80		35.49	.000
SLT diagnosis (harsh)	5.83	.381	4.00	.000
Group (control)	4.32	.408	4.01	.000
SRRS (work stress)	.0742	.244	2.84	.005
Number of coffees a day	.955	.272	3.09	.003
Alcohol consumption	-3.69	-.198	2.31	.023
ENT diagnosis (mass lesion)	3.206	.196	2.11	.037

As can be seen in table VI.29, 29% of the variation in the mean QFx for vowel [u] can be predicted on the basis of the significant positive variable work stress (t=3.03, p<.01) combined with the significant negative variable and the best predictor in the model, in employment (t= -3.86 p<.001), the significant negative variable alcohol consumption (t= -3.63, p<.001), the significant positive variable, number of coffees a day (t= 3.44, p<.001) and the negative variable, gastroesophageal disease (t=-2.43, p<.05).

16 % of the variation in the overall quality of the mean QFx during reading (table VI.29) can be predicted on the basis of the combined effects of the best significant independent predictor variable, number of years living in Lisbon (-.255) followed by in employment (.-241), harsh voice (.219), and number of coffees a day (.209).

As can be seen in table VI.29, 29 % of the variation in the overall quality of the mean QFx during conversation can be predicted on the basis of the following combined variables, harsh voice, normal voice, work stress, number of coffees a day, beer drinking habit, and mass larynx lesion. Within the model an exception is the variable alcohol consumption that is negative; all the other 5 predictor variables are positive independent variables. The best predictor is control group followed by harsh voice, number of coffees a day, work stress, drinking habit and larynx mass lesion pathology (as might be seen in the beta weight coefficients in table VI.29).

In summary, from the main effects in the study only speech had a significant statistical effect on the mean contact closed quotient ($F(3,195)=4$, $p<.05$). No significant statistical interactions were found.

The predictive analyses showed that 7 to 29% of the mean contact closed quotient variance during all speech tasks is explained by a combination between two and six variables. Larynx mass lesion pathology (beta weight=.244) is the best predictor variable for the QFx variance in the vowel [a], alcohol consumption (beta weight=-.311) for the vowel [i], in employment (beta weight=-5.10) for the vowel [u], number of years living in Lisbon (beta weight=.093) for reading and 'normal voice (beta weight=.408) for the conversation. Within the models the other combining variables are demographic (in employment), social habits (coffee intake and alcohol consumption) and health

related variables (stress, larynx diagnosis, diseases and voice diagnosis). The final multiple regression equations that predict the mean QFx for all speech tasks are respectively:

- Vowel [a] = 46.7 + 5.85 (mass larynx lesion) – 3.2 (in employment);

- Vowel [i] = 46.64 – 7.10 (alcohol consumption) + 1.31 (number of coffees a day) – 3.81 (in employment) + .103 (work stress) – 4.76 (distributed tissue larynx change);

-Vowel [u] = 47.7 – 5.1 (in employment)– 7.73 (alcohol consumption) + 1.51 (number of coffees a day) + .108 (work stress) – 3.89 (gastroesophageal disease);

- Reading = 47.6 + .093 (number of years living in Lisbon) – 2.8 (in employment) + 3. 53 (harsh voice) + .759 (number of coffees a day);

- Conversation = 38.8 + 4.32 (control group) + 5.8 (harsh voice) + .96 (number of coffees a day) + .07 (work stress)– 3.7 (alcohol consumption) + 3.21 (larynx mass lesion pathology).

6.2.7 –QFx standard deviation

Means of contact closed quotient standard deviation for all the speech tasks according to group, gender and age are presented in table VI.30.

Table VI.30 – Means of QFx standard deviation for all the speech tasks

	Vowel [a]	Vowel [i]	Vowel [u]	Reading	Conversation
Group					
Dysphonics	.18	.15	.17	6.8	8.0
Controls	.14	.13	.09	6.0	7.4
Gender					
Females	.16	.13	.10	6.1	7.3
Males	.10	.12	.11	5.9	7.5

Overall, means of QFx standard deviation of the dysphonics are higher than that of the controls and the females show higher values for the vowels [a], [i] and reading than the male speakers (table VI.30).

It is clear from table VI.31 that no statistical significant differences or interactions were observed in the multiple repeated analyses performed for the mean QFx standard deviation.

Table VI.31 – Between subjects' effects for QFx standard deviation

Source	df	Sum of Squares	Mean square	F-test	p (Sig. of F)
Main effects					
Group	1	2.319	2.319	1.906	.171
Gender	1	.763	.763	.627	.431
Age	1	3.977	3.977	3.269	.074
Interactions					
Group* Gender	1	.352	.352	.289	.592
Group*Age	1	1.381	1.381	1.135	.290
Gender* Age	1	1.317	1.317	1.082	.301
Group* Gender* Age	1	1.610	1.610	1.323	.254
Error	79	96.121	1.217		

The multiple comparison analyses revealed that within the groups there is a highly significant speech effect ($F(2,133)=1419$, $p<.001$). No statistically significant interactions were found. Data is presented in table VI.32.

Table VI.32 – Within subjects' effects for the mean QFx standard deviation

Source	df	Sum of Squares	Mean square	F-test	p (Sig. of F)
Within main effect					
Speech	2	3395.006	2021.764	1419.138	**.000**
Interactions					
Speech* Group	2	5.240	3.121	2.190	.124
Speech* Gender	2	4.191	2.496	1.752	.183
Speech* Age	2	5.553	3.307	2.321	.111
Speech* Group * Gender	2	1.045	.622	.437	.612
Speech* Group * Age	2	3.625	2.159	1.515	.225
Speech* Gender * Age	2	4.073	2.425	1.702	.191
Speech* Group * Gender* Age	2	3.479	2.072	1.454	.238
Error	133	188.992	1.425		

Post hoc paired t-test comparisons were carried out so that significant differences between the various speech tasks could be assessed. It is clear from table VI.33 that there is no statistically significant mean QFx standard deviation difference between the vowels. However, there is a statistically significant mean QFx standard deviation

difference between all the vowels and connected speech (reading and conversation) and between reading and conversation (except for the control subjects).

Table VI.33 – Related t-test for the QFx standard deviation speech tasks pairs

	Dysphonics			Controls		
F0 standard deviation mean	t	df	p	t	df	p
Vowel [a] versus vowel [i]	.084	38	.933	1.249	53	.217
Vowel [a] versus vowel [u]	1.020	38	.314	.066	52	.948
Vowel [i] versus vowel [u]	.654	45	.517	-1.709	55	.093
Vowel [a] versus Reading	4.996	38	**.000**	3.550	49	**.001**
Vowel [i] versus Reading	5.165	42	**.000**	5.822	52	**.000**
Vowel [u] versus Reading	5.410	42	**.000**	3.963	51	**.000**
Vowel [a] versus Conversation	5.122	39	**.000**	3.780	53	**.000**
Vowel [i] versus Conversation	5.567	45	**.000**	6.492	56	**.000**
Vowel [u] versus Conversation	5.798	45	**.000**	4.270	55	**.000**
Reading versus Conversation	-5.805	48	**.000**	-.669	52	.507

The results of the multiple regression analysis in theQFx standard deviation during all speech tasks (vowels, reading and conversation), t and p-values and also the overall significance ANOVA regression are presented in table VI.34.

The combining effect of the two positive independent variables, born in the center of Portugal, (t=3.92, p<.05) the best predictor (.254) followed by the number of cigarettes per day (t=2.44, p<.05) the second best predictor (.251) shown in table VI.34 accounted for 9% of the mean QFx standard deviation in vowel [a].

Results in table VI.34 indicate that none of the variables made a contribution to producing a model for vowel [i] mean QFx standard deviation.

The results shown in table VI.34 indicate that 41% of the variance in the mean QFx standard deviation of vowel [u] is accounted for by the combined effect of the variables: being born in the islands, harsh voice, mass larynx lesion and personal stress. All the variables are positive independent predictors. The best predictor is being born in the

islands (.440) followed by harsh voice (.309), mass larynx lesion pathology (.271) and personal stress (-.182).

Table VI.34 - Multiple regression analysis in the QFx standard deviation

Vowel [a]				
ANOVA (regression) F(2,87)=5.16, p=.008 Adjusted R Square = .085				
Predictor variables	Coefficient B	Beta weight	t-value	p-value
(Constant)	.097			.000
Being born in the center of Portugal	.141	.254	3.92	.016
Number of cigarettes a day	.0056	.251	2.47	.017
Vowel [i]				
No model				
Vowel [u]				
ANOVA (regression) F(4,92)=17.36, p=.000 Adjusted R Square = .405				
Predictor variables	Coefficient B	Beta weight	t-value	p-value
(Constant)	.103		6.44	.000
Born in the islands (Madeira or Azores)	.458	.440	5.48	.000
SLT diagnosis (harsh)	.134	.309	3.59	001
ENT diagnosis (mass lesion)	.122	.271	3.26	002
SRRS (personal stress)	.00055	-.182	-2.26	027
Reading				
ANOVA (regression) F(4,94)=8,13, p=.000 Adjusted R Square = .225				
Predictor variables	Coefficient B	Beta weight	t-value	p-value
(Constant)	6.56		14.02	.000
Group (control)	-1.18	-.436	-4.25	.000
SLT severity diagnosis (slight)	-1.49	-.385	-3.79	.000
Age	.024	.248	2.72	.008
Number of coffees a day	-.168	-.186	-2.01	.047
Conversation				
ANOVA (regression) F(6,96)=7.48, p=.000 Adjusted R Square = .276				
Predictor variables	Coefficient B	Beta weight	t-value	p-value
(Constant)	8.65		22.17	.000
Cause of dysphonia (tobacco)	2.14	.377	4.17	.000
Born in the islands (Azores or Madeira)	3.36	.324	3.81	.000
Cause of dysphonia (work)	1.00	.225	2.64	.010
Number of coffees a day	-.197	-.206	-2.43	.017
SRRS (stress total)	-.004	-.219	-2.44	.016
Years of education	-.052	-.176	-2.06	.042

The results shown in table VI.34 indicate that 23% of the variance in the mean QFx standard deviation during reading accounted by the combined effect of the variables, normal voice, slight dysphonia severity, age and number of coffees a day. The variables are positive and negative independent predictors. The best predictor is control

group (-.436) followed by slight dysphonia severity (-.385), number of coffees a day (.248) except age that is a positive independent variable (-.186).

It is clear from table VI.34, that 28% of the variance in the mean QFx standard deviation during conversation is accounted for by the combined effects of tobacco as a cause of dysphonia ($t=4.17$, $p<.001$) as the best predictor in the model followed by being born in the islands ($t=3.81$, $p<.001$), work as a cause of dysphonia ($t=2.64$, $p<.01$), number of coffees per day ($t=-2.4$, $p<.05$), stress total ($t=-2.44$, $p<.05$) and years of education ($t=-2.1$, $p<.05$).

In summary, speech has a highly significant statistical effect on the mean contact closed quotient standard deviation ($F(2,133)=1419.14$, $p<.001$). No other significant statistically main effects or interactions were found.

The predictive analyses showed that 9 to 41% of the QFx standard deviation variance (for all speech tasks except the vowel [i]) is explained by a combination of two and six variables. Being born in the centre of Portugal (beta weight=.254) is the best predictor variable for the QFx variance in the vowel [a], born in the islands (beta weight=.440) for the vowel [u], 'normal' voice (beta weight=-.436) for reading and tobacco as a cause of dysphonia (beta weight=.377) for conversation. Other variables brought out in the models are demographic (age and years of education), social habit (coffee intake) and health (stress, larynx pathology, voice diagnosis and voice severity diagnosis). The final multiple regression equations that predict mean QFx standard deviation are:

-Vowel [a] = .097 + .141 (being born in the center of Portugal) + .006 (number of cigarettes a day);

-Vowel [u] = .103 + .46 (being born in the islands) + .134 (harsh voice) + .122 (mass larynx lesion) + .0006 (personal stress);

-Reading = 6.6 – 1.2 (control group) – 1.49 (slight dysphonia severity) + .024 (age) - .17 (number of coffees a day);

-Conversation = 8.7 + 2.14 (tobacco use as a cause of dysphonia) + 3.4 (being born in the islands) + 1 (work as a cause of dysphonia) - .197 (number of coffees a day) - .004 (stress total) - .052 (years of education).

6.3 – Subjects' self-rated psychosocial voice impact

The VHI results will be presented according to dependent variables: (a) total score scale; (b) emotional sub-score scale; (c) functional sub-score scale and, (d) physical sub-score scale. Those variables were modelled against the independent variables: (a) group, (b) gender and (c) age.

Table VI.35 – Mean VHI

			Total	Emotional	Functional	Physical
Group		N	M ± SD*	M ± SD*	M ± SD*	M ± SD*
	Dysphonics	49	34.35 ± 3.19	8.45 ± 1.04	8.49 ± 1.20	17.41 ± 1.22
	Controls	56	10.54 ± 1.79	2.02 ± .60	3.34 ± .58	5.16 ± .72

* M (mean) ± SD (standard deviation)

According to the data presented in table VI.35 the mean and standard deviation total and sub-total scores for the subjects' self-rated psychosocial voice impact score are higher for dysphonics than for the control speakers. Although variability (standard deviation) is higher in the dysphonics than in the controls some outliers are also found in the controls as shown in figure VI.13.

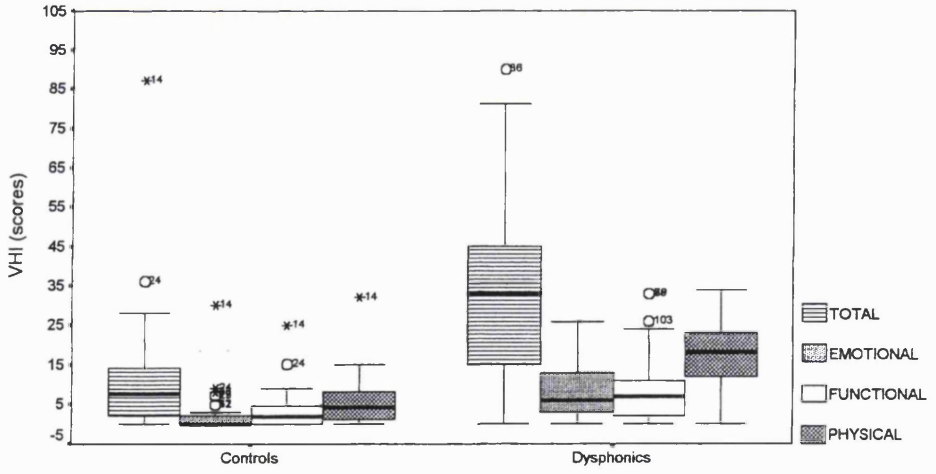


Figure VI.13 – VHI boxplots for the controls and dysphonics

It is clear from figure V.13 that the VHI scores distribution is: (1) more skewed (seen by the location of the thick black bar across the interior of the box that represents the median) in the control speakers; (2) the whiskers (vertical lines outside the box) show

that the largest and smallest VHI scores (that are not considered an outlier or an extreme score) are wider in the dysphonics than the controls; (3) both groups show outlier values (values that are more than 1.5 box-lengths away from the box and are represented with the symbol 'o'); (4) the controls show extreme values (more than 3 box-lengths away from the box, represented by the symbol '*').

In fact a detailed analysis of the frequency distribution of the ratings according to each subscale (in appendix H) shows that in the control group more than 80% of the speakers rate their answers (in all subscales) between 'never' and 'almost never' with an exception for the question 'F3 People have difficulty understanding me in a noisy room' where the frequency on the abovementioned answers ('never' and 'almost never') is 63% and in 'sometimes'nearly 27%. The rating answers for the dysphonic speakers vary widely along the possible 5 answers (from 'never' and 'always') as might be seen in appendix H.

Multiple repeated analyses were performed to examine whether statistical differences exist according to the main effects in the study (group, gender and age) and interactions. Only group had a highly statistically significant effect ($F(1,97)=28.16$, $p<.001$) (table V.36) confirming that the higher mean for the dysphonic speakers is statistically different than that of the controls.

Table VI.36 – Between subjects' effects for VHI

Source		df	Sum of Squares	Mean square	F-test	p (Sig. of F)
Main effects	Group	1	9143.555	9143.555	28.157	**.000**
	Gender	1	435.829	435.829	1.342	.250
	Age	1	19.449	19.449	.060	.807
Interactions						
	Group* Gender	1	3.202	3.202	.010	.921
	Group*Age	1	548.889	548.889	1.690	.197
	Gender* Age	1	226.980	226.980	.699	.405
	Group* Gender* Age	1	163.140	163.140	.502	.480
	Error	97	31499.188	324.734		

The multiple comparison analyses revealed that there is a highly significant difference within the VHI subscales ($F(1,130)=85.53$, $p<.001$) and a highly significant

interaction with the between subjects group ($F(1,130)=27.36$, $p<.001$). Data is presented in table VI.37. Comparison of means in table VI.35 shows that the physical scores are higher than the emotional and functional scores for both dysphonics and controls.

Table VI.37 – Within subjects' effects for VHI

Source	df	Sum of Squares	Mean square	F-test	p (Sig. of F)
Within main effect					
VHI	1	10854.323	8072.369	85.531	**.000**
Interactions					
VHI* Group	1	3472.495	2582.497	27.363	**.000**
VHI* Gender	1	198.463	147.597	1.564	.216
VHI* Age	1	9.418	7.004	.074	.856
VHI* Group * Gender	1	21.304	15.844	.168	.756
VHI* Group * Age	1	188.031	139.838	1.482	.231
VHI* Gender * Age	1	100.976	75.096	.796	.408
VHI* Group * Gender* Age	1	63.048	46.889	.497	.536
Error	130	12309.768	94.379		

The significant interaction (VHI-group) resulting from these analyses, was followed up by post hoc comparisons (shown in table VI.38) so that significant differences between the combinations could be assessed. Results confirmed a statistically significant difference between the total and sub-total scores for all the subjects with an exception for the difference between emotional and functional scores in the dysphonics.

Table VI.38 – Related t-test for the VHI

	Dysphonics			Controls		
VHI means	t	df	p	t	df	p
VHI total versus VHI emotional	11.601	48	**.000**	6.907	55	**000**
VHI total versus VHI functional	12.081	48	**000**	5.688	55	**000**
VHI total versus VHI physical	7.816	48	**000**	4.717	55	**000**
VHI emotional versus VHI functional	.153	50	.879	-4.276	55	**000**
VHI emotional versus VHI physical	-11.932	48	**000**	-7.889	55	**000**
VHI functional versus VHI physical	-9.359	48	**000**	-4.256	55	**000**

Table VI.39 presents the results of the multiple regression analysis in VHI, t and p-values and also the overall significance ANOVA regression.

Table VI.39 - Multiple regression analysis in VHI

VHI Total				
ANOVA (regression) F(2,56)=15.391, p=.000 R Squared (adjusted) = .332				
Predictor variables	Coefficient B	Beta weight	t-value	p-value
(Constant)	36.214		9.148	.000
Group (control)	-25.770	-.700	-5.524	.000
Cause of dysphonia (disease)	-15.770	-.316	-2.492	.016
VHI Emotional sub-scale				
ANOVA (regression) F(3,56)=10.25, p=.000 R Squared (adjusted) = .320				
Predictor variables	Coefficient B	Beta weight	t-value	p-value
(Constant)	2.108		1.503	.138
Dysphonia onset (gradual)	7.237	.485	4.487	.000
F0 standard deviation (vowel i)	.736	.342	3.153	.003
In employment	-3.427	-.242	-2.214	.031
VHI Functional sub-scale				
ANOVA (regression) F(1,58)=8.728, p=.005 R Squared (adjusted) =.116				
Predictor variables	Coefficient B	Beta weight	t-value	p-value
(Constant)	3.341		4.300	.000
Dysphonia onset (gradual)	4.080	.362	2.954	.005
VHI Physical sub-scale				
ANOVA (regression) F(2,56)=31.95, p=.000 R Squared (adjusted) =.516				
Predictor variables	Coefficient B	Beta weight	t-value	p-value
(Constant)	13.707		9.332	.000
Group (control)	-9.848	-.558	-5.897	.000
Jitter (vowel i)	1.820	.347	3.663	.001

As can be seen in table VI.39, 33% of the variance in the VHI total score can be predicted on the basis of the significant negative independent variables 'control group' ($t=-5.5$, $p<.001$) and subject's reported disease as a cause of dysphonia ($t=-2.5$, $p<.05$).

The overall combining model (gradual dysphonia onset, vowel [i] F0 standard deviation and economic activity) shown in table VI.39 is a highly significant predictor ($p<.001$) model of VHI emotional score, accounting for 32% of its variance. Each predictor in the model is an independent significant variable; dysphonia onset ($t=4.49$, $p<.001$), vowel [i] F0 standard deviation ($t=3.15$, $p<.01$), in employment ($t=-2.2$, $p<.05$).

Table VI.39 shows that 12% of variance of VHI functional score is accounted for by gradual dysphonia onset and that it is a significant positive predictor variable ($t=2.95$, $p<.01$).

52% of the variation in the overall VHI physical score (table VI.39) can be predicted on the basis of significant independent variables, controls ($t=-5.89$, $p<.001$) and jitter for vowel [i] ($t=3.66$, $p<.001$).

In summary, a significant statistical difference between the dysphonic and control speakers was found for the VHI total score with the dysphonics presenting higher score, therefore worse VHI than the controls. Within the groups the scores are statistically different except for the subscale emotional versus functional in the dysphonic speakers.

Results indicate that multiple prediction of VHI variance is between low (12%) for the functional score, moderate (33% and 32%) for the total and emotional score respectively and high (52%) for the physical score.
12% of the functional score variance is associated with gradual dysphonia onset meaning that dysphonics with a gradual onset will show a higher functional score (and consequently worst impact). Also, interestingly 33% of the emotional score variance is associated with 'gradual dysphonia onset', 'F0 standard deviation on vowel [i]' and 'in employment'. Nevertheless, the first two variables contribute to a higher emotional score (and consequently worst impact) whilst 'in employment' contributes to a lower emotional score (and consequently lighter impact).
33% of the total score variance is significantly associated with 'control speaker' and 'disease as a cause of dysphonia' meaning that the final total score will be lower when those variables are present in the model. Also, 52% of the physical score variance depends on the variable 'control', that contributes to its lowering (and consequently lighter impact) and 'jitter of vowel [i]' that contributes to its increase (and consequently worst impact). The final multiple regression equations that predict variance for the VHI are:

-VHI total score = 36.2 - 25.8 (control group) – 15.8 (subject's reported disease as a cause of dysphonia);

-VHI emotional score = 2.108 + 7.24 (gradual dysphonia onset) + .736 (F0 standard deviation on vowel [i]) – 3.43 (in employment);

-VHI functional score = 3.34 + 4.08 (gradual dysphonia onset);

-VHI physical score = 13.71 – 9.85 (controls) + 1.82 (jitter for vowel [i]).

Additionally for exploratory purposes data analyses of the VHI scores according to the SLT diagnostic severity and the ENT diagnosis was performed and is shown in table VI.40. Results show that scores are higher for the worst severity SLT diagnosis and within the ENT diagnosis 'healthy' and 'minor variations' present lower scores than other categories 'functional', 'mass' and 'tissue' disorders.

Table VI.40 – VHI data according to SLT and ENT diagnosis

		Total	**Emotional**	**Functional**	**Physical**
SLT severity	N	M ± SD*	M ± SD*	M ± SD*	M ± SD*
Normal	60	11.2 ± 14	2.1 ± 4.5	3.6 ± 4.8	5.5 ± 5.7
Slight	14	26.7 ± 17	6.7 ± 5.9	6 ± 5.1	14 ± 7.2
Moderate	17	34 ± 25	9.4 ± 9.6	9.8 ± 10.6	16.9 ± 8.3
Severe	16	44.8 ± 22	11.6 ± 7.8	10.6 ± 8.4	22.6 ± 7.7
ENT diagnosis					
Healthy larynx	35	12.7 ± 18	2.6 ± 6.1	3.6 ± 5.1	6.4 ± 7.7
Minor variation	25	9.7 ± 8.2	1.6 ± 2.4	3.0 ± 3.0	5.0 ± 4
Functional disorder	24	37 ± 25.3	9.6 ± 7.8	10.3 ± 9.7	16.8 ± 9.2
Mass disorder	13	33.5 ± 15.9	9.5 ± 8.8	8.1 ± 7.7	19.3 ± 6.6
Tissue change	10	33.6 ± 21.8	8.1 ± 7.7	8.1 ± 7.9	17.4 ± 8.2

* M (mean) ± SD (standard deviation)

A summary of the results of the present study is shown in table VI.41, the voice data obtained is presented in table VI.42 and the predictive formulas for the voice characteristics in table VI.43. Moreover, VHI data and the predictive formulas for the voice psychosocial impact obtained in the present study are shown in table VI.44 and VI.45 respectively.

Table VI.41 – Summary of the findings in the present study

The mean F0: - is significantly higher for the females than for the males; - is significantly higher for all vowels than for connected speech (reading and conversation); - is significantly higher for reading than for conversation in the female speakers; - values distribution although not significant are wider in the dysphonics than in the controls; - values overlap between females and males within the dysphonics whereas in the controls is only between the reading from males and reading and conversation from females.
The mean F0 standard deviation: - is not significantly different between all vowels; - is not significantly different between reading and conversation (except for the older control speakers); - is significantly different between all vowels and connected speech for all speakers; - values are wider for all vowels than for connected speech for all speakers; - values are wider in dysphonics than controls.
The 90% F0 range: - is significantly wider during conversation than reading; - is significantly narrower for the young controls than that for the older ones during reading and conversation; - is significantly narrower for the young dysphonics than that for the older ones during conversation.
The percentage of jitter: - is wider in the dysphonics than controls although not statistically significant; - is wider for the vowel [i] in the dysphonics although not statistically significant; - is wider for the vowel [a] in the female controls and for the [u] in the male controls although not statistically significant.
The percentage of irregularity: - is significantly higher for the dysphonics than for the controls in reading and conversation; - is significantly higher for the males than for the females during conversation; - is significantly higher for the older females than for the young ones during reading.
The mean contact quotient (QFx): - is significantly different between the vowel [a] and the vowel [i] for the dysphonics; - is significantly different between the vowel [i] and the connected speech (reading and conversation) for the dysphonics; - is significantly different between the vowel [u] and the connected speech for the controls; - is significantly different between the vowel [i] and conversation for the controls.
The QFx standard deviation: - is significantly different between all vowels and the connected speech (reading and conversation) for all speakers; - is significantly different between reading and conversation for the dysphonics.

Table VI.42 – Voice data obtained in the present study

	Dysphonics (n=52)					Controls (n=57)				
	[a]	[i]	[u]	Reading	Conversation	[a]	[i]	[u]	Reading	Conversation
Mean F0 (Hz)										
Females	199.49	212.67	213.98	183.38	179.74	210.51	220.96	225.06	190.33	186.61
Males	113.02	130.18	128.09	111.20	108.26	118.35	127.58	122.99	109.59	109.21
SD F0 (Hz)										
Females	4.8	4.9	4.9	.09	.11	5.6	3.6	3.6	.12	.12
Males	7.2	6.8	5.2	.11	.14	1.04	1.56	7.04	.09	.12
90% F0 range (octaves)										
Females				.85	1.1				.86	1.1
Males				1.0	1.4				.75	1.1
% Jitter										
Females	1.5	1.8	1.5			1.3	.73	.64		
Males	1.1	2.0	1.3			.61	.69	.81		
% Irregularity										
Females				14.50	17.3				7.6	12.3
Males				19.12	28.4				12.5	23.5
Mean QFx (%)	44.80	46.47	46.01	44.17	44.59	45.34	46.05	46.44	44.89	44.27
SD QFx (%)	.18	.15	.17	6.8	8.0	.14	.13	.09	6.0	7.4

Table VI.43 – Summary of the predictive formulas for the voice characteristics

Mean F0	**Vowel [a]** = 218.52 – 80.58 (male) – 1.49 (nº of cigarettes per day) + 39.14 (sudden dysphonia onset) – 18.98 (minor nose abnormalities); **Vowel [i]** = 214.87 – 81.84 (male) – 2.04 (nº of cigarettes per day) + 5.94 (nº of coffees a day) – 25.34 (disease as a cause of dysphonia) + 22.95 (hoarse voice); **Vowel [u]** = 223.40 – 91.06 (male) – 50.06 (tobacco as a cause of dysphonia) – 35.68 (slight dysphonia severity) - 21.99 (smoker) +. 104 (total stress); **Reading** = 217.22 – 70.38 (male) - 1.25 (nº of cigarettes per day) - .759 (age) + 27.65 (psychological factors as a cause of dysphonia) + 21.41 (sudden dysphonia onset); **Conversation** = 185.55 – 62.11 (male) - 1.09 (nº of cigarettes per day) – 26.16 (tobacco as a cause of dysphonia) + 20.34 (psychological factors as a cause of dysphonia).
Mean F0 standard deviation	**Vowel [a]** = -1.57 – 9.82 (larynx tissue change) + 11.40 (being born in Alentejo or the Algarve) + 8.82 (tobacco as a cause of dysphonia) + .131 (age) + .203 (nº of cigarettes per day); **Vowel [i]** = 5.81 + 6.93 (tobacco as a cause of dysphonia) + 5.37 (sudden dysphonia onset) – 5.07 (larynx distributed tissue change) + .028 (personal stress) + .047 (financial stress) + .084 (nº of years living in Lisbon) - 2.22 (normal voice) - 3.69 (alcohol consumption); **Vowel [u]** = 3.9 -.059 (financial stress) + 10.71 (being born in the islands); **Reading** = .087 + .0.19 (nº of pills per day); **Conversation** = .096 + .0006 (male) + .096 (age) - .015 (alcohol consumption).
90% F0 range	**Reading** = .802 + .254 (larynx distributed tissue change) +.216 (harsh voice) + .21 (menstruation) - .18 (alcohol consumption); **Conversation** = .77 + .297 (minor nose abnormalities) – .369 (breathy voice) + .013 (nº of cigarettes per day) - .07 (nº of coffees a day) + .007 (age) + .398 (psychological factors as a cause of dysphonia).
% Jitter	**Vowel [a]** = 2.8 + 3.5 (being born in Alentejo or the Algarve) - .12 (years of education) – 93 (minor larynx variation); **Vowel [i]** = 1.44 - 1.59 (control group) - 1.95 (slight dysphonia severity) + .033 (nº of years living in Lisbon); **Vowel [u]** = .86 + 1.92 (being born in the North) + 1.31 (larynx functional disorder) - .015 (financial stress) + 1.71 (ovulation phase).
% Irregularity	**Reading** = 8.09 + 8.8 (hoarse voice) + 12.74 (tobacco as a cause of dysphonia) + 5.39 (being born in the North) + 3.9 (male); **Conversation** = 19.18– 17.46 (tobacco as a cause of dysphonia) + 25.19 (being born in the islands) + 7.85 (male) – 2.09 (nº of coffees a day) + 7.17 (hoarse voice) – .034 (total stress).
Mean QFx (%)	**Vowel [a]** = 46.7 + (mass larynx lesion) – 3.2 (in employment); **Vowel [i]** = 46.64 – 7.10 (alcohol consumption) + 1.31 (nº of coffees a day) – 3.81 (in employment) + .103 (work stress) – 4.76 (distributed tissue larynx change); **Vowel [u]** = 47.7 – 5.1 (in employment)– 7.73 (alcohol consumption) + 1.51 (nº of coffees a day) + .108 (work stress) – 3.89 (gastroesophageal disease); **Reading** = 47.6 + .093 (number of years living in Lisbon) – 2.8 (in employment) + 3.53 (harsh voice) + .759 (nº of coffees a day); **Conversation** = 38.8 + 4.32 (control group) + 5.8 (harsh voice) + .96 (nº of coffees a day) + .07 (work stress)– 3.7 (alcohol consumption) + 3.21 (larynx mass lesion).
Mean QFx standard deviation	**Vowel [a]** = 46.7 + (mass larynx lesion) – 3.2 (in employment); **Vowel [i]** = 46.64 – 7.10 (alcohol consumption) + 1.31 (nº of coffees a day) – 3.81 (in employment) + .103 (work stress) – 4.76 (distributed tissue larynx change); **Vowel [u]** = 47.7 – 5.1 (in employment)– 7.73 (alcohol consumption) + 1.51 (nº of coffees a day) + .108 (work stress) – 3.89 (gastroesophageal disease); **Reading** = 47.6 + .093 (nº of years living in Lisbon) – 2.8 (in employment) + 3.53 (harsh voice) + .759 (number of coffees a day); **Conversation** = 38.8 + 4.32 (control group) + 5.8 (harsh voice) + .96 (nº of coffees a day) + .07 (work stress)– 3.7 (alcohol consumption) + 3.21 (larynx mass lesion).

Table VI.44 - VHI data obtained in the present study

VHI mean score	Dysphonics (n=52)	Controls (n=57)
Total	34.35	10.54
Emotional	8.45	2.02
Functional	8.49	3.34
Physical	17.41	5.16

Table VI.45 – Summary of the predictive formulas for the VHI

VHI
Total = 36.2 – 25.8 (control group) – 15.8 (disease as a cause of dysphonia);
Emotional = 2.11 + 7.24 (gradual dysphonia onset) + 0.74 (Fo standard deviation on vowel [i]) – 3.43 (in employment);
Functional = 3.34 + 4.08 (gradual dysphonia onset);
Physical = 13.71 – 9.85 (controls) + 1.82 (jitter for the vowel [i]).

Chapter VII - Discussion

7.1 – Introduction

In the following chapter the discussion of the findings is structured on the basis of the aims of this prospective exploratory study: (1) the voice characteristics of dysphonic Portuguese speakers compared to a group of 'normal' Portuguese speakers, (2) the variables that contribute most importantly to predicting variance in the voice characteristics, (3) the dysphonic Portuguese speakers' self-rated psychosocial voice impact compared to a group of non dysphonic Portuguese speakers, and (4) the variables that contribute most importantly to predicting variance in the speakers' self-rated psychosocial voice impact.

Discussion will follow the same order presented in the previous chapter according to the voice dependent measures, (a) mean F0; (b) mean F0 standard deviation; (c) F0 range; (d) jitter; (e) percentage of irregularity, (f) mean contact closed quotient; (g) mean contact closed quotient standard deviation and to the VHI dependent measures (a) total score, (b) emotional score, (c) functional score and (d) physical score.

Each voice parameter in the study will be individually interpreted and proposed explanations for the findings will be incorporated and compared with earlier findings reported in the literature review.

7.2 – Voice characteristics

7.2.1- Mean fundamental frequency (F0)

The mean F0 data in this study for all the control speakers is within the limits of the published data reported in the literature review (chapter III) but for other languages (e.g. English, German, Finnish and Swedish). Moreover, the gender differences (where females present higher mean F0 than males) and speech tasks differences (sustained vowels present systematically higher F0 mean than connected speech) found are also in agreement with the published data reported in the literature review (chapter III).

The results of this study failed to indicate a statistically significant difference between dysphonics and control speakers. Nevertheless, the mean F0 of the dysphonic

speakers was lower than that of the control speakers for all speech tasks in the females and for the vowel [a] and conversation in the male speakers. These findings are in line with previous research wherein it was stated that mean F0 did not statistically distinguish dysphonic speakers from 'normal' speakers (Hecker and Kreul, 1971; Murry, 1978; Orlikoff et al., 1997; Buekers, 1998a).

The observed F0 trends in this study showing that the dysphonic speakers have lower mean F0 than the 'normal' speakers is also in concert with some previous research (Fritzell, Sundberg and Strange-Ebbesen, 1982; Peppard, Bless & Milenkovic, 1988; Drew and Sapir, 1995). And also the the fact that the male dysphonics during reading show higher mean F0 than 'normals' (in the present study) is in line with the Murry (1978), Hecker and Kreul (1971) and Verdonck-de Leuuw (1998) studies.

Although overall, the results of this study are similar to those in the literature review (as already mentioned) it should be stressed that Hecker and Kreul (1971), Orlikoff et al. (1997) and Verdonck-de Leeuw (1998) studies involved only subjects with laryngeal cancer whereas the current study involves a wide range of slight to moderate laryngeal pathologies (including only two pre-cancerous pathologies). In addition, given the possible degree of variability amongst the different pathologies in this study collapsed into groups roughly approximating the different traditional categories (e.g. structural and dynamic alterations) it would be important to question the validity of comparing different pathologies within the same group to a 'normal' group. Moreover, within the same larynx pathology two different subjects may present simple or compounded alterations at different stages of evolution. Also, it seems pertinent to emphasise that voice quality is a continuum where it is not possible to establish a unique and perfect relationship between the larynx pathology type and the dysphonia (Koike, Takahashi & Calcaterra, 1977; Baken & Orlikoff, 1988; Miluntinovic, 1996).

The fact that gender had a significant effect on the mean F0 is not surprising, taking into consideration the well-known gender dimorphism and the high degree of consensus that female's F0 is generally higher than male's F0 (Linke, 1973; Krook, 1988; Greene and Mathieson, 1989; Titze, 1994b; Russell, Penny & Pemberton, 1995; Baken & Orlikoff, 2000). Examination of the gender differences in this study revealed that mean F0 for women are approximately between 5 to 6 semitones higher than those from men. In common with the current study, Snidecor (1943) and Linke's (1973)

findings are that the females' pitch level is somewhat less than one octave higher than that of the males.

The results of this study failed to indicate any significant age effect on the mean F0. Despite that, the young adults showed a slightly higher mean F0 than the older adults for almost all speech tasks except for conversation. This fact is in agreement with the published research wherein it was stated that both genders demonstrate a decrease of F0 mean level with increasing age (Honjo & Isshiki, 1980; Linville & Korabic, 1987; Linville, 1988; Krook, 1988; Biever & Bless, 1989; Russel, Penny & Pemberton, 1995).

It is also clear from the preceding results within this study that speech tasks have a highly significant effect on the mean F0. Post hoc comparisons of the results showed several trends. Firstly, the mean F0 for all the sustained vowels in this study is significantly higher compared to that in reading and conversation. Various studies that examined the speech material influence on F0 also found that a higher F0 in the sustained vowels exists compared to that in reading and conversation (Fitch, 1990; Higgins, Netsell & Schulte, 1994; Murry, Brown & Morris, 1995).

Secondly, although a significant vowel effect was found between the vowel [a] and the other two vowels [i] and [u], no significant difference was found between the vowels [i] and [u]. This data supports the previously well-established finding of the influence of vowel tongue height on the phonatory action where there is a highly significant intrinsic F0 for the high vowels (e.g. [i] and [u]) than for the low vowels such as [a] (Viana, 1987; Sapir, 1989; Sussman & Sapienza, 1994; Whalen & Levitt, 1995; Gelfer, 1995; Kingston et al., 1997; Whalen et al., 1998; Fischer-Jorgensen, 1990; Higgins, Netsell & Schulte, 1998).

Front and back vowel effect was not significant although the vowel [u] had a slightly higher F0 than the vowel [i] in the 'normal' female speakers. The results for 'normal' subjects are in agreement with Ewan's (1975 cited by Fischer-Jorgensen, 1990) findings.

Thirdly, a significant higher mean F0 during reading than conversation was restricted to female subjects, which is in agreement with Murry & Brown's (1982) findings. Although this study did not demonstrate a significant effect for the 'normal', 'dysphonics' and male subjects the mean F0 during oral reading is also slightly higher than during conversation and this finding is corroborated by the observations made by Mysak (1959), Sorensen & Horii (1982) and Hollien, Hollien & De Jong (1997).

The results of the multiple regression analysis show that some predictor variables are associated with a lower mean F0 (gender-male, age, number of cigarettes a day, smoker, disease and tobacco as a cause of dysphonia, minor nose abnormalities and slight dysphonia severity).

The gender result, meaning that males have lower mean F0 than females (as has been already mentioned) it is not surprising and corresponds to non-controversial evidence of gender F0 differences predominantly cited in the published research (Linke, 1973; Greene & Mathieson, 1989; Titze, 1994b; Hollien, 2000).

Age accounted for the mean F0 lowering only during reading. Since data from the present study showed a significant negative correlation of age with years of education (r=-.488), the lowering of F0 might simply reflect an effect of level of education.

The second most important predictor that accounts for a lower mean F0 (in all the models) is related to smoking habit (either the number of cigarettes a day, or the fact of being a smoker or even tobacco as aetiology of dysphonia).
It is interesting but not surprising to find that smoking behaviour is highly associated with the prediction of a lower fundamental frequency because the deleterious effects of tobacco on vocal fold histology and voice quality has been frequently mentioned (Sataloff, Spiegel & Hawkshaw, 1995; Stemple, Glaze & Gerdeman, 1995; Colton & Casper, 1996; Dworkin & Meleca, 1997) although statistically significant differences between the voice quality of nonsmokers and smokers have been difficult to find (Gilbert & Weismer, 1974; Rodrigues, 1998). Nevertheless, the found evidence is that the mean F0 of the smokers is lower than that of the non-smokers (Gilbert & Weismer, 1974; Sorensen & Horii, 1982; Rodrigues, 1998), which is in agreement with the findings in this study.

Other health related variables that are associated with a lower mean F0 are related to nasal cavity alterations and upper respiratory problems as aetiology of dysphonia and sligh dysphonia severity. In fact upper respiratory tract abnormalities cause inflammation of the mucosal, alter mucosa secretions and make the mucosa more vulnerable to injury and consequently may contribute to voice quality disturbance (Cooper, 1977; Benninger & Gardner, 1998).

Cooper's (1977) clinical experience is that an upper respiratory tract infection may serve as a precipitating factor in a voice disorder if (a) that infection is prolonged and/or severe, (b) occurs in combination with body fatigue and/or psychological depression and (c) if the subject protects his voice by lowering his pitch and likes that type of voice. However, no experimental evidence is available about the effect of upper respiratory problems on specific voice parameters.

The study findings suggest that the variables (number of coffee a day, sudden dysphonia onset, psychological factors as a cause of dysphonia, hoarse voice and stress) are associated with a higher mean F0.

When considering how caffeine may affect the voice the major possible explanation arise from the fact of it being a central nervous stimulant (Martin, 1983, 1984) and consequently affects the vocal fold vibration although experimental evidence is not available.

Psychological factors (such as emotional tension and anxiety) and stress are frequently cited as related to laryngeal tension and poor respiratory support (Morrison et al., 1994; Stemple, Glaze & Gerdeman, 1995) and also potential sources of stress related to psychological factors are linked to the onset and maintenance of dysphonia (Prater & Swift, 1984; Aronson, 1990; Haynes, Pindzola & Emerick, 1992; Jackson-Menaldi, 1992; Morrison et al., 1994; Koschkee & Rammage, 1997). Nevertheless, experimental evidence is still scarce (House & Andrews, 1988; Kinzl, Biebl & Rauchegger, 1988; Goldman et al., 1996).

Overall, the results obtained in this study indicate that the data from the history case, the subjects' self –rating stress, the ENT and the SLT diagnosis data played a major role in the prediction of the mean F0 variance for all the speech tasks.

7.2.2- F0 standard deviation

The mean F0 standard deviation in the present study for the 'controls' is lower than that published in the literature review (chapter III) but for other languages. It is quite conceivable that these discrepancies may be attributable to the differences in the type of calculation or to the sampling size used for analyses (especially for connected

speech where the discrepancies are greater). The published studies using reading (Stoicheff, 1981; Krook, 1988; Shipp et al., 1992; Russell, Penny & Pemberton, 1995; Söderstein, Hertegärd & Hammarberg, 1995) show that the time analyses vary between 9 and 35 seconds (whereas the present one used 3 minutes) or between 45 and 99 words (whereas the present one used 290 words). During conversation only Fitch (1990) reported the use of 3 minutes (as the present study) whereas the others (Hollien & Jackson, 1973; Sorensen & Horii, 1982; Britto & Doyle, 1990; Hollien, Hollien & De Jong, 1997) reported between 30 seconds and 2 minutes.

The ANOVA results in this study indicate that only age and speech and their interaction was statistically significant. Various studies indicated greater F0 variability as a function of age (Mysak, 1959; Linke, 1973; Stoicheff, 1981, Linville, 1988). According to speech results show that (1) no significant differences were found between all the vowels; (2) significant differences were found between the vowels and connected speech. Vowels show higher F0 standard deviation than connected speech (reading and conversation) and this finding seems to support the statements of several researchers in that vowels are less stable than connected speech (Askenfelt & Hammarberg, 1986; Laver, Hill & Beck, 1992; Scherer, Vail & Guo, 1995) and that variability of F0 tends to decrease as the time sample increases (Horii, 1975; Titze, Horii & Scherer, 1987; Laver, Hiller & Beck, 1992). (3) there is no significant F0 standard deviation difference between reading and conversation (except for the older control speakers). This finding is in agreement with the abovementioned finding from Horii (1975) and from other researchers (Ramig & Ringel, 1983; Snidecor, 1943, Mysak, 1959; Saxman & Burk, 1968; Holien & Jackson, 1973; Fitch, 1990; Higgins & Saxman, 1991; Hollien, Hollien & De Jong, 1997) where F0 variability between reading and speech is not large and with others (Stoicheff, 1981; Linville, 1988) where a tendency toward a greater pitch variability as a function of age was found.

The prediction of the mean F0 standard deviation variance shows that some demographic variables (age, gender-male, birth place and number of years living in Lisbon), social habit (smoking habit) and health related variables (consumption of medication, tobacco as a cause of dysphonia, sudden dysphonia onset and personal and financial stress) are associated with a higher mean F0 standard deviation whereas others like 'controls', 'alcohol consumption', ENT diagnosis (tissue larynx change) and 'financial stress' are associated with a lower higher mean F0 standard deviation.

Nevertheless it should be mention that 'financial stress' may also contribute to a lower mean F0 standard deviation (depending on the combined model variables as shown in table VI.11).

Consumed medication per day was the only significant predictor that accounts for 30% of the F0 standard deviation variance contributing for its increase during reading. Prescribed medication effects on voice has been focused by several researchers (Martin, 1983; 1984; Sataloff, 1994b; Colton & Casper, 1996) although scientific evidence is almost absent. However, considering that in the present study the majority of the reported prescribed medication was antihypertensives, antidepressants and hormonal compensation, some speculation may arise based on an increased phonation threshold pressure related to dehydration and possible vocal fold cover irregularities. Additionally, the subjects' sample in the present study not only shows a low level of daily water intake (1 litre or less) but also a considerable consumption percentage of drying agents (smoking and coffee intake). Nevertheless, it must be noted that this is merely tentative speculation and as such should be viewed with caution. However, this finding underscores the importance of having additional research to specify further the relationship mentioned.

Among the predictors for the other models some regional variables (Alentejo or Algarve or the islands and number of years living in Lisbon) contribute to a higher F0 standard deviation variance. It seems that the regional variables call attention to the importance of tongue-height on phonatory action once the vowel [a] from Alentejo or the Algarve (articulated similar to an open [ɛ]), the vowel [u] from Azores (articulated as [ü], centralized) or from Madeira (articulated as [ɑw], diphthongised). The question of articulatory environment interaction with the glottal source has been the subject of theoretical explanation (described in detail in chapter III) and experimental evidence is almost limited to sustained phonation with few references to connected speech. However, it remains to be seen if there are in fact tongue-height effects or if other factors are biasing the results (e.g. cultural or pitch contours). This question of articulatory environment requires a more in-depth investigation for Portuguese.

On the other hand nothing is known about the reason why there is a higher F0 standard deviation related to the number of years living in Lisbon. It is tempting to assume that is related to cultural or social factors or to a variety of combining unknown

factors. Further studies of Portuguese subjects with different linguistic backgrounds would be of interest.

Age and gender (male) contribute to a higher F0 standard deviation of the vowel [a] and conversation. As has already been mentioned, experimental evidence on the voice variance according to the ageing process seems reasonable but demonstrates some divergences. The contribution of gender for the conversation may be explained by the high percentage of irregularity found in this study.

Among the social habit variables, tobacco as a aetiology of dysphonia contributes to a higher F0 standard deviation on the vowel [a] and [i] whereas alcohol consumption contributes to a lower F0 standard deviation on the vowel [i] and reading. Tobacco has contributed for the mean F0 lowering and it seems obvious that it also contributes to a wider F0 standard deviation. Alcohol at high levels is considered a central nervous system (CNS) depressant causing evident speech and voice differences from that of a sober speaker (Sobell & Sobell, 1972; Martin, 1983, 1984). However, at low levels it is considered a CNS stimulant and its effects on speech and voice are controversial and experimental evidence is scarce (Sataloff, Spiegel & Hawkshaw, 1995; Dworkin & Meleca, 1997). This leads to the assumption that data (concerning this subject) in the present study is not clear enough to draw any conclusion. Meanwhile, considering the significant social role played by alcohol nowadays and the well known adverse effects of alcohol on larynx health, especially when added to smoking habits (Aronson, 1990; Hirano & Bless, 1993; Colton & Casper, 1996; Dworkin & Meleca, 1997) further investigation in this area would be important.

Overall, the results obtained in this study indicate that the data from the history case, the speakers' self-rating of stress and the ENT diagnosis data played a major role in the prediction of the F0 standard deviation variance for all the speech tasks.

7.2.3- F0 range

Data for male 'controls' (0.75 octaves) is within the limits of the published literature during reading, 0.79 octaves (Mysak, 1959) and 0.66 octaves (Horii, 1975); however, during speech the 90% F0 range is 1.18 octaves and that value is higher than the 0.78 octaves published by Mysak (1959). 90% F0 range for females was only found

during speech with university students (Linke, 1973), which is lower (0.77 octaves) than the value found in the present study (1.12 octaves). Possibly, the differences found between the abovementioned studies are biased not only by linguistic confounding relationships but also by the characteristics of the sample or other unknown variables.

ANOVA results reveal a significant main effect of speech tasks showing that conversation has a wider F0 range than that of reading. This outcome was expected considering that conversation is a less restrictive task in nature than reading in that it potentially allows the speaker to be more expressive in a more relaxed and informal situation (Baken & Orlikoff, 2000). Second, young controls show a significantly narrower range than older ones during reading and conversation. Third, young dysphonics show a significantly narrower range than the olders during conversation. This findings allow speculation that the restricted F0 range in young subjects may be attributable to psychological attitude in a 'non natural' situation.

Moderate (variance explained between 20% and 29%) associations were brought out for the 90% F0 range during reading and conversation. The demographic variable 'age', health related (menstruation, psychological factors as a cause of dysphonia), social habit (number of cigarettes), ENT diagnosis (larynx tissue change and minor nose abnormalities), and STL diagnosis (harsh voice) are associated with a wider 90% F0 range whereas the social habit (alcohol consumption and coffee intake) and the STL diagnosis (breathy voice) are associated with a more restricted 90% F0 range.

On the basis of the experimental evidence (Abitbol et al., 1995; Sataloff, Spiegel & Hawshaw, 1995; Chernobelsky, 1998) a slight hoarseness and larynx hypotension during the menstrual cycle was found so, it is conceivable that a wide F0 range during reading results from an unconscious larynx effort to compensate for the voice quality to obtain a better reading performance. However, the hypothesis being true and considering that those voices are on the boundary of the 'normal-disordered continuum' the maintenance of this effort may result in voice fatigue, a common reported symptom during the menstrual phase (Sataloff, Spiegel & Hawshaw, 1995; Chernobelsky, 1998). This is an interesting issue but it was not the specific aim of this study and more detailed data is needed in this regard.

Overall, the results obtained in this study indicate that the history case data, the ENT and the SLT diagnosis data played a major role in the prediction of the F0 range variance for all the speech tasks.

7.2.4 – Percentage of jitter

All the percentage jitter values (with an exception for the vowel [a] in the female speakers) gathered in this study for the 'controls' are within the typical values (between 0.5 and 1.0%) accepted by the researchers for 'normal' sustained phonations (Hollien, Michel & Doherty, 1973; Heiberger & Horii, 1982; Horii, 1979, 1982; Titze, Horii & Scherer, 1987; Deem et al., 1989) and in agreement with the range of jitter values published in the literature for other languages for the females (Linville, 1988; Deem et al., 1989; 1991) and for the males (Horii, 1982; Deem et al., 1989; 1991; Hertrich & Ackerman, 1995). Neverthless, two exceptions were observed. Firstly, the vowel [a] for the female speakers in the present study shows slightly higher percentage jitter (mean=1.28) than those published in the literature (between 0.38 and 1.15). In fact the mean value for the Portuguese [a] is approximately the same as the published for German (mean=1.15) by Hertrich & Ackerman (1995). Secondly, the vowel [u] for the male speakers in the present study shows slightly higher percentage (mean=0.81) that the published values (between 0.29 and 0.66).

In the present study jitter measures were not sensitive to the main effects of group, age, gender and speech. Notwithstanding that consensus about jitter measurements is still not achieved this finding is in line with the findings from Peppard, Bless & Milenkovic (1988) and Haji et al. (1986) where no statistically significant differences were found between 'dysphonics' and 'normals' but is not in agreement with the studies of Iwata & von Leden (1970), Hertrich & Ackerman (1995), Orlikoff et al. (1997) and Yiu et al. (2000) who found statistically significant differences between 'dysphonics' and 'normals'.

The results (non significant difference) may suggest that there was not enough information in the number of analysed samples in this study although the criteria used was the maximum number of samples from the steady sound. Also the duration of segments was not statistically different between the dysphonics and 'normal' subjects as described in the methodology chapter. Furthermore, the lack of significance may be due to the fact that approximately two thirds of the voice samples used are slight to

moderate dysphonics. There is another possible explanation, related to the collecting methods and extracting procedure differences between this study and the early ones (see chapter III section 3.6). Neverthless, it should be noted that in 20 roughness voices (from a large database of 389 pathological voices) Omori et al. (1997) found that 5 where within the normal jitter range.

Although the results failed to indicate any statistically significant differences some trends were noted. Dysphonic, male and older speakers present slightly higher jitter values during the sustained vowels than that of the 'normal', females (except in the vowel [a] of the females that is higher than that of the male speakers) and younger subjects. Overall, these findings corroborate some studies cited in the literature review wherein higher jitter magnitudes were found related to dysphonics (Peppard, Bless & Milenkovic, 1988; Rantala, Maatta & Vilkman, 1997), males (Titze, Horii & Scherer, 1987; Orlikoff & Baken, 1990; Colton & Casper, 1996) and ageing (Hollien, Michel & Doherty, 1973; Murray & Doherty, 1980; Ramig & Ringel, 1983; Biever & Bless, 1989; Colton & Casper, 1996; Baken & Orlikoff, 2000) although those differences were not significant.

Examination of the jitter values also indicate that vowel [i] has the highest value for the dysphonics, males and adults subjects whereas the vowel [a] has the highest value for the 'normal', females and older adults. However, an issue of potential bias (interference with F0) should be noted.

As in a high percentage of studies (Horii, 1975; Koike, Takahashi & Calcaterra, 1977; Linville, 1988; Peppard, Bless & Milenkovic, 1988, Kotby et al., 1993; Sussman & Sapienza, 1994; Orlikoff, 1991; Orlikoff et al., 1997; Chae et al., 2000) the subjects in the present study were instructed to phonate at a 'comfortable' pitch and loudness. However, 'confortable' level for both frequency and intensity can vary over time (Gelfer, 1995) and habitual conversational pitch elevation is observed after vocal excessive loading (Gelfer, Andrew & Schmidt, 1991; Stemple, Glaze & Lee, 1995). In the current study according to methodological sequence (stroboscopy examination followed by the questionnaire case history interview) the speakers produced the sustained vowels after more than one hour of clinical evaluation.

Also, several researchers acknowledge a relationship between F0 level and jitter (Horii, 1979; Heiberg & Horii, 1982; Hollien, Michel & Doherty, 1973) although there is not a straightforward consensus on whether jitter increases or decreases as the F0 level

increases. In fact examination of the present study data reveals that males and young adults show the highest jitter value in the highest F0 level vowel ([i]) and this is in line with Hollien, Michel & Doherty (1973) findings but in contrast with Stemple, Stanley & Lee's (1995) findings whereas the 'normal', females and older subjects show their highest jitter value in the lowest F0 vowel ([a]) and this is in agreement with Horii's (1982) findings. This results suggest that the present information is not enough and robust and more research is needed in this regard.

Results of the regression analysis performed produced combining models (of three or four variables) for the three sustained vowels in the study that account for a moderate prediction (between 21% and 23%) of the jitter variance.

The demographic (born in the Alentejo or in the Algarve or in the North and number of years living in Lisbon) as well as the health related variables (ovulation phase and larynx functional disorder) contribute to a higher jitter percentage whereas the demographic (number of years of education) and the health related variables ('control', minor larynx variation, slight dysphonia severity and financial stress) contribute for a lower jitter.

Once again as in the F0 standard deviation the results suggest a correlation between tongue-height variance and increasing jitter variance as referred to in the previous literature (Horii, 1979; Linville & Korabic, 1987; Deam et al., 1989; Sussman & Sapienza, 1994).

Overall, the results obtained in this study indicate that the history case data, the speakers' self-rating stress, the ENT and the SLT diagnosis data played a major role in the prediction of the jitter variance for all the speech tasks.

7.2.5- Percentage of irregularity

It is difficult to establish if the percentage of irregularity data obtained in this study for the 'controls' is within the limits of the published data because this measure is only available for the Speech Studio software (Laryngograph Ltd.) and only a few references were found (Carlson, 1995a & b; Bowen, 1997). Neverthless, the results of the present study are in line with the Carlson's (1995a) statement just for females during

reading (mean=7.6%) once she considers that the values should be within 5 and 10% for 'normal' speakers and in disagreement for reading in the male speakers (mean=12.5%) and conversation in both genders (12.3% for females and 23.5% for males). Moreover in comparison to Bowen (1997) (5.83 for males and 5.60 for females during reading) the present study presents higher values. However, data from reading (both genders) and conversation (female speakers only) in the present study is in agreement with data presented by Carlson (1995a) for 'normal' speakers during reading (mean=13.9%) and conversation (18.4%). Moreover, Carlson (1995a) notes that within 'normal' speakers the percentage of irregularity ranges between 4 and 6% and between 40 and 60% were found (showing the highly intersubject variability). Discrepancies may be possible related to methodology differences between the studies, for example, first, the increased levels of irregularity in the present study may be related to the subjects' body constitution which possibly interfered in the ELG data gathering. In this study the mean body mass index for the control subjects is 23.66 (ranging from 17.2 to 36.0) meaning that subjects vary between the ideal weight and overweight. Second, the increased levels of irregularity in the present study may be related to the effect of larynx-vocal tract interaction for voiced fricatives. Third, the higher levels of irregularity could be indicative of higher percentages of pervasive habits (e.g. smoking and alcoholic habits) or voice use (eg. using voice abuses such as hard glottal attacks or others and being anaware of their negative effects) low hydration habits (e.g. controls reported a mean of 0.8 litres a day) or associated to diseases (e.g. 12.3% of controls reported gastroesophageal disease and 52.6% minor nose abnormalities) other unknown factors.

Statistical ANOVA testing confirmed that the factors of gender and speech task did exert a main effect on increased irregularity as well as the interactions between the factors speech, group, gender and age. This confirms the existence of higher percentages of irregularity in the dysphonics and males than in the 'normal' and female subjects for both speech tasks (reading and conversation) and older females during reading.

No normative data is yet available for the percentage of irregularity although it appears obvious that dysphonics would present higher levels due to the lack of vocal fold integrity and phonation biomechanics and the multiple components of health status (e.g. individual body integrity, emotional state and voice use). In fact Carlson (1995a)

found higher percentage of irregularity in laryngeal cancer patients than in 'normals' during reading and conversation.

Although the older subjects present higher percentages of irregularity than the young ones results showed that this is only significant for the females during reading. The possible explanation for this trend may be related to the limited older males sample size as have already been mentioned.

A combination of four and six variables produced a significant set of predictors accounting for a moderate prediction (between 31% and 39%) variance of the percentage of irregularity during reading and conversation. Within the models the demographic variables (male, born in the North or in the islands) and health related variables (tobacco as a aetiology of dysphonia and hoarse voice) are associated to a higher percentage of irregularity during reading and conversation. However, the variables 'the number of coffees a day' and 'total stress' are associated with a lower percentage of irregularity during conversation.

Overall, the results obtained in this study indicate that the history case data, the speakers' self-rating stress and the SLT diagnosis data played a major role in the prediction of the percentage of irregularity variance for reading and conversation.

7.2.6-Mean contact closed quotient (QFx)

The QFx results for the 'normal' subjects were similar to those reported by Rossiter, Howard & Comins (1995) and within the limits stated by Orlikoff (1998) which are 40% to 60% suggesting that the 'normal' subjects display relatively controlled vocal fold behaviour for all the speech tasks as measured by ELG. Also in agreement with the abovementioned researchers no significant difference according to gender was found in the 'normal' subjects.

It is also interesting to note that our findings seem to reflect the evidence cited by Rossiter, Howard & Comins (1995) and Verdolini et al. (1998) in which non-vocally trained subjects will exhibit lower contact quotient than that of long term vocally trained subjects. Within the 'normal'subjects sample there were only 6 subjects with a voice hobby (singing in a choir on a non systematic basis for at least 7 years) and one singer

student for at least three years and the majority of the subjects had non significant vocal training. Neverthless, this is only speculative observation as this was not the aim of the present study and deserves further investigation.

ANOVA results indicate that the contact closed quotient in this study is not sensitive to the main effects of group, gender and age. Although experimental evidence in this area is still insufficient the findings in this study concurs with previous research in which no gender effect was observed (Bowen, 1997; Orlikoff, 1998).

The only main effect on the mean QFx found in this study was speech task. Post hoc comparisons revealed that speech effect for the dysphonics is restricted to a significantly higher mean QFx for the vowel [i] than that of the vowel [a], reading and conversation whereas for the 'normal' subjects is restricted to a significantly higher mean QFx for the vowel [i] than that of conversation and a significantly higher mean QFx for the vowel [u] than that of reading and conversation. These findings raise the question as to the reasons for the values which may be related either to a higher intrinsic F0 in those vowels or to vowel specific variation in loudness. Orlikoff (1991) states that the magnitude of the contact quotient increases with increasing vocal intensity.

Multiple regression analysis indicate significant combining models relating different variables that account for a low prediction variance (between 7% and 16%) for the mean QFx during the sustained vowel [a] and reading and a moderate prediction variance (between 24% and 29%) for the mean contact closed quotient for the other speech tasks. Among all the predictors that account for a higher QFx variance index are demographic (number of years living in Lisbon), social habits (number of coffees per day) and health related variables (normal voice, harsh voice, larynx mass lesion and work stress).

It appears obvious that harsh voice, larynx mass lesion and work stress be related to a relatively increment of the vocal folds contact due to the compensatory increased larynx adductory effort applied by the subjects either as a result of activity or to compensate for the reduced integrity of the mucosal vocal fold contact area (e.g. mass lesion). Whether the number of coffees acts as a CNS stimulant or as dehydration agent when in combination with other variables is not known. Additionally, 'normal' voice was within the expectations of higher contact closed quotients than the other situations.

Among all the predictors that account for a final lower QFx variance index are demographic (economic activity), social habits (alcohol consumption) and health related variables (distributed tissue larynx change lesion and gastroesophageal disease).

There is a substantial body of research (Stemple, Stanley & Lee, 1995; Buekers et al., 1995; Sapir, Mathers-Schmidt & Larson, 1996; Rantala. Maata & Vilkman, 1997) confirming the effects of occupation on voice quality indicating signs of voice fatigue and hoarseness and also poorer vocal fold contact efficiency after loud talking (Buekers, 1998a).

Whether alcohol acts as a CNS stimulant or depressant or any other unknown effect due to the combination with the other variables is not known. As has already been mentioned this is an area that deserves further investigation.

Gastroesophageal disease has been identified to either be causative or a cofactor in the larynx mucosal changes (e.g. acute laryngitis and Reinke's oedema) and various researchers (Woo, 1996; Colton & Casper, 1996; Dworkin & Meleca, 1997; Benninger & Gardner, 1998; Ross, Noordzji & Woo, 1998) indicate incomplete and irregular vocal fold closure in laryngitis which seems to be in the line of the findings of this study.

Overall, the results obtained in this study indicate that the history case data, the speakers' self-rating of stress, the ENT and the SLT diagnosis data played a major role in the prediction of the mean contact quotient variance for all the speech tasks.

7.2.7- QFx standard deviation

The findings of this study revealed that, firstly, reading and conversation present significant higher contact closed quotient standard deviation than that for all the vowels and secondly, conversation shows higher QFx standard deviation than reading for the dysphonic subjects. Although not significant 'normal' subjects present the same trend as the dysphonics.

Given dysphonics/normal, males/females and young/old subjects did not differ significantly in QFx standard deviation does not discount intergroups variability in QFx standard deviation noted across the various speech tasks. Dysphonics, males and older subjects yielded higher QF standard deviation for all speech tasks.

The results indicate that the regression analysis brought out combining models for almost all the speech tasks except for the vowel [i]. Low (variance explained at most 8%) and moderate (variance explained between 23% and 41%) correlations were brought out for the QFx standard deviation. The demographic (age, born in the center or islands), social habits (number of smoked cigarettes a day) and health related variables (tobacco and work as an aetiology of dysphonia, harsh voice, mass lesion pathology and personal stress) contribute to a higher final QFx standard deviation variance. On the basis of the male-female coalescence model (Krook, 1988; Yamazawa and Hollien, 1992) it seems obvious that QFx standard deviation increases not only with chronological age but also with the other abovementioned variables.

Moreover, a lower final QFx standard deviation index results from the combined effect of the demographic (years of education), social habits (number of coffees a day) and health related variables ('controls', slight dysphonia severity and total stress).

Overall, the results obtained in this study indicate that the history case data, the speakers' self-rating stress, the ENT and the SLT diagnosis data played a major role in the prediction of the contact quotient standard deviation variance for all the speech tasks.

7.3 – Psychosocial voice impact

It is not possible to establish if the VHI data obtained in this study is within the limits of the published data because there is only a modest number of studies using this tool and either data is not mentioned (Benninger et al., 1998; Rosen et al., 2000) or it is for other larynx pathologies (Stewart, Chen & Stach, 1998) or singers (Rosen & Murry, 2000) both absent from the present study. Moreover, until now there is no data for 'normal' speakers. Nevertheless, within the limitations comparison between the mean

group data in this study and those obtained in Jacobson et al. (1997) and Rosen & Murry (2000) will be exposed.

Jacobson et al. (1997) shows higher VHI scores for the self-rated voice severity categories than those obtained in the present study in accordance to the SLT severity categories. Also, mean group data for the nonsinger dysphonics in Rosen & Murry (2000) is higher than the data of dysphonics from the present research. These discrepancies may be attributable to unknown factors (eg. cultural) or methodological differences between the studies. In fact in Jacobson et al.'s (1997) study, 26% of the patients are laryngectomees which it seems obviously a more serious voice disorder than the voice disorders studied in this research. Also, Rosen & Murry (2000) in their nonsinger group considered an age range between 6 and 79 years and because voice is aged related results may be biased by voice characteristics common to different age stages (infance, puberty and elderly).

Comparison between the groups showed that dysphonic subjects present statistical significant more marked psychosocial voice impact than the 'normal' subjects. This finding shows that the questionnaire is sensitive to dysphonia and the results, although the methodological differences are in line with the previous research (Smith et al., 1996; Verdonck-de Leeuw, 1998; Hogikyan et al., 2000).

No significant effect of age or gender was found although females and older subjects showed higher psychosocial voice impact scores. The non-significant effect of age is in line with Stewart, Chen & Stach's (1998) and Rosen & Murry's(2000) (both used the VHI) but in contradition with Smith et al's (1996) findings. These discrepancies may be related to bias in the samples related to employment status, living alone or other unknown factors.

Within the sub-scales (functional, emotional and physical) there was no statistically significant difference between the emotional and functional sub-scores for the dysphonic subjects. This finding allows speculation about the duration of the voice symptoms (chronicity) of the speakers in the present research (more or less 44% of the subjects presented dysphonia between one and five years and 35% for more than 10 years) and the possible coping strategies used by them and the acceptability of the disorder. Smith et al. (1996) write that some people seem to tolerate their vocal symptoms for a considerable time.

Multiple regression analyses applied to VHI data indicate that the variable 'controls' contributes to predicting lower total and physical VHI scores which

corroborate the data comparisons already mentioned and the published data (Smith et al., 1996; Verdonck-de Leeuw, 1998; Hogikyan et al., 2000).

Moreover, the acoustic measures F0 standard deviation and jitter (both in the vowel [i]) made a contribution to predicting the VHI emotional and physical scores (contributing to worst voice impact). Overall, these results are in line with those from Wuyts et al. (2000) who found a high correlation between VHI and objective measures of vocal quality. Unfortunately as this is the only published article on this subject and specific values are not presented it is not possible to establish in depth associations.

Also, the variable 'disease as a cause of dysphonia' is significantly associated with a lower total VHI score (consequently a lower total voice impact). For the purpose of this study the label of 'disease as a cause of dysphonia' was considered if the dysphonia onset was related to an upper respiratory infection (eg. cold, allergies) and associated with continuing vocal use (as might be seen in chapter V). These results seem to confirm the clinical research observation that many people do not recognize the relevance of the early symptoms of chronic or acute conditions such as upper respiratory infection, allergies (Morrison et al., 1994; Freeman, 2000).

Moreover, 'gradual dysphonia onset' is significantly associated with a higher emotional and functional VHI score (worse emotional and functional voice impact) whereas 'in employment' is associated to a lower emotional VHI score (contributing to a lighter emotional impact). Freeman (2000) says that people can react very differently to different situations and it can be speculated if 'in employment' is an effective coping strategy for emotional voice impact.

7.4 – Strengths and limitations encountered in the study

This study represents the first attempt to determine the nature of the dysphonic Portuguese speakers involving a large sample of subjects, and a comparison 'normal' group of subjects. It uses a standardized multidimensional evaluation protocol (combining qualitative and quantitative measures) from the professionals' and speakers' perspectives and multivariate statistical analyses. Useful information about Portuguese voice has been provided by the investigation which would otherwise have been unobtainable.

However, care should be taken in the interpretation and generalization of the findings considering the exploratory nature of this study and the potential sample and methodological limitations. Strengths and limitations encountered in the study are outlined below.

Among the published literature, studies that used large samples of subjects are scarce (Hollien & Shipp, 1972; Horii, 1979; Hudson & Holbrook, 1981, Krook, 1988; Hollien, Hollien & De Jong, 1997) especially involving dysphonic subjects (Kotby et al., 1993; Verdonck-de-Leeuw, 1998; Wuyts et al., 2000; Wolfe, Martin & Palmer, 2000) and this fact limits the possibility of gathering a representative pattern of voice quality due to the large inter-speaker variability.

Although strict subject selection criteria were previously defined it was not possible to control for all potential variables so there is the problem of selection bias. Subjects differ in age (only for the male subjects) social habits (smoke, water and coffee consumption) and health variables (gastroesophageal disease and level of stress). Neverthless, although the statistical measures applied reduce the intrinsic variability, the interpretation and generalization of the findings must be carefully drawn.

It was not feasible to gather a wide range of larynx pathologies, as patients in all conditions were not available at the time of data gathering. Pathologies were collapsed into groups roughly approximating the categories (defined in the methodology) therefore, it is difficult to extract meaningful information about the variability exhibited by each specific laryngeal pathology. According to Haji et al. (1986) attempts to discriminate between dysphonic and 'normal' speakers must take into consideration the severity of the voice problem and there is no question that differentiation between extremely dysphonic voices and 'normal' voices will be remarkably good. On the other hand if the group includes mainly slight and moderate dysphonia (which is the case of the present study, approximately 65% of the speakers present slight to moderate dysphonia), differentiation will be rather poor.

A feature worth noting is that the proportion of females (75%) higher than males (25%) for both dysphonic and control speakers is similar to other researchers' findings (Herrington-Hall et al., 1988; Arnoux-Sindt et al., 1994; Fritzell, 1996; Ferreira et al., 1996; Guimarães, 1997; Carvalho et al., 1995).

Moreover, in a high percentage of the published studies 'normal' voice classification is based on a perceptual voice evaluation and/or lack of voice complaints and 'normal' speakers are not submitted to laryngostroboscopic examination (Linville, 1988) whereas in this study all the 'normal' as well as the 'dysphonic' subjects were assessed by stroboscopy and perceptual voice evaluation. Although the criteria followed in the present study may be, in a sense, considered a strength of the research design for once stroboscopic examination proved to be essential for the differential diagnosis between 'normal', 'minor abnormalities' and 'pathology' (Pemberton et al., 1993) in a practical sense it became a limitation in gathering 'normal' subjects (especially men). However, issues of potential bias must be considered. Firstly, no quantitative F0 and intensity measurement control of the subjects' voice production was performed although they were instructed to produce at a comfortable conversational pitch and loudness level and this is the common procedure in the published studies (Casper, Brewer & Colton, 1987; Colton et al., 1995; Södersten, Hertegård & Hammarberg, 1995; Elias et al., 1997). Nevertheless as reported in the methodology section the fiberscope was connected to a stroboscope with the fundamental frequency (in Hz) displayed on the videomonitor. Secondly, a great number of subjects complained about discomfort during the stroboscopy due to fiberscope and the topical anaesthesia. Although Peppard & Bless (1991 cited by Stemple, Glaze & Gerdeman, 1995) demonstrated that topical anaesthesia did not affect the subjects' phonatory ability it cannot be confirmed that it did not have an effect in our study. Nevertheless, it is possible that due to the type of examination used in this study anaesthesia contributed to a lower degree of discomfort (Södersten & Lindestad, 1990) and to a better control of the gag reflex.

Although voice quality measurement is nowadays, a common procedure, still data on connected speech presented in research literature shows that it is rather meager compared to data on sustained vowel phonation as well as studies using the two different voice uses (sustained phonation and connected speech). The strengths and limitations of the use of a unique form of voice production in the voice study (sustained or connected phonation) have long been the subject of debate in the voice science community (presented in detail in chapter III). In order to obtain more useful quantitative voice measures this study used two levels of data: (1)'micro' level, short duration samples of sustained phonation and (2) 'macro' level, two to three minutes samples of connected speech (reading and conversation).

Furthermore, nowadays studies that use combined forms of voice evaluation from the professionals' and subjects' perspectives as the present study are still scarce (Carlson, 1995a; Bough et al., 1996; Verdonck-de Leeuw, 1998; Buekers, 1998a; Wuyts et al., 2000). This study also attempted to fill some of the gaps in the research literature as not only using the traditional tools of voice evaluation in the clinical setting but also by identifying the speakers' self-rating of stress and psychosocial voice impact.

Electrolaryngography was the instrumental technique chosen for voice measurement since the waveform obtained is unaffected by the acoustic resonances of the vocal tract and more suitable for the voice signal perturbation analysis (Haji et al., 1986; Bough et al., 1996) and it has also the advantages of not being invasive and immune to noise (a special advantage in the clinical setting where data was collected, as explained in the methodology section). One of the major constraints of this technique is the potentially high percentage of unreliable ELG signals (that happened in the pilot study but not in the main study, 96.5% of all subjects recruited whose data could be used).

Subjective forms of evaluation on the professional perspectives were also used (e.g. stroboscopy and perceptual voice evaluation) but potential bias may be found. Firstly, this study used two experienced ENT surgeons' examinations but intra- and interjudge reliability was not determined. Nevertheless, a unique rating protocol was used and was previously tested in a pilot study. Moreover, all the examinations were performed in the presence of the researcher and were recorded on videotape. This procedure was similar to Colton et al. (1995) study where four experienced specialists were used as observers and judgments were based as well in the agreement between them and no formal reliability was determined, and to Verdonck-de Leeuw (1998) where the patients stroboscopic examinations '...were performed by or under the supervision of an experienced phoniatrician.' Secondly, only one rater for the perceptual voice analyses was used like other studies (Silverman & Zimmer, 1978; Linville, 1988; Gelfer & Bultemeyer, 1990; Fitch, 1990; Pemberton et al., 1993). Nevertheless, several important criteria focused on in the previous literature were taken into consideration such as rater experience in perceptual evaluation, precise and defined terminology scales and use of running speech sample (Buekers, 1998b; Fex, 1992; Orlikoff et al., 1999; Carding et al., 2000). Moreover high intra-rater judge reliability and correlation

with the ELG measurements were observed. According to Kreiman et al. (1993), when high intra-reliability is observed, the rater is considered reliable. However, it would be important to demonstrate the generalisation of these findings for other comparable trained raters.

The speakers' self-rating scales (level of stress measured with SRRS and psychosocial voice impact measured by VHI) showed a significant difference between the groups although within each group a large variability of scores were found. Several factors may be either related to the subjects or to the sensitivity of the scales for this type of subjects. However, the use of a combination of measures (multiple regression) in this study proved the importance of the level of stress and psychosocial voice impact when in combination with other variables to the final voice parameters variance. As acknowledged previously this results offer preliminary information. Instead of examining only the events per se, future studies if time, sequence and duration between the life events occurrence and the dysphonia onset and severity are also considered, would certainly give more important insights in this area. However, there is the problem of the formal validity for both instruments (SRRS and VHI). Unlike most validation studies of new questionnaires translation was made according to the published literature and according to the specific requirements for VQoL tools (www.qelmed.org). Moreover, they have been previously tested in a pilot study (described in detail in appendix A) and refined by reference to a group of adult speakers without voice complaints and also used in an adequate size for statistical purposes. Within the published articles using the VHI tools (Jacobson et al., 1997; Benninger et al., 1998; Stewart, Chen & Stach, 1998; Rosen & Murry, 2000; Rosen et al., 2000; Wuyts et al., 2000) the present study and Wuyts et al. (2000) are the only ones that report VHI data in comparison to 'normal' speakers. Neverthless, full consideration of the implications of the present investigation must await the outcome of a future investigation with more refined methodology and a larger sample.

Furthermore, in addition to a traditional univariate statistical analyses (ANOVA) of the data it also used multivariate statistical analyses (multiple regression analyses) to determine which combination of variables would best predict the dependent measures. As has already been mentioned significant differences between dysphonics and 'normal' subjects are difficult to obtain by the exclusive use of a univariate approach because values are often within 'normal' limits. So in addition to that traditional approach and

considering that voice is a complex phenomenon this study used a multivariate statistical approach that enabled the prediction of a score or index for each dependent measure on the basis of combining variables models.
In contrast to the enormous amount of studies on voice using univariate analysis, there are hardly any data available about prediction variance of voice parameters on the voice study (Eskenazi, Childers & Hicks, 1990; Piccirillo et al., 1998; Verdonck-de Leeuw, 1998; Wuyts et al., 2000, Wolfe, Martin & Palmer, 2000; Parsa & Jamieson, 2000) although it is not new in health care (e.g. The TNM system for the classification of the morphologic extent of tumors, the APGAR score for the classification and description of newborns and the body mass index for the weight/height2 relationship).

In the present study the multiple regression analyses showed a moderate association between the professionals evaluations (SLT and ENT) and some instrumental measurements (mean F0, F0 standard deviation, 90% F0 range, jitter, percentage of irregularity and QFx mean and standard deviation) as well as between the speakers's self-ratings of psychosocial voice impact and stress and some instrumental measures (F0 standard deviation, jitter, percentage of irregularity and QFx).

7.5 – Suggestions for future research

This investigation presented preliminary data on Portuguese speakers; however, there is still the potential for further research on many areas not fully analysed once it was not in its aims.

The results confirm the undoubted importance of a multifactorial voice evaluation where the information gathered from the professionals and speakers' perspectives would produce a better knowledge of the voice characteristics. Regarding this it would be interesting to study in detail intra and inter-rater reliability (between professionals) for the different tools of voice evaluation and compare it to instrumental measures. The advantages would be a more detailed, standardized voice diagnosis which consequently contributes to a more effective intervention and reassessment that is still unfortunately rare in the clinic.

As the number of males and range of larynx pathologies in this investigation is reduced, it remains an area for future investigation whether the findings can be generalized to a larger group which might reveal important trends on Portuguese voice.

Furthermore, in future, it may be of interest to carry out investigations to expand the knowledge of potential variables effect on voice measures variance (e.g. alcohol quantitative consumption, hormonal cycle and stress) once specific clear and detailed aspects are taken into consideration (as have already been mentioned along the discussion), which were not in the aims of the present study. The information gained from that will be useful in voice problem prevention and therapy by clarifying misinformation and identifying risk factors.

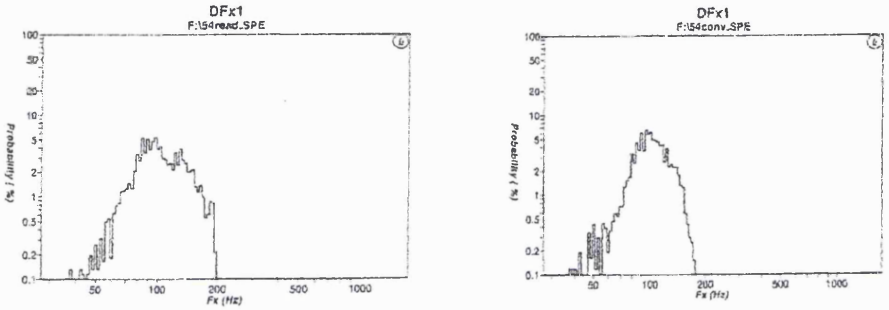
It would also be informative to carry out a further study in which individual data is analysed using quantitative and qualitative information (to clarify the potential benefits of the links that can be made between the different tools of voice evaluation) and not only group data as was the aim of the present study thereby dealing with individual variability, which would be of interest in a clinical setting (by providing a conceptual approach to the voice clinician's diagnosis and consequently intervention).

An example of the possibility of individual analysis is presented with the speaker (54) a 39-year-old woman, administrative, whom the ENT examination revealed a slight Reinke's oedema, the SLT diagnosis revealed severe hoarseness and the self-rated psychosocial voice impact (VHI, Jacobson et al., 1997) has a total score of 52 (for a possible maximum of 120).

The use of the laryngograph recordings (during 2-minute reading and 3-minute conversation) shows, in the figure VII.1 below, that there is a broad frequency distribution for both speech tasks that are often associated with mild vocal fold vibration irregularity (Fourcin, 2000). Moreover, Abberton (1976) and Fourcin (1981) found that low frequency components with a broader distribution of peaks is common in speakers with prolonged smoking habit which is the case of speaker 54 who smokes for more than 20 years, 40 cigarettes a day.

The main part of the DFx plot distribution indicates a useful voice range which seems in agreement with the speaker' self-rated psychosocial voice impact in the functional sentences 'F1-my voice makes difficult for people to hear me' and 'F5-my family has difficulty hearing me when I call them throughout the house' rated with 'almost never'

and the sentence 'F3-people have difficulty understanding me in a noisy room' rated with 'sometimes'. Although the overall shape of distribution is large the most proeminent peak (modal frequency) is around 100 Hz for both speech tasks (as might be seen in figure VII.1) which is a low pitch for a female but it is in line with the type of larynx pathology (Reinke's oedema) (Le Huche & Allali, 1990; Colton & Casper, 1996; Dworkin & Meleca, 1997) and the effects of prolonged smoking habit (Gilbert & Weismer, 1974; Stoicheff, 1981; Sorensen & Horii, 1982; Rodrigues, 1998) where researchers usually found a lower F0.



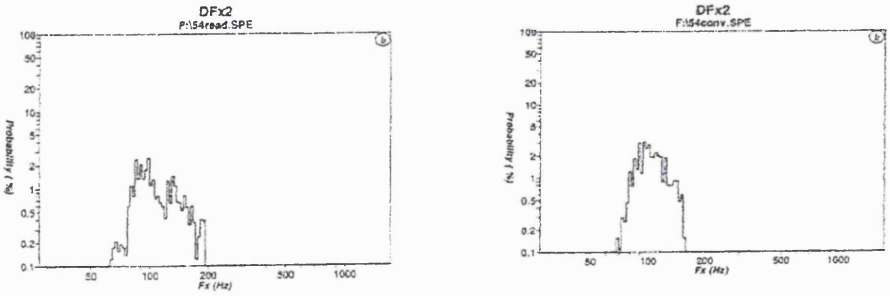
During reading

During conversation

Figure VII.1 – DFx1 distributions for the speaker 54

(from Speech Studio, Laryngograph Ltd.)

DFx1 plots (figure VII.1) comparison with DFx2 plots (figure VII.2) shows that these last ones are reduced in size and in the amount of frequencies covered, highlighting the fact that there is irregularity in the DFx1 that has been discarded on the DFx2. This fact is in line with Fourcin's (1981, 2000) comment in that 'pathological voice' does not show similar shape and size DFx1 and DFx2 distributions.



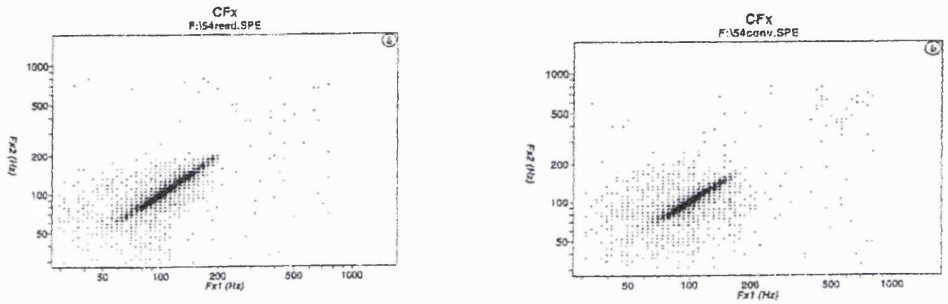
During reading

During conversation

Figure VII.2 – DFx2 distributions for the speaker 54

(from Speech Studio, Laryngograph Ltd.)

The CFx plots (figure VII.3) cover a wide range of frequencies (seen by the length of the diagonal) and shows a strong black diagonal line, but has a great deal of scatter away from the diagonal confirming the existence of limited periodicity (irregularity) of the vocal fold vibration (Carlson, 1993). Comparison with the self-rated psychosocial voice impact shows that the patient rated the physical sentence 'P13-my voice sounds creaky and dry' with 'almost always'.



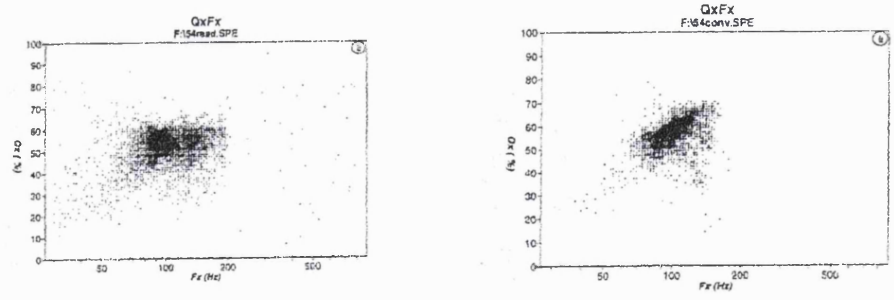
During reading

During conversation

Figure VII.3 – CFx distributions for the speaker 54

(from Speech Studio, Laryngograph Ltd.)

The QxFx plots (figure VII.4) show scatter throughout the plot showing a limited control over the closed phase of the cycle confirming the irregularities already discussed about the DFx and CFx plots. The points are concentrated in the area of 90-100Hz.



During reading

During conversation

Figure VII.4 – QxFx distributions for the speaker 54

(from Speech Studio, Laryngograph Ltd.)

During reading some low frequencies are produced with irregularly high percentage closure quotient whereas some with an irregularly low low closure quotient. Therefore, voice may be perceived as sometimes strained and sometimes breathy. During

conversation there is an increasing contact phase ratios with increase in vocal fold frequency. This laryngographic evaluation is in line with the SLT diagnosis of hoarseness and in agreement with the speaker's self-rated psychosocial voice impact once she rated as 'sometimes' that 'P20- I use a great deal of effort to speak' and 'almost always' that 'P14- I feel as though I have to strain to produce voice' and 'P17- the clarity of my voice is unpredictable'.

Nevertheless, although descriptive analysis of F0 measures (e.g. DFx, CFx and QxFx distributions) have been frequently mentioned as advantageous for clinical purposes where there is typically one rater only (Wechsler, 1977; Fourcin, 1981, 2000; Abberton & Fourcin, 1984, 1997; Carlson, 1993, 1995b) it has been shown to have poor correlation from different observers (Wechsler, 1975), therefore intra and inter-rater reliability studies with this type of evaluation would be important.

7.6 - Conclusions

This research represents a preliminary attempt at determining aspects of the nature of the Portuguese voice and although the results of this study should not be hastily generalized to other populations or other methodological conditions, overall they are consistent with what other researchers have found. Thus, there is converging evidence that:

- data for 'normal' speakers in almost all voice dependent variables in the study (mean F0, 90% F0 range, % jitter, % irregularity and mean QFx) are within the limits of the published 'normative' data for other languages. However, F0 standard deviation outcome in the present study is lower than the published literature. No reports of QFx standard deviation and the speakers' self-rating of psychosocial voice impact (VHI) data for 'normal' speakers were found in the literature;

- no one of the voice dependent measures reached statistical significance to distinguish 'controls' from 'dysphonics'; however, when interacting with other variables some statistically significant differences were found in the F0 standard deviation and the percentage of irregularity (with speech) and in the 90% F0 range (with age). The main effect of group ('controls' versus dysphonics) reached statistical significance in the self-rating of psychosocial voice impact (VHI) data. Neverthless,

data obtained in the present study showed similar trends to those published in the literature (as shown in the tables between VI. 41 and 45 in the results chapter);

- outcomes of predictive analyses (that varied between low, moderate and high) suggested important associations between the voice dependent measures and the speakers' self-rating of psychosocial voice impact, the voice data questionnaire, the speakers' self-rating stress, the SLT and the ENT diagnosis.

Chapter VIII - References

Abberton, E.R.M. (1976). *A laryngographic study of voice quality.* Unpublished PhD Thesis, University of London.

Abberton, E. Fourcin, A.J.; Rosen, S.; Walliker, J.R.; Howard, D.M.; Moore, B.C.J.; Douek, E.E. & Frampton, S. (1985). Speech perceptual and productive rehabilitation in electro-cochlear stimulation. In R. A. Schindler & M. M. Merzenich (Eds), *Cochlear Implants* (pp. 527-537). New York: Raven Press.

Abberton, E.; Fourcin, A.J. & Hazan, V. (1986). Phonological competence with profound hearing loss. *Speech, Hearing and Language: Work in progress*, UCL, 2, 1-14.

Abberton, E.R.M.; Howard, D.M. & Fourcin, A.J. (1989). Laryngographic assessment of normal voice: A tutorial. *Clinical Linguistics and Phonetics,* 3, 281-296.

Abberton, E. & Fourcin, A. (1997). Electrolaryngography. In M. J. Ball & C. Code (Eds.), *Instrumental clinical phonetics* (pp. 119-148). London: Whurr Publishers, Ltda.

Abberton, E.; Fourcin, A.; Hu, X.; Bootle, C.; Miller, D. (1999). Speech pattern element analysis and display – 'Lx Speech Studio'. V. Hazan & M. Holland (Eds), *Proceedings of ESCA/Socrates Tutorial and research workshop on Method and Tool innovations for Speech Science Education (Matisse)* (pp.77-80). London: University College London.

Abberton, E.; Fourcin, A. (2001). Vocal fold contact ratios and voice quality in the connected speech of young women. *PEVOC IV conference*, Stockholm.

Abercombie, D. (1967). *Elements of general phonetics.* Edinburgh: U.P. Edinburgh.

Abitbol, J.; Brux, J.; Millot, G.; Masson, M.F.; Mimoun, O.L.; Pau, H. & Abitbol, B. (1995). Does a hormonal vocal cord cycle exist in women? Study of vocal premenstrual syndrome in voice performers by videostroboscopy-glottography and cytology on 38 women. *Journal of Voice,* 3, 157-162.

Andrews, M.L. (1999). *Manual of voice treatment, pediatrics through geriatrics.* (2nd ed.). San Diego: Singular Publishing Group.

Arnoux-Sindt, B.; Guerrier, B.; Owhadi-Richardson, A. & Daures, J. P. (1994). Enquete sur la voix de l'enseignant sur l'academie de Montpellier 1991-1993. *Bulletin d'Audiophonologie,* 5 & 6, 519-528.

Aronson, A.E. (1990). *Clinical voice disorders: An interdisciplinary approach* (3rd ed.). New York: Thieme.

Askenfelt, A.G. & Hammarberg, B. (1980). Speech waveform perturbation analysis. *Speech Transmission Laboratory – Quarterly Progress and status report*, 4, 40-49.

Askenfelt, A.G. & Hammarberg, B. (1986). Speech perturbation analysis: a perceptual-acoustical comparison of several measures. *JSHR*, 29, 50-64.

Atkinson, J.E. (1976). Inter- and intraspeaker variability in fundamental voice frequency. *J.Acoust.Soc.Am.*, 60 (2):440-445.

Baken, R.J. & Orlikoff, R.F. (1988). Changes in vocal fundamental frequency at the segmental level: control during voiced fricatives. *JSHR*, 31:207-211.

Baken, R.J. & Orlikoff, R.F. (2000). *Voice clinical measurement.* San Diego; Singular Publishing Group.

Baken, R.J. (1992). Electroglottography (special article). *Journal of Voice*, 6(2), 98-110.

Baken, R.J. (1996). *Clinical measurement of speech and voice.* San Diego; Singular Publishing Group.

Ball, V.; Faulkner, A. & Fourcin, A. (1990). The effects of two different speech-coding strategies on voice fundamental frequency control in deafened adults. *British Journal of Audiology*, 24, 393-409.

Bassich, C.J. & Ludlow, C.L. (1986). The use of perceptual methods by new clinicians for assessing voice quality. *JSHD*, 51, 125-133.

Becker, W.; Naumann, H.H.; Pfaltz, C.R. (1994). *Ear, nose, and throat diseases* (2nd ed.). New York: Thieme.

Behrman, A.; Orlikoff, R.F. (1997). Instrumentation in voice assessment and treatment: What's the use?. *American Journal of Speech-Language Pathology*, 6(4), 9-16.

Benninger, M.S.; Gardner, G.M. (1998). Medical and surgical management in Otolaryngology (pp.497-528). In A.F. Johnson and B.H. Jacobson, *Medical Speech-Language Pathology, a pratitioner's guide.* New York: Thieme.

Benninger, M.S.; Ahuja, A.S.; Gardner, G.M. & Grywalski, C. (1998). Assessing outcomes for dysphonic patients. *Journal of voice*, 12 (4), 540-550.

Berzon, R.A. (1998). Understanding and using health-related quality of life instruments within the clinical research studies. In M. J. Staquet, R. D. Hays and P. M. Fayers (Eds.), Quality of life assessment in clinical trials. Oxford: Oxford University Press.

Bielamowicz, S.; Kreiman, Gerratt, B.R.; Dauer, M.S. & Berke (1996). Comparison of voice analysis systems for perturbation measurement. *JSHR*, 39, 126-134.

Biever, D.M. & Bless, D.M. (1989). Vibratory characteristics of the vocal folds in young adult and geriatric women. *Journal of Voice*, 3(2), 120-131.

Bless, D.M. & Baken, R.J. (1992). Assessment of voice (special article). *Journal of Voice*, 6, (2), 95-97.

Bless, D.M. et al. (1992). International Association of Logopedics and Phoniatrics (IALP) Voice committee discussion of assessment topics. (Special article). *Journal of Voice*, 6(2), 194-210.

Bless, D.M.; Glaze, L.E.; Lowery, D.B.; Campos, G. & Peppard, R.C. (1993). Stroboscopic, acoustic, aerodynamic, and perceptual analysis of voice production in normal speaking adults. In Moya Andrews (1999). *Annual of voice treatment, pediatrics through geriatrics* (pp. 17). San Diego: Singular Publishing Group.

Blomgren, M.; Chen, Y.; Ng, Manwa L. & Gilbert, H. (1998). Acoustic, aerodynamic, physiologic, and perceptual properties of modal and vocal fry registers. *J. Acoust.Soc. Am.*, 103(5), 2649-2658.

Boone, D. R. (1982). *The Boone voice program for adults.* Austin: Pro-Ed.

Borden, G.J.; Harris, K.S. & Raphael, L.J. (1994) *Speech science primer* (pp.72-89). Baltimore: Lippincott Williams & Wilkins.

Bough, D.; Heuer, R.J.; Sataloff, R.T.; Hills, J.R. & Cater, J.R. (1996). Intrasubject variability of objective measures. *Journal of Voice*, 10(2), 166-174.

Bowen, F. (1997). *The effects of ageing on the human voice: a laryngographic study.* Msc in Speech and Language Pathology, Department of Human Communication Sciences, UCL.

Britto, A.I.; Doyle, P.C. (1990). A comparison of habitual and derived optimal voice fundamental frequency values in normal voice adult speakers. *JSHD*, 55: 476-484.

Broad, D.J. (1973). Phonation. In Fred D. Minifie; Thomas J. Hixon & Frederic Williams (Eds.), *Normal aspects of speech, hearing and language* (pp.127-167). Englewood Cliffs: Prentice-Hall, Inc.

Brodnitz, F.S. (1979). Menstrual cycle and voice quality (letter to the editor). *Arch Otolaryngol.*

Brown, W.S.; Murry, T. & Hughes, D. (1976). Comfortable effort level: an experimental variable. *J. Acoust. Soc. Am.*, 60, 696-699.

Brown, W.S.; Hollien, H. (1981). Effect of menstruation on fundamental frequency of female voices. In Laurence V. (Ed.), *Transcripts of the tenth symposium care of the professional voice* (pp.94-101). New York: Voice Foundation.

Brown, W.S.; Hollien, H. (1982). Effect of menstruation on the singing voice. In Laurence V. (Ed.), *Transcripts of the eleventh symposium care of the professional voice* (pp.140-147). New York: Voice Foundation.

Brown, W.S.; Hollien, H. (1983). Effect of menstruation on the singing voice. Part I: history and current status. In Laurence V. (Ed.), *Transcripts of the twelfth symposium care of the professional voice* (pp.112-123). New York: Voice Foundation.

Brown, W.S.; Morris, R.J. & Michel, J.F. (1989). Vocal jitter in young adult and aged female voices. *Journal of Voice,* 3(2), 113-119.

Buekers, R.; Bierens, E.; Kingma, H. et al. (1995). Vocal load as measured by the voice accumulator. *Folia Phoniatrica Logopedica,* 47, 252-261.

Buekers, R. (1998a). Are voice endurance test able to assess vocal fatigue? *Clin Otolaryngol,* 23, 533-538.

Buekers, R. (1998b). Perceptual evaluation of vocal behaviour. *Log Phon Vocol,* PEVOC-II supplement, 1, 23-27.

Carding, P.N. & Horsley, I.A. (1992). An evaluation study of voice therapy in non-organic dysphonia. *European Journal of Disorders of Communication,* 27, 137-158.

Carding, P.N.; Horsley, I.A.; Docherty, G.J. (1998). The effectiveness of voice therapy for patients with non-organic dysphonia. *Clin. Otolaryngol,* 23, 310-318.

Carding, P.N; Carlson, E.I.; Epstein, R.; Mathieson, L. & Shewell, C. (2000). Formal perceptual evaluation of voice quality in the United Kingdom. *Log Phon Vocol,* 25, 133-138.

Carlson, E.I. (1993). Accent method plus direct visual feedback of electroglottographic signals. In J. C. Stemple (Ed.), *Voice Therapy: Clinical studies* (pp.57-68). St. Louis: Pub. Mosby-Year Book.

Carlson, E.I. (1995a). *A study of voice quality in irradiated laryngeal cancer patients tumour stages T1 and T2.* Unpublished PhD Thesis, University of London.

Carlson, E.I. (1995b). Electrolaryngography in the assessment and treatment of incomplete mutation (puberphonia) in adults. *European Journal of Disorders of Communication,* 30, 140-149.

Carvalho, P.; Correia, P.; Quintel, A.; Pimentel, J. (1995). Vocal cord nodules. In the *abstract book of the 1st World Voice Congress*, April, Oporto, Portugal.

Casper, J.K.; Brewer, D.W. & Colton, R.H. (1987). Variations in normal human laryngeal anatomy and physiology as viewed fiberscopically. *Journal of Voice,* 1(2), 180-185.

Chae, S.W.; Choi, G.; Kang, H.J.; Choi, J.O.; Jin, S.M. (2001). Clinical analysis of voice change as a parameter of premenstrual syndrome. *Journal of Voice*, 15 (2), 278-283.

Chernobelsky, S. (1998). Effect of the menstrual cycle on laryngeal muscle tension of singers and nonsingers. *Log Phon Vocol,* 23, 128-132.

Cheyne, H.A; Nuss, R.C. & Hillman, R.E. (1999). Electroglottography in the pediatric population. *Arch Otolaryngol Head Neck Surgery,* 125, pp 1105-1108.

Childers, D.G.; Hicks, D.M.; Moore, G.P.; Eskenazi, L. & Lalwani, A.L. (1990). Electroglottography and vocal fold physiology. *JSHR,* 33, 245-254.

Clark, J. & Yallop, C. (1995). *An introduction to phonetics and phonology.* Oxford: Blackwells.

Coleman, R. F.; Markham, I. W. (1991). Normal variations in habitual pitch. *Journal of Voice,* 5 (2), 173-177.

Colton, R.H. & Conture, E.G. (1990). Problems and Pitfalls of Electroglottography. *Journal of Voice,* 4 (1), 10-24.

Colton, R.H.; Woo, P.; Brewer, D.W.; Griffin, B. & Casper, J. (1995). Stroboscopic signs associated with benign lesions of the vocal folds. *Journal of Voice*, 9 (3), 312-325.

Colton, R.H (1996). Physiology of voice quality. In Wm.S. Brown; B. P. Vinson; M. A. Crary (Eds.), *Organic voice disorders, assessment and treatment* (pp. 49-88). San Diego: Singular Publishing Group.

Colton, R.H. & Casper, J.K. (1996). *Understanding voice problems.* A physiological perspective for diagnosis and treatment. Baltimore: Williams & Wilkins.

Cooper, M. (1977). Voice disorders and symptoms. In Charles C. Thomas Publisher, *Modern Techniques of Vocal Rehabilitation* (pp.5-52). Springfield: Charles C. Thomas.

Cordes, A.K. (1994). The reliability of observational data: I. Theories and methods for Speech-Language Pathology. *JSHR,* 37, 264-278.

Correia, P.; Quintel, A.; Pimentel, J. (1995). Voice disorders. In the *abstract book of the 1st World Voice Congress*, April, Oporto, Portugal.

Crary, M.A. & Glowaski, A.L. (1996). Vocal fold immobility. In Wm.S. Brown; Betsy P. Vinson; Michael A. Crary. *Organic voice disorders* (pp.301-321). San Diego: Singular Publishing Group.

Cruz, M.C. (1986). Disfonia e saúde mental. *Boletim informativo da APTF*, 1, 18-22.

Cruz, M.C. & Barros, E. (1989). Disfonia por trauma cirurgico do nervo recorrente – relato de um caso. *Boletim Clínico dos HCL*, 46(3-4), 175-178.

Cruz, M.C. (1991). O meu filho está rouco. *Rua Sésamo – Guia de pais*, 44-45.

Cruz, M.C. & Guimarães, I. (1997). Saúde vocal. In Manuela de Sá. *Segredos da voz* (pp. 128-141). Mem Martins: Sebenta Editora.

Cruz, M.C. (2000). Reabilitação vocal. In S. B. Ruah (Ed.), *Manual de ORL* (pp.116-122). Lisboa: Edições Roche Farmacêutica Química Lda., 116-122.

Cummings, K.E. & Clements, M.A. (1995). Analysis of the glottal excitation of emotionally styled and stressed speech. *J. Acoust. Soc. Am.*, 98 (1), 88-98.

Cunha, C. & Cintra, L. (1984). *Nova gramática do Português Contemporâneo.* Lisboa: Edições João Sá da Costa.

Daniloff, R.; Schuckers, G. & Feth, L. (1980). *The physiology of speech and hearing, an introduction.* Englewood Cliffs: Prentice-Hall, Inc.

De Krom, G. (1995). Some spectral correlates of pathological breathy and rough voice quality for different types of vowel fragments. *JSHD*, 38, 794-811.

Debruyne, D.; Decoster, W. (1999). Acoustic differences between sustained vowels perceived as young or old. *Log Phon Vocol*, 24(1), 1-5.

Deem, J.F.; Manning, W.H.; Knack, J.V.& Matesich, J.S. (1989). The automatic extractions of pitch perturbation using microcomputers: some methodological considerations, *JSHR*, 32, 689-697.

Deem, J.F.; Manning, W.H.; Knack, J.V.& Matesich, J.S. (1991). Comparison of pitch perturbation extraction procedures with adult male and female speakers. *Folia Phoniatr*, 43, 234-245.

Dejonckere, P.H. & Lebacq, J. (1985). Electroglottography and vocal nodules. *Folia Phoniat*, 37, 195-200.

Dejonckere, P.H. (1998). Effect of louder voicing on acoustical measurements in dysphonic patients. *Log Phon Vocol*, 23(2), 79-84.

Doherty, E.T.; Shipp, T. (1988). Tape recorder effects on jitter and shimmer extraction. *JSHR*, 31, 485-490.

Drew, R.; Sapir, S. (1995). Average speaking fundamental frequency in soprano singers with and without symptoms of vocal attrition. *Journal of Voice*, 9(2), 134-141.

Dworkin, J.P. & Meleca, R.J. (1997). *Vocal pathologies, diagnosis, treatment and case studies.* San Diego: Singular Publishing Group.

Elias, M.E.; Sataloff, R.T.; Rosen, D.C.; Heuer, R.J. & Spiegel, J.R. (1997). Normal strobovideolaryngoscopy: variability in healthy singers. *Journal of Voice,* 11, 104-107.

Emanuel, F.W.; Sansone, F.E. (1969). Some spectral features of 'normal' and simulated 'rough' vowels. *Folia phoniat,* 21, 401-415.

Enderby, P. & Philipp, R. (1986). Speech and language handicap: towards knowing the size of the problem. *BJDC,* 21, 151-165.

Enderby, P. & Emerson, J. (1995). *'Does speech and language therapy work?' A review of the literature* (pp.151-165). London: Whurr Publishers.

Enderby, P. (1997). *Therapy outcome measures.* San Diego: Singular Publishing Group.

Epstein, R. (1999). *The impact of botulinum toxin injection in adductor spasmodic dysphonia – a cross sectional and longitudinal study.* Unpublished PhD Thesis, University of London.

Eskenazi, L.; Childers, D.G. & Hicks, D.M. (1990). Acoustic correlates of vocal quality. *JSHR,* 33, 298-306.

Eustace, C.S.; Stemple, J.C. & Lee, L. (1996). Objective measures of voice production in patients complaining of laryngeal fatigue. *Journal of Voice,* 10 (2), 146-154.

Fairbanks, G. (1960). *Voice and articulation drillbook.* New York: Harper.

Faure, M. & Muller, A. (1992). Stroboscopy. (Special article). *Journal of Voice,* 6 (2), 139-148.

Fawcus, M. (2000). The causes and classification of voice disorders. In M. Freeman & M. Fawcus (Eds.), *Voice disorders and their management.* (pp.47-68). Kent: Croom Hehn, Ltd.

Ferreira, A.; Cruz, M.C. (1986). A mastigação como forma de intervenção terapêutica. *Boletim informativo da APTF,*1, 23-25.

Ferreira, A.; Cruz, M.C.; Zagalo, C.; Montalvão, P.; Leandro, J.; Monteiro, L.; Vidal, M. (1996). Afonias e disfonias psicogénicas, análise de 100 casos clínicos. *Revista Portuguesa de Otorrinolaringologia,* 34 (1), 81-85.

Fex, S. (1992). Perceptual Evaluation. (Special article). *Journal of Voice,* 6 (2), 155-158.

Firmin, H.; Reilly, S. & Fourcin, A. (2000) Non-invasive monitoring of reflexive swallowing. *Work in progress,* UCL,1-14.

Fischer-Jorgensen, E. (1990). Intrinsic F0 in tense and lax vowels with special reference to German. *Phonetica*, 47, 99-140.

Fitch, J.L.; Holbrook, A. & Tallahassee, F. (1970). Modal vocal fundamental frequency of young adults. *Arch. Otolaryng*, 92, 379-382.

Fitch, J.L. (1990). Consistency of fundamental frequency and perturbation in repeated phonations of sustained vowels, reading, and connected speech. *JSHD*, 55:360-363.

Fourcin, A.J. & Abberton, E. (1971). First applications of a new Laryngograph, *Medical and Biological Illustration*, 21, 172-182, (reprinted in The Volta Review, vol 74, 3, 1972).

Fourcin, A.J. (1974). Laryngographic examination of vocal fold vibration. In B. Wyke (Ed.), *Ventilatory and Phonatory Control Systems* (pp. 315-333). New York: University Press.

Fourcin, A.J. & Abberton, E. (1976). The Laryngograph and the voiscope in speech therapy (pp.116-122). In E. Loebell (Ed.), *XVIth Int. Congr.Logopedics and Phoniatrics.*

Fourcin, A.J. (1981). Laryngographic assessment of phonatory function. Proceedings of the Conference on the Assessment of Vocal Pathology. *American Speech-Language-Hearing Association Reports*, 11, 116-127.

Fourcin, A.J. (1986). Electrolaryngographic assessment of vocal fold function. *Journal of Phonetics*, 14, 435-442.

Fourcin, A.J. (1989). An aspect of voice quality in speech pathology. *Work in progress*, UCL, 3, 255-261.

Fourcin, A.; Abberton, E. & Ball, V. (1993). Voice and intonation – analysis, presentation and training. In B. A.G. Elsendoorn & F. Coninx (Eds.), *Interactive learning technology for the deaf* (pp.137-150). Berlin: Springer-Verlag.

Fourcin, A.J. (1993). Normal and pathological speech: phonetic, acoustic and laryngographic aspects. Singh W., Sontar D.S. (Eds.), *Functional Surgery of the pharynx and larynx* (pp. 31-54). Butterworth Heineman.

Fourcin, A.J.; Abberton, E.; Miller, D. & Howells, D. (1995). Laryngography. *European Journal of Disorders of Communication*, 30, 101-115.

Fourcin, A.J. (2000). Voice quality and electrolaryngography. In R. D. Kent and M. J. Ball (Eds.), *Voice quality and electrolaryngography* (pp.285-306). San Diego: Singular Publishing Group.

Fratalli, C.M. (1998a). Assessing functional outcomes: an overview. *Seminars in speech and language*, 19 (3), 209-221.

Fratalli, C.M. (1998b). Outcomes measurement: definitions, and perspectives. In C. M. Fratalli (Ed.), *Measuring outcomes in speech-language pathology* (pp.1-27). New York: Thieme.

Frèche, Ch.; Dejean, Y.; Demard, F.; Demard, D.; Morgon, A.; Traissac, L.; Romanet, Ph.; Boudin, G.; Cornut, G.; Bouchayer, M.; Pfauwadel, M.C.; Jakobowicz, M.; Verhulst, J.; Yana, M. (1984). *La voix humaine et ses troubles.* Paris: Livrarie Arnette.

Freeman, M. (2000). Psychogenic, psychological and psychosocial issues in diagnosis and therapy. In M. Freeman & M. Fawcus (Eds.), *Voice disorders and their management* (pp.137-155). Kent: Croom Hehn, Ltd.

Fritzell, B.; Sundberg, J. & Strange-Ebbesen, A. (1982). Pitch change after stripping oedematous vocal folds. *Folia Phoniatr*, 34, 29-32.

Fritzell, B. (1996). Voice disorders and occupations. *Log Phon Vocol*, 21, 7-12.

Fujimura, O. (1981) Body-cover theory of the vocal fold and its phonetic implications. In K. Stevens and M. Hirano (Eds.). *Vocal fold physiology* (pp.271-288). Tokyo: University of Tokyo Press.

Garner, P.E. & Howard, D.M. (1999). Real-time display of voice source characteristics. *Log Phon Vocol*, 24, 19-25.

Garrett, K.L. & Healey, E.C. (1987). An acoustic analysis of fluctuations in the voices of normal adult speakers across three times of day. *J. Acoust. Soc. Am.*, 82, 58-62.

Gauffin, J. & Hammarberg, B. (1991). *Vocal fold physiology.* London: Whurr Publishers.

Gelfer, M.P. (1988). Perceptual attributes of voice: development and use of rating scales. *Journal of Voice*, 2, (4), 320-32.

Gelfer, M.P. & Bultemeyer, D.K. (1990). Evaluation of vocal fold vibratory patterns in normal voices. *Journal of Voice*, 4 (4), 335-345.

Gelfer, M.P.; Andrews, M.L. & Schmidt, C.P. (1991). Effects of prolonged loud reading on selected measures of vocal function in trained and untrained singers. *Journal of Voice*, 5(2), 158-167.

Gelfer, M.P. (1995). Fundamental frequency, intensity and vowel selection: effects on measures of phonatory stability. *JSHR*, 38, 1189-1198.

Gelfer, M.P.; Fendel, D.M. (1995). Comparisons of jitter, shimmer, and signal-to-noise ratio from directly digitised versus taped voice samples. *Journal of Voice*, 9 (4), 378-382.

Gerrat, B.R.; Kreiman, J. (2000). Theoretical and methodological development in the study of pathological voice quality. *Journal of Phonetics*, 28:335-342.

Gilbert, H. & Weismer, G. (1974). The effects of smoking on the fundamental frequency of adult women. *Journal of Psycholinguistic Research*, 3, 225-331.

Gleeson, M.J. & Fourcin, A.J. (1983). Clinical analysis of laryngeal trauma secondary to intubation. *Journal of the Royal Society of Medicine*, 73,928-932.

Goldman, S.L.; Hargrave, J.; Hillman, R.E.; Holmberg, E. & Gress, C. (1996). Stress, anxiety, somatic complaints, and voice use in women with vocal nodules: preliminary studies. *American Journal of Speech-Language Pathology*, 5, 44-54.

Gomes, F.; Santos, M.; Quintal, A.; Correia, P.; Pimentel, J.M. (2000). Disfonias Funcionais. *Rev. Port. ORL*, 38(1), 41-45.

Gordon, M. (1991). Assessment of dysphonic patient. M. Fawcus (Ed.) *Voice disorders and their management* (pp.39-72). London: Chapman & Hall.

Gould, W.J. & Korovin, G.S. (1994). Laboratory advances for voice measurements. *Journal of Voice*, 8 (1), 8-17.

Graddol, D.; Swann, J. (1983). Speaking fundamental frequency: some physical and social correlates. *Language and speech*, 26, 4, 351-366.

Greene, M.C.L. & Mathieson, L. (1989). *The voice and its disorders.* London: Whurr Publishers.

Guimarães, I. (1992). A voz um bem a preservar. *Fisioterapia*, 3(1), 21-25.

Guimarães, I. (1995). Child dysphonia: a case study. In *Abstract book of the 1st World Voice Congress*, April, Oporto, Portugal.

Guimarães, I. & Cruz, M.C. (1995). *Voz-Intervenção.* Lisboa: Fisiopraxis.

Guimarães, I. (1997). 'Um pouco de nós sobre voz' – relato de uma experiência pessoal- 1992/1996. *Pró-Fono*, 9, 35-40.

Guimarães, I. & Cruz, M.C. (1997a). *A disfonia na criança.* Lisboa: Fisiopraxis.

Guimarães, I. & Cruz, M.C. (1997b). Voice prevention: a Portuguese experience. In *Abstract book of the 3rd. European Congress of LCSTL (Standing Liaison Committee of E.U. Speech and Language Therapists and Logopedists)*, May, Lisbon, Portugal.

Haji, T.; Horiguchi, S.; Baer, T. & Gould, W. (1986). Frequency and amplitude perturbation analysis of electroglottograph during sustained phonation. *J. Acoust. Soc. Am.*, 80, 58-62.

Hajioff, D.; Rattenbury, H.; Carrie, S.; Carding, P. & Wilson, J. (2000). The effect of Isshiki type 1 thyroplasty on quality of life and vocal performance. *Clin. Otolaryngol*, 25, 418-422.

Hammarberg, B. & Gauffin, J. (1995). Perceptual and acoustic characteristics of quality differences in pathological voices related to physiological aspects. In O. Fujimura & M. Hirano. *Vocal fold physiology*. San Diego: Singular Publishing Group.

Hammarberg, B. (2000). Voice research and clinical needs. *Folia Phoniatr Logop*, 52, 93-102.

Harmon, D.K.; Masuda, M. & Holmes, T.H. (1970). The social readjustment rating scale: a cross-cultural study of western Europeans and Americans. *Journal of Psychosomatic Research*, 14, 391-400.

Harris, T.M. (1992). The pharmacological treatment of voice disorders. *Folia Phoniatr.*, 44, 143-154.

Haynes, W.O.; Pindzola, R.H. & Emerick, L.L. (1992). *Diagnosis and evaluation in speech pathology*. (4th edition). Englewood Cliffs: Prentice-Hall.

Hecker, M.H. & Kreul, E.J. (1971). Description of speech of patients with cancer of the vocal folds. Part I: measures of fundamental frequency. *J. Acoust. Soc. Am.*, 49, 1275-1282.

Heiberger, V.L. & Horii, Y. (1982). Jitter and shimmer in sustained phonation. *Speech and Language*, 7, 299-332.

Herrington-Hall, B.L.; Lee, L.; Stemple, J.C.; Niemi, K.R. & McHone, M.M. (1988). Description of laryngeal pathologies by age, sex, and occupation in a treatment-seeking sample. *JSHD*, 53, 57-64.

Hertgård, S. (1994). Normal and pathological glottal closure patterns at different ages. *Voice*, 3, 17-26.

Hertrich, I. & Ackermann, H. (1995). Gender-specific vocal dysfunctions in Parkinson's disease: electroglottographic and acoustic analyses. *Ann Otol Rhinol Laryngol*, 104, 197-202.

Higgins, M.B. & Saxman, J.H. (1989). Variations in vocal frequency perturbation across the menstrual cycle. *Journal of Voice*, 3(3), 233-243.

Higgins, M.B. & Saxman, J.H. (1991). A comparison of selected phonatory behaviors of healthy aged and young adults. *JSHR*, 34: 1000-1010.

Higgins, M.B.; Netsell, R. & Schulte, L. (1994). Aerodynamic and electroglottographic measures of normal voice production: intrasubject variability within and across sessions. *JSHR*, 37, 38-45.

Higgins, M.B.; Netsell, R. & Schulte, L. (1998). Vowel related differences in laryngeal articulatory and phonatory function. *JSHR*, vol. 41, 712-724.

Hill, D.S.; Akhtar, S.; Corroll, A. & Croft, C.B. (2000). Quality of life issues in recurrent respiratory papillomatosis. *Clin Otolaryngol*, 25, 153-160.

Hillenbrand, J.; Getty, L.A.; Clark, M. & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *J. Acoust. Soc. Am.*, 97 (5), 3099-3109.

Hillenbrand, J. & Houde, R.A. (1996). Acoustic correlates of breathy vocal quality: dysphonic voices and continuous speech. *JSHR*, 39, 311-321.

Hillman, R.E.; Holmberg, E.B.; Perkell, J.S.; Walsh, M.; Vaughan, C. (1990). Phonatory function associated with hyperfunctionally related vocal fold lesions. *J. Voice*, 4, 52-63.

Hirano, M. (1981). *Clinical Examination of voice.* New York: Sprienger-Verlag.

Hirano, M. (1989). Objective evaluation of the human voice: clinical aspects. *Folia Phoniatr*, 41, 89-144.

Hirano, M. & Bless, D.M. (1993). *Videostroboscopic examination of the Larynx.* London: Whurr Publishers.

Hogikyan, N.D. & Sethuraman, G. (1999). Validation of an instrument to measure voice-related quality of life (V-RQOL). *Journal of Voice*, 13 (4), 557-569.

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

Hollien, H. & Michel, J. (1968). Vocal fry as a phonational register. *JSHR*, 11, 600-604.

Hollien, H. & Shipp, T. (1972). Speaking fundamental frequency and chronologic age in males. *JSHR*, 15, 155-159.

Hollien, H. & Jackson, B. (1973). Normative data on the speaking fundamental frequency characteristics of young adult males. *Journal of Phonetics*, 1, 117-120.

Hollien, H.; Michel, J. & Doherty, E.T. (1973) A method of analysing vocal jitter in sustained phonation. *Journal of Phonetics*, 1, 85-91.

Hollien, H. (1974). On vocal registers. *Journal of Phonetics*, 2, 125-143.

Hollien, H. (1983). A review of vocal registers. In Laurence V. (Ed.), *Transcripts of the twelfth symposium on care of professional voice* (pp.1-5). New York: Voice Foundation.

Hollien, H. (1987). Old voices: what do we really know about them? *Journal of Voice,* 1, 87-101.

Hollien, H.; Hollien, P. & De Jong, G. (1997). Effects of three parameters on speaking fundamental frequency. *J. Acoust. Soc. Am.,* 5, 2984-2992.

Hollien, H. (2000). The concept of ideal voice quality. In R. D. Kent and M. J. Ball, *Voice Quality Measurement* (pp. 13-24). San Diego: Singular Publishing Group.

Holmberg, E.B.; Perkell, J.S.; Hillman, R.E.; Gress, C. (1994). Individual variation in measures of voice. *Phonetica,* 51, 30-37.

Holmes, T.H. & Rahe, R.H. (1967). The social readjustment rating scale. *Journal of psychosomatic research,* vol. II, n° 2, 213-217.

Holmes, T.H. & David, E.M. (1989). *Life change, life events, and illness.* New York: Praeger.

Honjo, I. & Isshiki, N. (1980). Laryngoscopic and voice characteristics of aged persons. *Arch Otolaryngol,* 106, 149-150.

Horii, Y. (1975). Some statistical characteristics of voice fundamental frequency. *JSHR,* 18, 192-201.

Horii, Y. (1979). Fundamental frequency perturbation observed in sustained phonation. *JSHR,* 22, 5-19.

Horii, Y. (1982). Jitter and shimmer differences among sustained vowel phonations. *JSHR,* 25, 12-14.

House, A.O. & Andrews, H.B. (1987). The psychiatric and social characteristics of patients with functional dysphonia. *Journal Psychosomatic Research,* 31 (4), 483-490.

Howard, D.M. & Lindsey, G. (1988). Conditioned variability in voicing offsets. *IEE transactions on acoustic speech, and signal processing,* 36 (3), 406-407.

Howard, D.M.; Lindsey, G.A.; Allen, B. (1990). Toward the quantification of vocal efficiency. *Journal of Voice,* 4(3), 205-212.

Howard, D.M. (1995). Variation of electrolaryngographically derived closed quotient for trained and untrained adult female singers. *Journal of Voice,* 9(2), 163-172.

Howard, D.M. (1998). Instrumental voice measurement: uses and limitations. In T. Harris; S. Harris; J.S. Rubin & D. M. Howard (Eds). *The voice clinic handbook* (pp. 323-382). London: Whurr Publishers.

Howell, D.C. (1997). *Statistical methods for Psychology.* Belmont: Wadsworth Publishing Company.

Hudson, A.I. & Holbrook, A. (1981). A study of the reading fundamental vocal frequency of young black adults. *JSHR*, 24, 197-201.

Hudson, A.I. & Holbrook, A. (1982). Fundamental frequency characteristics of young black adults: spontaneous speaking and oral reading. *JSHR*, 25, 25-28.

Ingrisano, D.R-S.; Perry, C.K.; Jepson, K.R. (1998). Environmental noise: a threat to automatic voice analysis. *American Journal of Speech-Language Pathology*, 7(1): 91-96.

Iwata, S. & von Leden, H. (1970). Pitch perturbations in normal and pathologic voices. *Folia Phoniat*, 22, 413-424.

Jackson-Menaldi, M.C. A. (1992) *La voz normal* (pp.15-26). Madrid: Editorial Panamericana.

Jacobson, B.H.; Johnson, A.; Grywalski, C.; Silbergleit, A.; Jacobson, G.M; Benninger, M. & Newman, C.W. (1997). The Voice Handicap Index (VHI): Development and validation. *American Journal of Speech-Language Pathology*, 6, 66-70.

Kaplan, H.M. (1971) *Anatomy and Physiology of speech* (pp.203-274). New York: McGraw-Hill Book Company.

Karnell, M.P. (1991) Laryngeal perturbation analysis: minimum length of analysis window. *JSHR*, 34, 544-548.

Kempster, G.; Kistler, D. & Hillenbrand, J. (1991). Multidimensional scaling analysis of dysphonia in two speaker groups. *JSHR*, 34, 534-543.

Kent, R.D. (1977). *The speech sciences.* San Diego: Singular Publishing Group.

Kingston, J.; Macmillan, N.A.; Dickey, L.W.; Thorbun, R. & Bartels, C. (1997). Integrality in the perception of tongue root position and voice quality in vowels. *J. Acoust. Soc. Am.*,101(3), 1696-1709.

Kinzl, J.; Biebl, W. & Rauchegger, H. (1988). Functional Aphonia: Psychosomatic aspects of diagnosis and therapy. *Folia Phoniatr*, 40, 131-137.

Kitajima, K. (1981). Quantitative evaluation of the noise level in the pathologic voice. *Folia Phoniatr*, 33, 115-124.

Kitzing, P. (1990). Clinical applications of electroglottography. *Journal of Voice*, 4(3), 238-249.

Klatt, D.H. & Klatt, L.C. (1990). Analysis, synthesis, and perception of voice quality variations among female and male talkers. *J. Acoust. Soc. Am.*, 87 (2), 820-857.

Klingholz, F.; Martin, F. (1985). Quantitative spectral evaluation of shimmer and jitter. *JSHR*, 28, 169-174.

Klingholz, F.; Penning, R.; Liebhardt, E. (1988). Recognition of low-level alcohol intoxication from speech signal. *J. Acoust. Soc. Am.*, 84(3), 929-935.

Klingholtz, F. (1990). Acoustic recognition of voice disorders: a comparative study of running speech versus sustained vowels. *J. Acoust. Soc. Am.*, 87(5), 2218-2224.

Koike, Y.; Takahashi, H. & Calcaterra, T.C. (1977). Acoustic measures for detecting laryngeal pathology. *Acta Otolaryngol*, 84, 105-117.

Koschkee, D.L. & Rammage, L. (1997). *Voice care in the medical setting.* San Diego: Singular Publishing Group.

Kotby, M.N.; Titze, I.R.; Saleh, M.M. & Berry, D.A. (1993). Fundamental frequency stability in functional dysphonia. *Acta Otolaryngol (Stockh)*, 113, 439-444.

Kreiman, J.; Gerratt, B.R.; Precoda, K. & Berke, G.S. (1992). Individual differences in voice quality perception. *JSHR*, 35, 512-520.

Kreiman, J.; Gerratt, B.R.; Kempster, G.B.; Erman, A. & Berke, G.S. (1993). Perceptual evaluation of voice quality: review, tutorial, and a framework for future research. *JSHR*, 36, 21-40.

Krom, G. (1995). Some spectral correlates of pathological breathy and rough voice quality for different types of vowel fragments. *JSHR*, 38(4), 794-811.

Krook, M.I.P. (1988). Speaking fundamental frequency characteristics of normal swedish subjects obtained by glottal frequency analysis. *Folia Phoniatr*, 40, 82-90.

Ladefoged, P. (1971). *Preliminaries to linguist phonetics.* Chicago : University Chicago Press.

Ladefoged, P. (1996). *Elements of acoustic phonetics* (2nd edition). Chicago: University Chicago Press.

Laver, J. (1980). *The phonetic description of voice quality.* Cambridge: Cambridge University Press.

Laver, J. (1991). *The gift of speech.* Edinburgh: Edinburgh University Press.

Laver, J.; Hiller, S. & Beck, J.M. (1992). Acoustic waveform perturbations and voice disorders. Special article. *Journal of Voice*, 6 (2), 115-126.

Le Huche, F. & Allali, A. (1990) *La voix* (volume 1 & 2). Paris: Masson.

Leff, J. & Abberton, E. (1981). Voice pitch measurements in schizophrenia and depression. *Psychological Medicine*, 11, 849-852.

Lieberman, P. (1961). Perturbations in vocal pitch. *J. Acoust. Soc. Am.*, 33, 597-602.

Lieberman, P. (1963). Some acoustic measures of the fundamental periodicity of normal and pathologic larynges. *J. Acoust. Soc. Am.*, 33, 597-602.

Linke, S.E. (1973). A study of pitch characteristics of female voices and their relationship to vocal effectiveness. *Folia Phoniatr. Logop,* 25, 173-185.

Linville, S.E. & Korabic, E.W. (1987). Fundamental frequency stability characteristics of elderly women's voices. *J. Acoust. Soc. Am.* 81(4), 1196-1199.

Linville, S.E. (1988). Intraspeaker variability in fundamental frequency stability: an age-related phenomenon? *J. Acoust. Soc. Am.*, 83(2), 741-745.

Linville, S.E. (1995). Changes in configuration in women after loud talking. *Journal of voice,* 9, 1, 57-65.

Linville, S.E. (1996). The sound of senescence. *Journal of Voice,* 10(2), 190-200.

Lockhart, M. & Martin, S. (1997). The Victoria infirmary voice questionnaire. In S. Martin (Ed.), *Working with Dysphonics* (pp.14-24). Oxon: Winslow Press.

Loveday (1981). Pitch, politeness and sexual role: an exploratory investigation into the pitch correlates of English and Japanese politeness formulae. *Lang. Speech,* 24:71-89.

Ludlow, C.L. (1981). Research needs for the assessment of phonatory function. *ASHA Reports,* 11, 3-8.

Ma, E.P-M.; Yiu, E. M-L. (2001) Voice activity and participation profile: assessing the impact of voice disorders on daily activities. *JSLHR,* 44, 511-524.

MacCurtain, F. & Fourcin, A. (1981). Application of the electro-laryngograph wave form display. *Tenth Symposium care of professional voice.* New York, 51-56.

Majewski, W.; Hollien, H. & Zalewski, J. (1972). Speaking fundamental frequency of Polish adult males. *Phonetica,* 25, 119-125.

Martin, D.; Fitch, J. & Wolfe, V. (1995). Pathologic voice type and the acoustic prediction of severity. *JSHR,* 38(4), 765-771.

Martin, F.G. (1983). Drugs and the voice. *Transcripts of the twelfth symposium care of the professional voice* (pp.124-132). New York: The voice foundation.

Martin, F.G. (1984). Drugs and the voice – Part II. *Transcripts of the twelfth symposium care of the professional voice* (pp. 191-201). New York: The voice foundation.

Martin, S. (1993). *Working with dysphonics.* London: Whurr Publishers.

Masuda, M. & Holmes, T. (1967). The social readjustment rating scale: a cross-cultural study of Japanese and Americans. *Journal of Psychosomatic Research,* 11, 227-237.

Mathieson, L. (2000). Normal-disordered continuum. In R. D. Kent and M. J. Ball (Ed.), *Voice Quality Measurement* (pp.1-12). San Diego: Singular Publishing Group.

Mathieson, L. (2001). *Greene & Mathieson's the voice & its disorders* (6th ed.). London: Whurr Publishers.

McAllister, A.; Sundberg, J. & Hibi, S.R. (1998). Acoustic measurements and perceptual evaluation of hoarseness in children's voices. *Log Phon Vocol,* 23(1), 27-38.

McGlone, R.E. & Hollien, H. (1963). Vocal pitch characteristics of aged women. *JSHR,* 6(2), 164-170.

Milenkovic, P. (1987). Least mean square measures of voice perturbation. *JSHR,* 30, 529-538.

Millar, J.B. & Wagner, M. (1983). The automatic analysis of acoustic variance in speech. *Language and speech,* 26, 2, 145-158.

Miller, T. (1989). *Stressful life events.* Madison: International University Press, Inc.

Millet, B. & Dejonckere, P.H. (1998). What determines the differences in perceptual rating of dysphonia between experienced raters? *Folia Phoniatr,* 50:305-310.

Milutinovic, Z. (1996). Classification of voice pathology. *Folia Phoniatr Logop,* 48, 301-308.

Morrison, M.; Rammage, L.; Nichol, H.; Pullan, B.; May, P. & Salked, L. (1994). *The management of voice disorders.* London: Chapman & Hall Medical.

Morton, V. & Watson, D.R. (1998). The teaching voice: problems and perceptions. *Log Phon Vocol,* 23, 52, 133-139.

Morton, R. & Tandon, D. (2000). Validation of quality-of-life measures in head and neck cancer patients. *Aust. J. Otolaryng.,* 3 (5), 479-487.

Murdoch, B.E.; Manning, C.Y.; Theodoros, D.G. & Thompson, E.C. (1997). Laryngeal and phonatory dysfunction in Parkinson's disease. *Clinical Linguistics & Phonetics,* 11 (3), 245-266.

Murphy, C.H.; Doyle, P.C. (1987). The effects of cigarette smoking on voice-fundamental frequency. *Otolaryngolog-Head and Neck Surgery,* 97 (4), 376-380.

Murray, T. & Doherty, E.T. (1980) Selected acoustic characteristics of pathologic and normal speakers. *Journal of Speech and Hearing Research,* 23, 361-369.

Murty, G.E.; Carding, P.N. & Lancaster, P. (1991). An outpatient clinic system for glottographic measurement of vocal fold vibration. *British Journal of Disorders of Communication*, 26, 115-123.

Murry, T. (1978). Speaking fundamental frequency characteristics associated with voice pathologies. *JSHD*, XLIII, 374-379.

Murry, T.; Brown, W.S. & Morris, R.J. (1995). Patterns of fundamental frequency for three types of voice samples. *Journal of Voice*, 3, 282-289.

Mysak, E.D. (1959). Pitch and duration characteristics of older males. *JSHR*, 2(1), 46-54.

Neil, W.F.; Wechsler, E. & Robinson, J.M.P. (1977). Electrolaryngography in laryngeal disorders. *Clinical Otolaryngology*, 2, 33-40.

Newman, C.; Kersner, M. (1998). Voice problems of aerobics instructors: implications for preventative training. *Log Phon Vocol*, 23(4), 177-180.

Niedzielska, G.; Pruszewicz, A. (1996). Acoustic and morphological lesions of the larynx in alcoholics. In M. Pais Clemente (Ed.), *Voice update*, (pp. 357-362). Amsterdam: Elsevier.

Nittrouer, S.; McGowan, R.S.; Milenkovic, P.H. & Beehler, D. (1990). Acoustic measurements of men's and women' voices: a study of context effects and covariation. *JSHR*, 33, 761-775.

Oates, J. (2000). Voice disorders associated with hyperfunction. In M. Freeman & M. Fawcus (Eds.), *Voice disorders and their management.* (pp.110-136). Kent: Croom Hehn, Ltd.

Omori, K.; Kojima, H.; Kakani, R.; Slavit, D.H. & Blaugrund, S.M. (1997). Acoustic characteristics of rough voice: subharmonics. *Journal of Voice*, 11(1), 40-47.

Orlikoff, R.F. & Baken, R.J. (1990). Consideration of the relationship between the fundamental frequency of phonation and vocal jitter. *Folia Phoniatr*, 42(3), 31-40.

Orlikoff, R.F. & Kahane, J.C. (1991). Influence of mean sound pressure level on jitter and shimmer measures. *Journal of Voice*, 5(2), 113-119.

Orlikoff, R.F (1991). Assessment of the dynamics of vocal fold contact from the electroglottogram: data from normal male subjects. *JSHR*, 4, 1066-1072.

Orlikoff, R.F (1995). Vocal stability and vocal tract configuration: an acoustic and electroglottographic investigation. *Journal of Voice*, 9, 173-181.

Orlikoff, R.F.; Kraus, D.H.; Harrison, L.B.; Ho, M.L. & Gartner, C.J. (1997). Vocal fundamental frequency measures as a reflection of tumor response to

chemotherapy in patients with advanced laryngeal cancer. *Journal of Voice*, 11, 33-39.

Orlikoff, R.F (1998). Scrambled EGG: the uses and abuses of electroglottography. *Phonoscope*, 37-53.

Orlikoff, R.F; Dejonckere, P.H.; Dembowski, J. et al. (1999). The perceived role of voice perception in clinical practice. *Phonoscope*, 2, 89-106.

Painter, C. (1988). Electroglottogram waveform types. *Oto-Rhino-Laryngology*, 245,116-121.

Parsa, V. & Jamieson, D.G. (2000). Identification of pathological voices using glottal noise measures. *JSLHR*, 43, 469-485.

Pemberton, C.; Russel, A.; Priestley, J.; Havas, T.; Hooper, J. & Clark, P. (1993). Characteristics of normal larynges under flexible fiberscopic and stroboscopic examination: an Australian perspective. *Journal of Voice*, 7(4), 382-389

Peppard, R.C.; Bless, D.M. & Milenkovic, P. (1988). Comparison of young adult singers and nonsingers with vocal nodules. *Journal of Voice*, 2(3), 250-260.

Perkins, W.H. & Kent, R.D. (1986). *Textbook of functional anatomy of speech, language, and hearing*. London: Taylor & Francis.

Perlman, A.L. & Liang, H. (1991). Frequency response of the Fourcin Electroglottograph and measurement of temporal aspects of laryngeal movement during swallowing. *JSHR*, 34, 791-795.

Perry, C.K.; Ingrisano, D.R-S. & Blair, W.B. (1996). The influence of recording systems on jitter and shimmer estimates. *American Journal of Speech-Language Pathology*, 5 (2), 86-90.

Piccirillo, J.F.; Painter, C.; Haiduk, A; Fuller, D. & Fredrickson, J.M. (1998). Assessment of two objective voice function indices. *Ann Otol Rhinol Laryngol*, 107, 396-400.

Pindzola, R.H. (1987). *A voice assessment protocol for children and adults.* Austin: Pro-Ed.

Pinto, O. de & Hollien, H. (1982). Speaking fundamental frequency characteristics of Australian women: then and now. *J. Phonetics*, 10, 367-375.

Poburka, B.J. (1999). A new stroboscopy rating form. *Journal of Voice*, 13 (3), 403-413.

Pontes, P.A.L.; Gonçalves, M.I. & Behlau, M. (1999). Vocal fold cover minor structural alterations: diagnostic errors. *Phonoscope*, 2 (4), 175-185.

Prater, R.J. & Swift, R.W. (1984). *Manual of voice therapy* (pp.1-13). Boston: Little, Brown and Company.

Pruszewicz, A.; Obrebowski, A.; Swidzinski, P.; Demenko, G.; Wika, T. & Wojciechowska, A. (1991). Usefulness of acoustic studies on the differential diagnosis or organic and functional dysphonia. *Acta Otolaryngol (Stockh)*, 111, 414-419.

Rabinov, C. R.; Kreiman, J.; Gerratt, B.R. & Bielamowicz, S. (1995). Comparing reliability of perceptual ratings of roughness and acoustic measures of jitter. *JSHR*, 38, 26-32.

Ramig, L.A. & Ringel, R.L. (1983) Effects of physiological ageing on selected acoustic characteristics of voice. *JSHR*, 26, 22-30.

Rammage, L.A.; Peppard, R.C. & Bless, D.M. (1992). Aerodynamic, laryngoscopic and perceptual acoustic characteristics in dysphonic females with posterior glottal chinks. A retrospective study. *Journal of Voice*, 6 (1), 64-78.

Rantala, L.; Lindholm, P. & Vilkman, E. (1998). Fo change due to voice loading under laboratory and field conditions. A pilot study. *Log Vocol Phon*, 23(4), 164-168.

Rantala, L.; Määttä, T. & Vilkman, E. (1997). Measuring voice under teachers' working circumstances: F0 and perturbation features in maximally sustained phonation. *Folia Phoniatri Logop*, 49, 281-291.

Rodrigues, A. (1998). *Estudo comparativo dos parâmetros acústicos (F0, Jitter, Simmer and NNE) em mulheres fumadoras e não fumadoras.* Bsc. in Speech Therapy. Alcoitão: Escola Superior de Saúde do Alcoitão.

Rontal, E.; Rontal, M.; Jacob, H.J. & Rohnic, M.I. (1983). Quantitative and objective evaluation of vocal cord function. *Annals of Otolaryngology, Rhinology and Laryngoscopy*, 92, 421-423.

Rosen, C. & Murry, T. (2000). Voice handicap index in singers. *Journal of voice*, 14 (3), 370-377.

Rosen, C.A.; Murry, T.; Zinn, A.; Zullo, T. & Sonbolian, M. (2000). Voice handicap index change following treatment of voice disorders. *Journal of Voice*, 14 (4), 619-623.

Rosen, D.C. & Sataloff, R.T (1997). *Psychology of voice disorders.* San Diego: Singular Publishing Group.

Ross, J.; Noordzji, J.P. & Woo, P. (1998). Voice disorders in patients with suspected laryngo-pharyngeal reflux disease. *Journal of voice*, 12,1, 84-88.

Rossiter, D.; Howard, D.M. & Comins, R. (1995). Objective measurement of voice source and acoustic output change with a short period of vocal tuition. *Voice*, 4, 16-31.

Rubin, J. (1998). The structural anatomy of the larynx and supraglottic vocal tract: a review. In T. Harris, S. Harris, J.S. Rubin & D.M. Howard (Eds.), *The voice clinic handbook* (pp. 15-33). London: Whurr Publishers, Ltd.

Ruch, L.O. & Holmes, T.H. (1971). Scaling of life change: comparison of direct and indirect methods. *Journal of Psychosomatic Research*, 15, 221-227.

Russel, A.; Penny, L. & Pemberton, C. (1995). Speaking fundamental frequency changes over time in women: a longitudinal study. *JSHR*, 38, 101-109.

Ryan, W.J. (1998). Long-term variability in speaking fundamental frequency. *Phonoscope*, 3, 155-164.

Sapir, S. (1989). The intrinsic pitch of vowels: theoretical, physiological, and clinical considerations. *Journal of Voice*, 3(1), 44-51.

Sapir, S.; Keidar, A. & Mathers-Schmidt, B. (1993). Vocal attrition in teachers: survey findings. *EJDC*, 28, 177-185.

Sapir, S.; Mathers-Schmidt, B. & Larson, G.W. (1996). Singers' and non-singers' vocal health, vocal behaviours, and attitudes towards voice and singing: indirect findings from a questionnaire. *EJDC*, 31, 193-209.

Sataloff, R.T.; Spiegel, J.R.; Hawkshaw, M.J. & Heuer, R.J. (1994a). Professional voice users: Obtaining the history. In M. S. Benninger; B. H. Jacobson & A. F. Johnson (Eds.), *Vocal Arts Medicine, the care and prevention of professional voice disorders* (pp.72-78). New York: Thieme.

Sataloff, R.T; Lawrence, V.L.; Hawkshaw, M.J. & Rosen, D.C. (1994b). Medications and their effects on the voice. In M. S. Benninger; B., H. Jacobson; A. F. Johnson (Eds.), *Vocal arts medicine, the care and prevention of professional voice disorders* (pp.216-225). New York: Thieme.

Sataloff, R.T.; Spiegel, J.R. & Hawkshaw, M.J. (1995). History and physical examination of patients with voice disorders. In J. Rubin, R.T. Sataloff et al. (Eds.), *Diagnosis and treatment of voice disorders* (pp. 247-261). New York: Igaku-Shoin.

Sataloff, R.T. (1997). *Professional voice: The science and art of clinical care.* (2nd ed.). San Diego: Singular Publishing Group.

Sataloff, R.T. (1998). *Vocal health and pedagogy.* San Diego: Singular Publishing Group.

Saxman, J.H. & Burk, K.W. (1967). Speaking fundamental frequency characteristics of middle-aged females. *Folia Phoniatrica,* 19, 167-172.

Scherer, R. C.; Vail, V.J. & Guo, C.G. (1995). Required number of tokens to determine representative voice perturbation values. *JSHR,* 38, 1260-1269.

Scherer, R.C.; Gould, W.J.; Titze, I.R.; Arlen, D.M. & Sataloff, R.T. (1988). Preliminary evaluation of selected acoustic and glottographic measures for clinical phonatory function analysis. *J. Voice,* 2, 230-244.

Scott, S.; Robinson, K.; Wilsom, J.A. & Mackenzie, K. (1997). Patient-reported problems associated with dysphonia. *Clin. Otolaryngol,* 22, 37-40.

Seikel, J.A.; King, D.W. & Drumright, D.G. (1997). *Anatomy and physiology for speech and language.* San Diego: Singular Publishing Group.

Selby, J. & Wilson, G. (1997). Laryngographic assessment of voice changes with altered hydration status. *Speech, Hearing and Language: Working in Progress,* UCL, 10, 153-170.

Shipp, T.; Qi, Y.; Huntley, R. & Hollien, H. (1992). Acoustical and temporal correlates of perceived age. *Journal of Voice,* 6 (3), 211-216.

Silverman, E.-M.; Zimmer, C.H. (1978). Effect of the menstrual cycle on voice quality. *Arch Otolaryngol,* 104, 7-10.

Simberg, S.; Laine, A.; Sala, E. & Ronnemaa, A-M. (2000). Prevalenve of voice disorders among future teachers. *Journal of Voice,* 14, 231-235.

Smith, E.; Verdolini, K.; Gray, S. et al. (1996). Effect of voice disorders on quality of life. *Journal of Medical Speech-Language Pathology,* 4 (4), 223-244.

Snidecor, J.C. (1943). A comparative study of pitch and duration characteristics of impromptu speaking and oral reading. *Speech Monographs,* 10, 50-57.

Sobell, L.C. & Sobell, M.B. (1972). Effects of alcohol on the speech of alcoholics. *JSHR,* 15, 861-868.

Sobell, L.C. & Sobell, M.B.; Coleman, R.F. (1982). Alcohol-induced dysfluency in nonalcoholics. *Folia Phoniatric,* 34, 316-323.

Söderstein, M. & Lindestad, P.A. (1990). Glottal closure and perceived breathiness during phonation in normally speaking subjects. *JSHR,* 33, 601-611.

Söderstein, M.; Hertegård, S. & Hammarberg, B. (1995). Glottal closure, transglottal airflow, and voice quality in healthy middle-aged women. *Journal of Voice,* 9(2), 182-197.

Sonninen, A. & Damste, P.H. (1971). An international terminology in the field of Logopedics and Phoniatrics. *Folia Phoniat,* 23, 1-32.

Sonninen, A. & Hurme, P. (1992). On the terminology of voice research. (Special article). *Journal of Voice*, 6(2), 188-193.

Sorensen, D.; Horii, Y.; Leonard, R. (1980). Effects of laryngeal topical anesthesia on voice fundamental frequency perturbation. *JSHR*, 23, 274-83.

Sorensen, D. & Horii, Y. (1982). Cigarette smoking and voice fundamental frequency. *J. Comm. Disord.*, 15, 135-144.

Stemple, J.C. (1993). *Voice therapy, clinical studies.* St. Louis: Mosby-Year Book, Inc.

Stemple, J.C.; Glaze, L.E. & Gerdeman, B.K. (1995). *Clinical voice pathology.* (2nd edition) San Diego: Singular Publishing Group.

Stemple, J.C.; Stanley, J. & Lee, L. (1995). Objective measures of voice production in normal subjects following prolonged voice use. *Journal of Voice*, 9(2), 127-133.

Stevens, K. (1998). *Acoustic phonetics.* Massachusetts: Massachusetts institute of technology.

Stewart, M.G.; Chen, A.Y. & Stach, C.B. (1998). Outcomes analysis of voice and quality of life in patients with laryngeal cancer. *Arch Otolaryngol Head Neck Surg.*,124, 143-148.

Stoicheff, M.L. (1981). Speaking fundamental frequency characteristics of nonsmoking female adults. *JSHR*, 24, 437-441.

Stone, R.E. & Rainey, C.L. (1991). Intra- and intersubject variability in acoustic measures of normal voice. *Journal of Voice*, 5(3), 189-196.

Strik, H. & Boves, L. (1992). Control of fundamental frequency, intensity and voice quality in speech. *Journal of Phonetics*, 20, 15-25.

Sulter, A.M.; Schutte, H.K. & Miller, D.G. (1996). Standardized laryngeal videostroboscopic rating: differences between untrained and trained male and female subjects, and effects of varying sound intensity, fundamental frequency, and age. *Journal of Voice*, 10 (2), 175-189.

Sussman, J.E. & Sapienza, C. (1994). Articulatory, developmental, and gender effects on measures of fundamental frequency and jitter. *Journal of Voice*, 2, 145-156.

Svec, J.G.; Shutte, H.K. & Miller, D.G. (1996). A subharmonic vibratory pattern in normal vocal folds. *JSHR*, 39, 135-143.

Terrel, J.E.; Fisher, S.G.; Wolf, G.T. (1998). Long term quality of life after treatment of laryngeal cancer. *Arch Otolaryngol Head Neck Surg*, 124, 964-971.

Titze, I.R.; Horii, Y. & Scherer, R.C. (1987). Some technical considerations in voice perturbation measurements. *JSHR*, 30, 252-260.

Titze, I.R. & Winholtz, W.S. (1993). Effect of microphone type and placement on voice perturbation measurements. *JSHR*, 36, 1177-1190.

Titze, I.R. (1994a). Toward standards in acoustic analysis of voice. *Journal of Voice*, 8 (1), 1-7.

Titze, I.R. (1994b). *Principles of voice production.* Englewood Cliffs: Prentice –Hall, Inc.

Trojan, F.; Kryspin-Exner, K. (1968). The decay of articulation under the influence of alcohol and paraldehyde. *Folia Phoniatr.*, 20, 217-238.

Tucker, H.M. (1987) *The Larynx.* New York: Thieme Medical Publishers, Inc., 1-32.

Tveteras, G. (1992). Perceptual analysis of voice quality. *Scand J Log Phon*, 17(3-4): 145-149.

Van Riper, C. & Irwin, J.V. (1958). *Voice and articulation.* Englewood Cliffs: Prentice-Hall.

Verdolini, K. (1994). Voice disorders. In Bruce Tomblin; Hughlett L.Morris; D.C. Spriestersbach (Eds), *Diagnosis in speech-language pathology* (pp. 247- 306). San Diego: Singular Publishing Group.

Verdolini, K; Palmer, P.M. (1997). Assessment of a 'Profiles' approach to voice screening. *Journal of Medical Speech-Language Pathology*, 5(4), 217-232.

Verdolini, K.; Drucker, D.G.; Palmer, P.M. & Samawi, H. (1998). Laryngeal adduction in resonant voice. *Journal of Voice*, 12 (3), 315-327.

Verdolini, K.; Ramig, L.; Jacobson, B. (1998). Outcomes measurement in voice disorders. In C. M. Frattali (Ed.), *Measuring outcomes in speech-language pathology* (pp.354-386). New York: Thieme.

Verdolini, K. & Ramig, L.O. (2001). Review: Occupational risks for voice problems. *Log Phon Vocol*, 26, 37-46.

Verdonck-de Leeuw, I.M. (1998). *Voice characteristics following radiotherapy: the development of a protocol.* Amsterdam: Studies in language and language use.

Verstraete, J.; Forrez, G.: Mertens, P. & Debruyne, F. (1993). The effect of sustained phonation at high and low pitch on vocal jitter and shimmer. *Folia Phoniatr*, 45, 223-228.

Viana, M.C. (1987). *Para a síntese da entoação do Português.* Unpublished dissertation, Centro de Linguística da Universidade de Lisboa.

Vilkman, E. (1996). Occupational risk factors and voice disorders. *Log Phon Vocol*, 21(1), 137-141.

Vilkman, E.; Lauri, E-R.; Alku, P.; Sala, E. & Shivo, M. (1998). Ergonomic conditions and voice. *Log Phon Vocol*, 23(1), 11-19.

Vilkman, E. (2000). Voice problems at work: a challenge for occupational safety and health arrangement. *Folia Phoniatr Logop*, 52, 120-125.

Watson, C. (1995). Quality analysis of laryngography in a busy hospital ENT voice clinic. *European Journal of Disorders of Communication*, 30, 132-139.

Wechsler, E. (1975). *The use of the Laryngograph in the study of some patients with voice disorders.* M.Sc. in Human Communication, University of London.

Wechsler, E. (1977). Laryngographic study of voice disorders. *British Journal of Disorders of Comunication*, 12, pp. 9-22.

Wendler, J. (1992). Stroboscopy. (Special article). *Journal of Voice*, 6 (2), 149-154.

Whalen, D.H. & Levitt, A.G. (1995). The universality of intrinsic Fo of vowels. *Journal of Phonetics*, 23, 349-366.

Whalen, D.H.; Gick, B.; Kumada, M. & Honda, K. (1998). Cricothyroid activity in high and low vowels: exploring the automatic of intrinsic Fo. *Journal of Phonetics*, 27, 125-142.

Wilson, K. (1987). *Voice problems in children* (pp.110-124). Baltimore: Williams & Wilkins.

Winding, O.; Clamp, S.; Ohmann, C. (1993). Quality and its evaluation. *Technology and Health Care*, 1, 101-105.

Winstanley, S. & Wright, H. (1991). Vocal fold contact area patterns in normal speakers: an investigation using the electrolaryngograph interface system. *BJDC*, 26, 25-39.

Wolfe, V.I. & Steinfatt, T.M. (1987). Prediction of vocal severity within and across voice types. *JSHR*, 30, 230-240.

Wolfe, V.; Cornell, R. & Palmer, C. (1991). Acoustic correlates of pathologic voice types. *JSHR*, 34, 534-543.

Wolfe, V.; Cornell, R. & Fitch, J. (1995). Sentence/ Vowel correlation in the evaluation of dysphonia. *Journal of Voice*, 3, 297-303.

Wolfe, V.; Fitch, J. & Cornell, R. (1995). Acoustic prediction of severity in commonly occurring voice problems. *JSHR*, 38, 273-279.

Wolfe, V. & Martin, D. (1997). Acoustic correlates of dysphonia: type and severity. *J. Commun. Disord.*, 30, 403-416.

Wolfe, V.; Fitch, J. & Martin, D. (1997). Acoustic measures of dysphonic severity across and within voice types. *Folia Phoniatr Logop*, 49: 292-299.

Wolfe, V.; Martin, D. & Palmer, C.I. (2000). Perception of dysphonic voice quality by naïve listeners. *JSLHR*, 43, 697-705.

Woo, P. (1996). The irritated larynx: edema and inflammation. In Wm.S. Brown, B.P. Vinson & M.A. Crary (Eds.), *Organic voice disorders, assessment and treatment* (pp. 119-170). San Diego: Singular Publishing Group.

World Health Organization (1980). ICDH. *International classification of impairment, disability and handicap.* Geneva, Switzerland: Author.

World Health Organization (1999). ICDH-2. *International classification of functioning, disability and health.* Beta-2 draft (full version). Geneva, Switzerland: Author.

Wuyts, F.L.; De Bodt, M.S.; Molenberghs, G.; Remacle, M.; Heylen, L.; Millet, B.; Van Lierde, K.; Raes, J. & Van de Heying, P.H. (2000). The dysphonia severity index: an objective measure of vocal quality based on a multiparameter approach. *JSLHR*, 43, 796-809.

Yamazawa, H. & Hollien, H. (1992). Speaking fundamental frequency patterns of Japanese women. *Phonetica*, 49, 128-140.

Yanagisawa, E.; Godley, F. & Muta, H. (1987). Selection of video cameras for stroboscopic videolaryngoscopy. *Ann Otol Rhinol Laryngol*, 96, 578-585.

Yiu, E.; Worrall, L.; Longland, J. & Mitchell, C. (2000). Analysing vocal quality of connected speech using Kay's computerized speech lab: a preliminary finding. *Clinical Linguistics & Phonetics*, 14, 295-305.

Yumoto, E.; Sasaki, Y.; Okamura, H. (1984). Harmonics-to-noise ratio and psychophysical measurement of the degree of hoarseness. *JSHR*, 27, 2-6.

Zawadzki, P. & Gilbert, H.R. (1989). Vowel fundamental frequency and articulatory position. *Journal of Phonetics*, 17, 159-166.

Zemlin, W. (1988). *Speech and hearing science anatomy and physiology.* (3rd edition). Englewood Cliffs: Prentice-Hall, Inc.

Chapter IX – Appendices

Appendix A – Pilot study

The pilot study

1 – Introduction

During a period of three and a half months (15 June to 1 October 1999) a pilot study was conducted to practice and develop a research design in order to get insight on Portuguese speakers in different voice conditions. It should be emphasised that this was an exploratory pilot study involving small groups of subjects, which was designed to determine the validity of the final research design and to identify and solve problems associated with recording procedures and data collection.

2 - Subjects

Thirty six subjects participated in this pilot study but the result was a set of thirteen subjects (or 36% of all subjects recruited) whose data could be used.

Table 1 – Subjects' demographic information

	Dysphonics (n=7) M ± SD*	Controls (n=6) M ± SD*
Age	46.7 ± 9.2	44.5 ± 12.83
Age range	34-64	34-67
% Females	57.1%	33.3%
% Males	42.9%	66.7%
Years of education	8.5 ± 3.99	5.7 ± 4.1
% Caucasian	85.7%	83.3%

* M ± SD (Mean ± Standard deviation)

As is evident from table 1, the mean age characteristics appear to represent middle age subjects, however this is not really the case because standard deviations are wide. It had been hoped that all the subjects in the study would have been under 65 years of age but unfortunately, considering the few cases referred to the out-patients ENT unit (during the period of the pilot study) one female had to be accepted aged more than 65.

Subjects were not specifically matched in a pairwise fashion (as might be seen by the percentage of genders in each group present in the table1).

Although all subjects speak Portuguese, two subjects were African. However, the majority of the subjects had been living in Lisbon for more than 25 years (except two subjects in the control group who live in Lisbon for 8 years only).

Table 1 shows that the mean years of education in both groups represent less than 9 years of study (less than elementary school). Nevertheless, two subjects in the control group had completed elementary school, and two others from the dysphonic and control group had completed university (more than 17 years of study).

The majority of females are not in paid employment with the exception of two subjects (an audiometry technician, and a cook). All men are employed but the majority are in different categories of work.

With the exception of two subjects in the dysphonic group who are professional voice users (a salesperson, subject 9 and a teacher, subject 21), all the other subjects in the sample are non-professional voice users. Only dysphonic subject 9 reported singing as a hobby for approximately 10 years.

Health history information gathered during the interview shows that the dysphonics reported a higher percentage of number of diseases (100%) than the control group (50%) and consequently a higher quantity of medication consumed per person. Within the dysphonic females the highest reported disease is allergy, and dysphonic males' highest reported diseases are gastroesophageal disease and high blood pressure.

Smoking habits shows some similarity between groups according to gender. 69% of all subjects do not smoke, and from those who do smoke, men reported a higher number of cigarettes a day than women (20 versus 1 to 9 cigarettes a day respectively).

85% of all subjects reported less than 0.5 to 1.0 litres of average daily water intake and drink coffee on a daily basis.

An inspection of the figure 1 reveals that ENT diagnoses for the dysphonics shows 42.8% of functional disorders, 28.5% of mass lesions, 14.2% of tissue changes and neurological disorder; and for the controls 66.6% present a healthy larynx and 33.3% minor larynx variations.

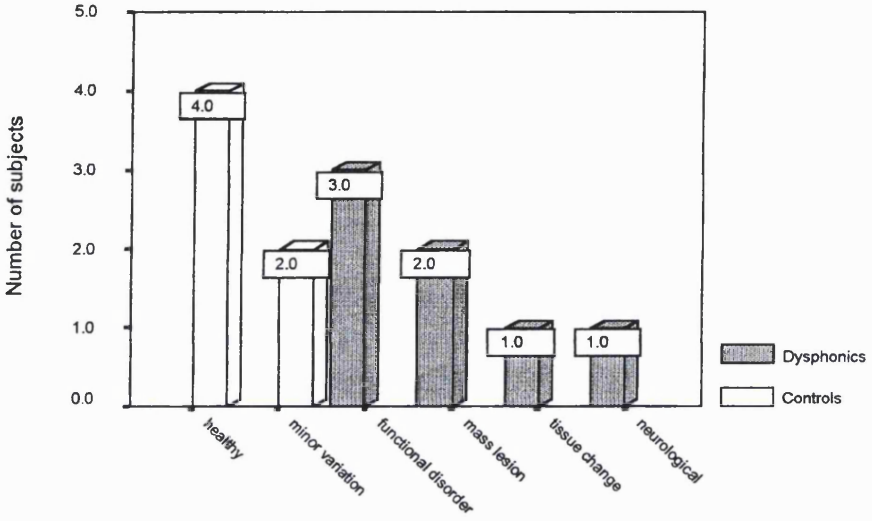


Figure 1 – ENT diagnoses

Also according to SLT voice diagnoses (figure 2) harsh voice quality is the most prevalent (42.8%) followed by breathy and hoarse with the same percentage (28.5%). The severity diagnoses indicate 57% of moderate dysphonia and 43% of slight dysphonia.

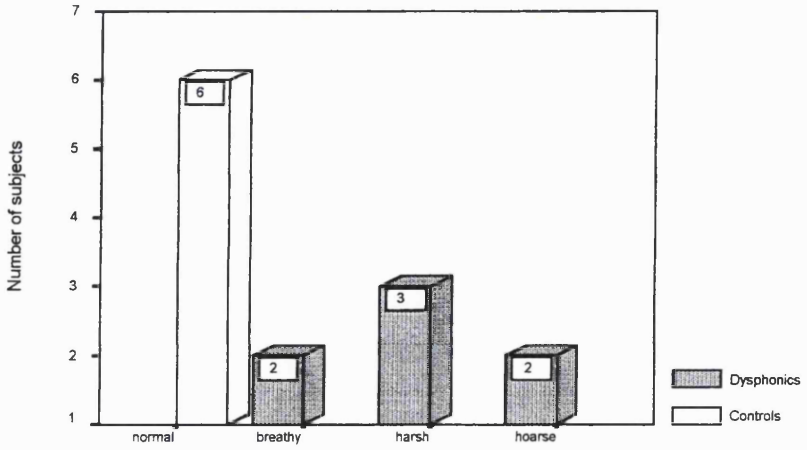


Figure 2 - SLT diagnoses

3 - Materials

All the materials described in the methodology (chapter V) were used in the pilot study with slight differences in the following: (1) the earliest SRRS form (Holmes & Rahe, 1967) was used here because it is one of the most commonly used measures of general stress in the literature, and it was already available in Portuguese in the Psychology Department of Lisbon University. (2) the rising and falling pitch contours on the vowel [a] was used. The rationale behind the use of it is that during the production of rising and falling pitch contours the vocal folds are forced to change in mass and stiffness through different settings. Laryngeal pathology may impede or limit the capacity to control the appropriate pitch contour of a speaker's voice (Fourcin & Abberton, 1971).

4 – Equipment and experimental environment

The laryngographic recordings were carried out in an audiometric cubicle in the Audiology department. Data collection used a DAT recorder (Sony TCD-D10), the Laryngograph processor (Laryngograph Ltd.), an electret condenser omnidirectional microphone (Knowles EK- 3132) and a portable oscilloscope connected to the main electric supply. The microphone was located below the speaker's mouth (like a tie clip), attached to velcro and grouped together with the leads from the Laryngograph.

5 - Experimental procedure

The procedures were the same as described in the methodology (chapter V) with the exception of the production of the rising and falling pitch contours with the vowel [a] after the sustained vowels. Each subject was asked to produce rising and falling pitch contours with the vowel [a] after a demonstration. If the productions were considered short or not precise the subject was asked to repeat them again either with or without a second demonstration. If the production was still considered to have been produced incorrectly after two attempts, no further attempts were requested.

6 - Data analyses

Data from the screening examinations (questionnaire, stroboscopy, auditory voice assessment) and the speakers'self-ratings was analysed subject by subject using the criteria described in the methodology (chapter V) and was edited in a SPSS (version 10) database.

In the acoustic and ELG analyses, first, each Sp and Lx waveform was examined for each task and several problems were identified in both signals. The Sp signal was overloaded in some subjects due to laughing, throat clearing or increases in speaking intensity in reading or conversation. Although recordings were made in a cubicle, background noise was detectable. During conversation the researcher's voice was detectable on the Sp signal but not, of course, on the Lx signal. Mostly Lx waveforms were contaminated by noise and in some cases the Lx amplitude was too weak, and no stable phonation could be identified for analysis, or the stable phonation portion was too short for reliable analysis. It was, therefore, decided to bandpass filter all Lx signals with a filter (Kemo type VBF8) between 10 Hz and 10 kHz in order to remove low as well as high frequency noise.

Although amplitude values can be calculated from Speech (Sp) data, for the purpose of this pilot study that was not considered because some of the general procedural rules were not achieved or considered during data collection. No attempt was made to calibrate the instruments before recordings had been done, and subjects' mouth and microphone distances were not constant. The analyses of isolated pitch contours were not used, considering the doubts about its naturalness. Therefore measures chosen for analyses were based on Lx data using all values in the original voice sample.

From the thirty six subjects two subjects were initially excluded from the pilot study: subject 25 was excluded due to technical problems (the Lx cable was mis-connected) and subject 17 was excluded because she was the only women referred who was only 18 years old and the majority of female subjects were aged more than thirty. Only 13 subjects (6 females and 7 males) produced data that was considered reliable. The 21 excluded subjects were 16 females and 6 males. From those, 14 subjects had laryngeal pathology ranging from laryngeal dysfunction, nodules, granuloma, arytenoid fixation, leukoplakia, and pachydermia.

The likely reasons for this are: (1) many of the recorded waveforms were quite seriously distorted possibly by the effects of the mains 'hum' due to the electrical connection of the equipment to the main supply or due to radio frequency pickup by electrodes (oscillation produced by other equipment in the hospital); (2) some of the recorded waveforms had insufficient amplitude to be considered a reliable Lx waveform possibly from an inaccurate placement of the neck electrodes or due to innapropriate recording control. It also can be due to subjects' anatomical characteristics (eg. is more difficult to obtain ELG waveforms in some female subjects) or pathological situations (eg. subject 7 had scar tissue after a tracheostomy, subject 3 had a large intubation granuloma and three other female subjects had longitudinal open glottis); (3) some speech-induced artifacts were found especially in the male subjects. Vertical movements of the larynx following throat clearing, swallowing and laughing caused large shifts of the Lx baseline and distorted the waveform.

7 - Statistical analyses

There are several limitations to the internal validity of this pilot study that make it impossible to attempt to find any statistical significance in the data. The selected sample was too small (N=13) and subjects were not matched in number, age and gender. Data was not obtained in ELG measurements for all tasks, especially vowels. Subjects were not randomly selected from a total dysphonic population and the pathologies included are not a representative sample from the universe of interest. For these reasons only descriptive statistics were used.

8 – Results and discussion

Although no statistically significant changes were found, due to the methodological problems previously described, the information gathered will be presented according to the questions in the study (see methodology). Also, as already mentioned, results for pitch contours and multiple regression analyses will not be presented due to the methodological problems and sparse data in the study respectively.

It must be stressed that at this stage the methodological inadequacies make it very difficult to generalise the results and to evaluate the specific nature of Portuguese pathological voice although interesting features appeared. Descriptive analyses will be

made along the presentation of the results; however, no normative, comparative and predictive discussion of the results will be attempted.

8.1 – Data for the criterion variables under diferent speaking conditions

8.1.1 – Fundamental frequency

An inspection of table 2 shows that the overall mean F0 for all speech tasks of the dysphonics, females and young adults is higher than that of the controls, males and older adults. Within the groups several trends can be noted, firstly the mean F0 of the vowels is higher than that of the reading and conversation. Secondly, vowel [i] present the highest mean F0 of the vowels according to group and age and vowel [u] has the highest mean F0 according to gender. Thirdly, the mean F0 of conversation is higher than that of the reading for all groups.

Table 2 – Means of F0 (in Hertzs) for all the speech tasks according to groups.

	Vowel [a]		Vowel [i]		Vowel [u]		Reading		Conversation
Group	N		N		N		N		N
Dysphonics	5	153.63	5	188.98	6	183.17	7	156.40	7	160.53
Controls	5	144.38	5	148.65	5	146.68	6	129.62	6	136.03
Gender										
Females	4	211.71	5	228.95	5	230.77	6	191.37	6	197.58
Males	6	107.19	5	108.68	6	113.09	7	103.47	7	107.77
Age										
Young adults	3	183.24	3	194.04	3	192.36	4	151.10	4	158.79
Older adults	7	134.33	7	158.00	8	156.92	9	140.90	9	144.96

However, an individual inspection of the mean values according to groups (figure 3 for dysphonics and figure 4 for controls) shows intersubject variability. In fact it is clear from figure 3 that: (1) although the group data shows a higher mean F0 during the sustained vowels than connected speech, subject 4 presents a higher mean F0 during conversation than that of the vowel [a]. Subject 4 is a 64 years old female with laryngeal dysfunction; (2) vowel [u] did not show the highest mean in the subject 23 for whom the highest mean F0 vowel was [i]; (3) the mean F0 of conversation is lower than that of reading for the dysphonics (9, 23 and 24). All speakers have 4 years of education and their professions are salesperson, housewife and cook, respectively; (4) within the dysphonics, females (4, 6, 23 and 24) show wider variability than males (9, 21 and 27); (5) within the males, subject 9 shows a slightly lower mean F0 than the subjects 21 and

27. Subject 9 is older (50 years old) than the other two (44 years old both). Also, subject 9 has vocal fold nodules whereas subjects 21 and 27 present minor larynx variations.

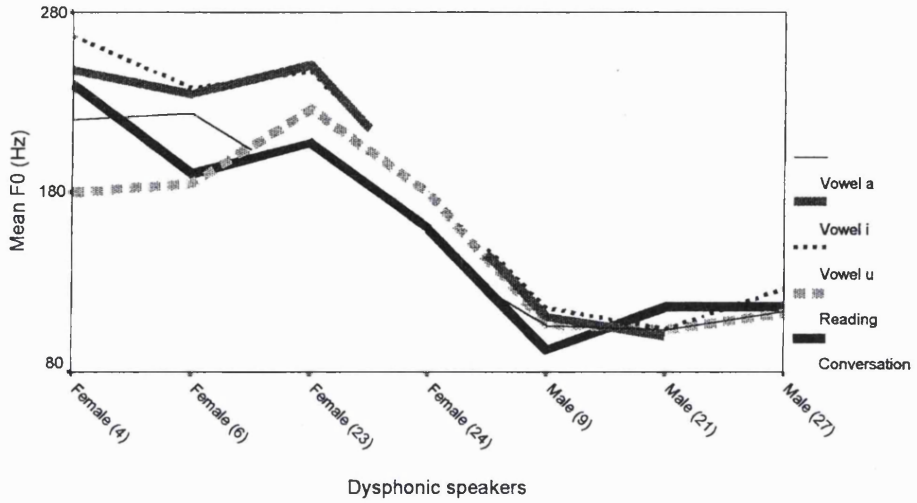


Figure 3 – Mean F0 for the dysphonic speakers

It is clear from figure 4 that: (1) data for the sustained vowels is non-existent for some speakers; (2) females (2 and 22) show a higher mean F0 than males (1, 14, 28 and 32); (3) speakers 1 and 14 show a lower mean F0 than the speakers 28 and 32. This fact may be related to age (the first two subjects are in their 48 years old and the other two are in their 34 and 36 years old). Moreover, the first two subjects are ex-smoker and smoker respectively and the following two had never smoked.

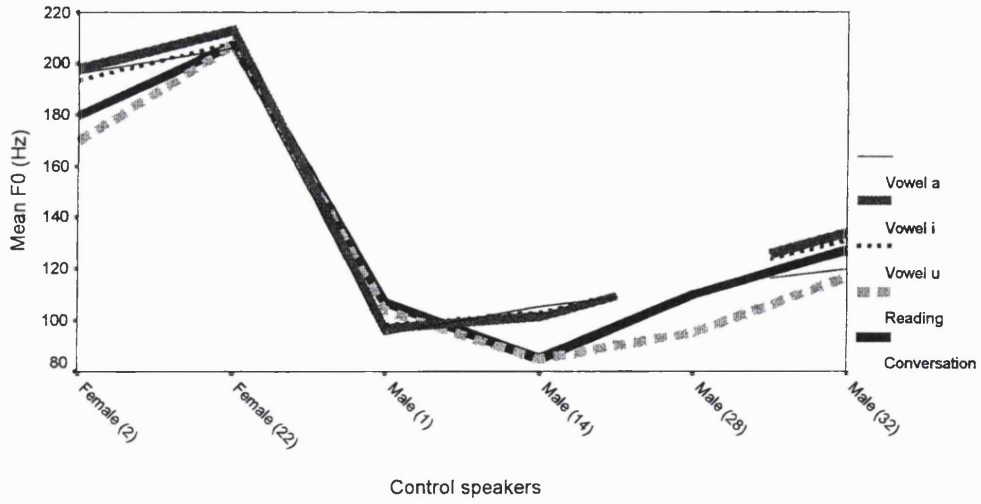


Figure 4 – Mean F0 for the control speakers

Although as a group data the mean F0 of the vowels is higher than the connected speech (table 2) as already been mentioned intersubject variability exists. For example subject 1 shows lower mean F0 for all the vowels than connected speech, subject 32 shows lower mean F0 for the vowel [a] than conversation, subject 9, 21 and 27 show similar mean F0 during the sustained vowel [a] and reading. The mean F0 of the vowel [u] is not always the highest mean vowel (e.g. in subjects 14 and 32). The mean F0 during conversation is lower than during reading for subject 9 and similar for subject 14.

8.1.2 – Fundamental frequency standard deviation

It is clear from table 3 that the overall mean F0 standard deviation of the dysphonics and females is higher than that of the controls and males with an exception for conversation (were the values are similar). According to age there is no robust trend, for example the younger adults show higher means for vowel [a] and vowel [i] than that of the older adults whilst for the other speech tasks is the other way round. Moreover, vowel [u] shows the highest mean F0 standard deviation for the dysphonics, females and older adults. Conversation presents higher values than reading for all groups. These findings raise the question as to whether age related mean F0 discrepancies are explained on the basis of the sample bias.

Table 3 – Means of F0 standard deviation (Hz) according to groups.

	Vowel [a]		Vowel [i]		Vowel [u]		Reading		Conversation
Group	N		N		N		N		N	
Dysphonics	5	3.2	5	4.3	6	4.5	7	.08	7	.11
Controls	5	3.0	5	2.6	5	2.4	6	.07	6	.10
Gender										
Females	4	5.2	5	5.6	5	5.5	6	.08	6	.10
Males	6	1.7	5	1.2	6	1.9	7	.07	7	.10
Age										
Young adults	3	4.0	3	4.5	3	3.2	4	.06	4	.09
Older adults	7	2.7	7	3.0	8	3.6	9	.08	9	.11

Inspection of the speakers F0 standard deviation distribution (figure 5 and 6) confirms that intersubject variability is greater for the sustained vowels than connected speech and also greater for the females than the males.

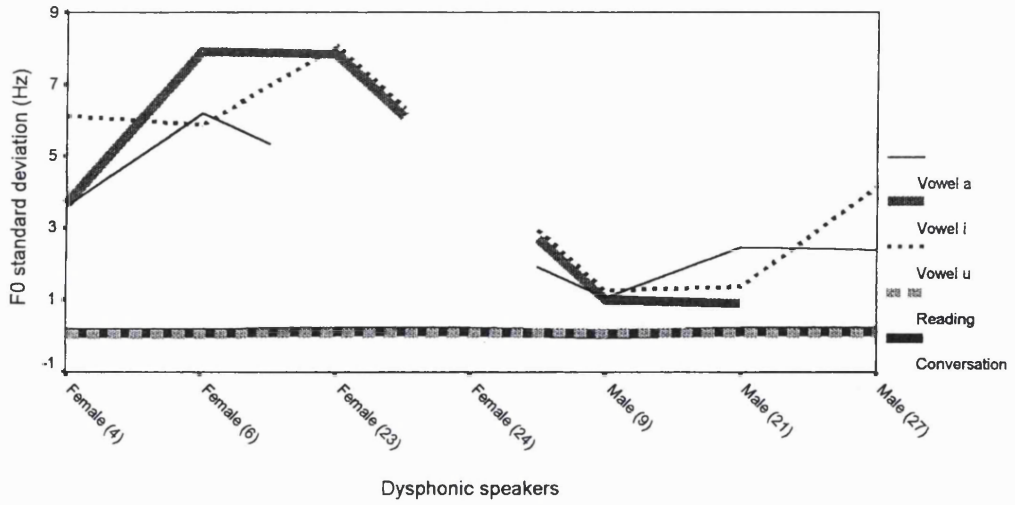


Figure 5 – Mean F0 standard deviation for the dysphonic speakers

It is clear from the figure 5 that data was not obtained for all dysphonic speakers and also for a control speaker as might be seen in figure 6.

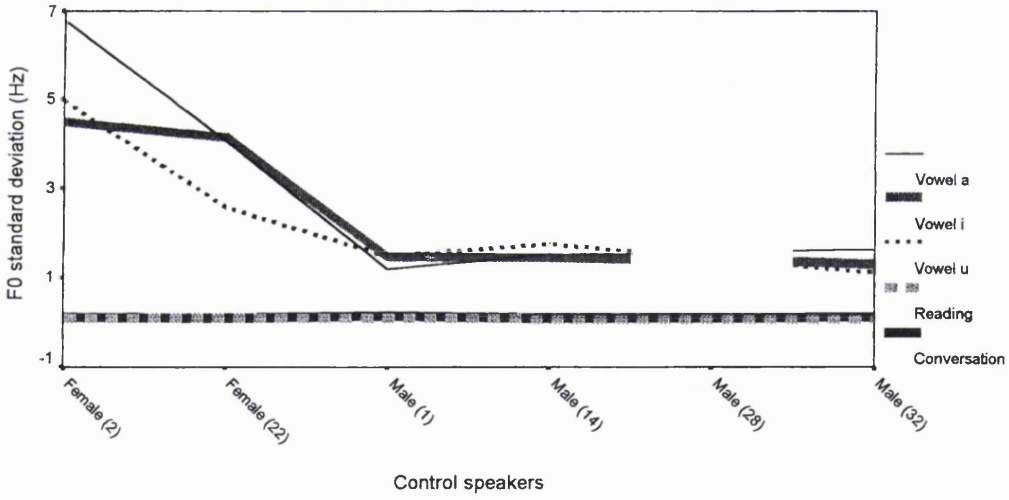


Figure 6 – Mean F0 standard deviation for the control speakers

8.1.3 – 90% Fundamental frequency range

An examination of table 4 shows several trends for the 90% mean of F0 ranges during reading and conversation. Firstly, dysphonics, males and older adults show wider 90% ranges than that of the controls, females and young adults. Secondly, both ranges are wider during conversation than reading.

Table 4 – 90% means of F0 range (in octaves).

			Reading	Conversation
Group		N		
	Dysphonics	7	.71 ± .15*	.96 ± .26*
	Controls	6	.58 ± .10	.94 ± .31
Gender				
	Females	6	.64 ± .14	.86 ± .16
	Males	7	.66 ± .16	1.0 ± .33
Age				
	Young adults	4	.51 ± .09	.82 ± .25
	Older adults	9	.72 ± .14	1.0 ± .27

*M (mean) ± SD (standard deviation)

Nevertheless, as shown in figure 7 and 8 intersubject variability exists: (1) for example although conversation presents wider range than reading there are two speakers that show the reverse, subject 9 (figure 7) and subject 22 (figure 8). Subject 9 is 50 years old, works as a salesperson and sings in a choir for ten years. One year ago started with intermittent voice complaints. The ENT diagnosis is kissing nodules and the SLT diagnosis, moderate harsh voice. Subject 22 is a 32 years old female without voice complaints; (2) subject 21 (figure 7) and subject 1 (figure 8) show the highest F0 range during conversation. Subject 21 is 44 years old and he is a physical teacher with intermittent voice complaints. Subject 1 is 48 years old and although without voice complaints the ENT diagnoses is 'arytenoids arepartially pulled forward during phonation'.

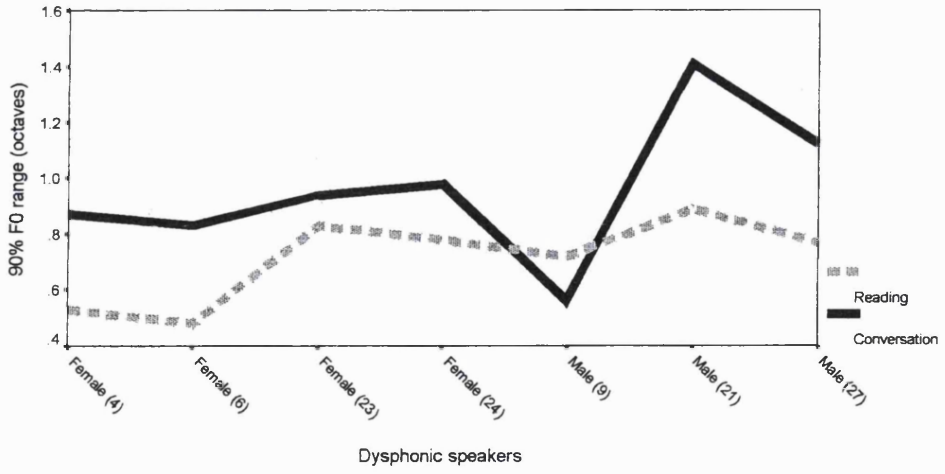


Figure 7 – 90% F0 range for the dysphonic speakers

As might be seen in figures 7 and 8 the intersubject variability is more evident during conversation than reading for both groups and the likely reason for that may be the fact that conversation is different from subject to subject which does not occur in reading (Baken & Orlikoff, 2000). Another possible explanation may be the speakers' communicative attitude.

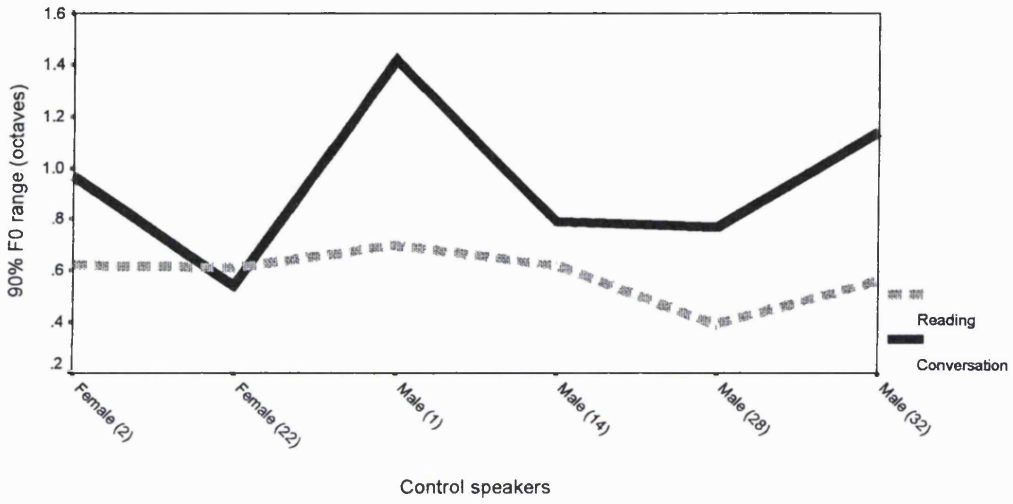


Figure 8 – 90% F0 range for the control speakers

Additionally an inspection of the DFx and CFx distribution of the subject 28 (as might be seen in the figure 9) corroborates the hypothesis of 'communication attitude' for a restricted F0 range once this speaker had a negative attitude towards reading aloud (as was referred in the limitations of the pilot study).

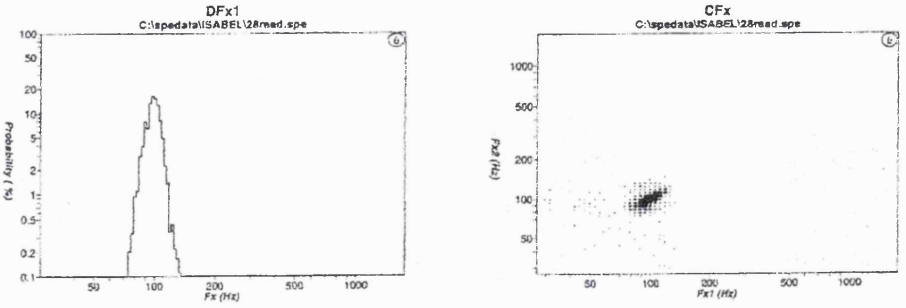


Figure 9 – DFx and CFx distribution during reading for speaker 28

Nevertheless, Leff & Abberton (1981) found a restricted frequency range with a high degree of 'peakedness' of the histogram related to monotonous voice in depression

and schizophrenia. From the history data is known that this speaker has a child with a serious disease (a syndrome) that obliged them to live in the capital far way from their family and friends.

8.1.4 – Percentage of jitter

The overall mean jitter presented in table 5 for all vowels of the females and young adults is higher than that of the males and older adults. Also, controls show higher values for the vowels [a] and [i] than the dysphonics. Moreover, vowel [a] presents the higher jitter value for the controls, females and older adults whilst vowel [i] presents the higher jitter value for the dysphonics and young adults.

Table 5 – Jitter (%) for all vowels according to groups

		Vowel [a]		Vowel [i]		Vowel [u]
Group	N		N		N	
Dysphonics	5	1.4 ± .94*	5	1.4 ± 1.3*	6	1.4 ± .84*
Controls	5	1.9 ± .69	5	1.6 ± .48	5	1.3 ± .47
Gender						
Females	4	2.1 ± .89	5	1.9 ± 1.1	5	1.5 ± .88
Males	6	1.3 ± .61	5	1.1 ± .51	6	1.2 ± .48
Age						
Young adults	3	1.9 ± 1.1	3	2.2 ± 1.45	3	1.5 ± 1.0
Older adults	7	1.5 ± .71	7	1.2 ± .49	8	1.3 ± .56

* M (mean) ± SD (standard deviation)

Figure 10 and 11 shows the intersubject variability for the dysphonics and controls respectively.

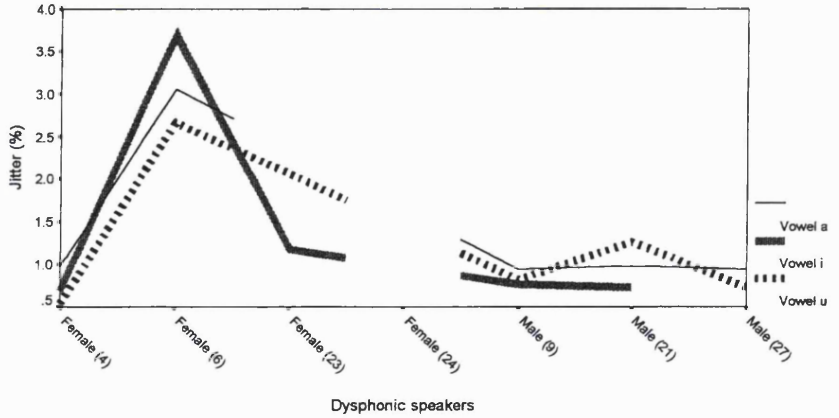


Figure 10 – % jitter for the dysphonic speakers

Although as a group data all vowels present similar jitter values, figure 10 shows a high intersubject variability and also some missing data for dysphonics. A similar

intersubject variability was found in the controls (figure 11). In the dysphonics subject 6 shows the highest percentage of jitter for all vowels. She is 34 years old and shows right vocal fold immobility in paramedian position with posterior open glottis. In the controls (figure 11) subjects 1, 2 and 22 show the highest percentage of jitter for the vowel [a] whereas subjects 14 and 22 is for the vowel [i] and subject 14 for the vowel [u]. Subject 14 is 48 years old, smokes 20 cigarettes a day and does not present voice complaints.

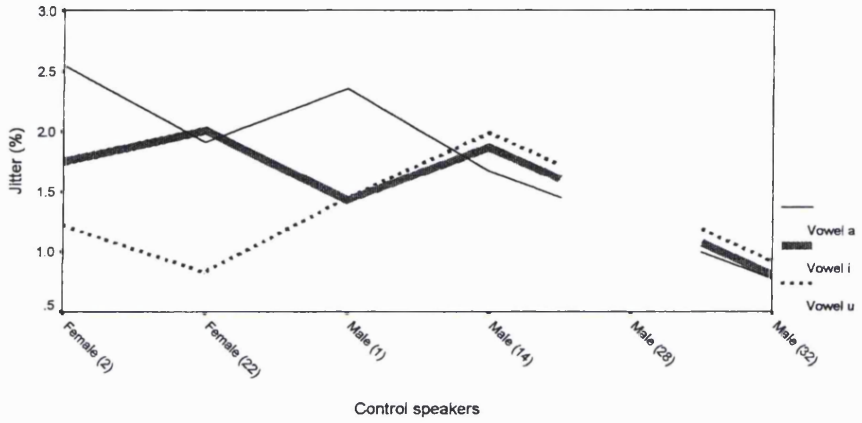


Figure 11 – % jitter for the control speakers

8.1.5 – Percentage of irregularity

Table 6 shows that the percentage of irregularity of the dysphonics, males and older adults is higher for both speech tasks than that of the controls, females and young adults. Also, conversation shows higher percentage of irregularity than reading.

Table 6 – Percentage of irregularity

			Reading	Conversation
Group		N		
	Dysphonics	7	13.5 ± 9.9*	17.6 ± 15.9*
	Controls	6	8.7 ± 2.5	14 ± 7.1
Gender				
	Females	6	9.5 ± 3.3	13.3 ± 6.6
	Males	7	12.2 ± 9.4	17.6 ± 14.9
Age				
	Young adults	4	8.2 ± 4.1	10.1 ± 4.4
	Older adults	9	12.2 ± 8	18.1 ± 13.1

* M (mean) ± SD (standard deviation)

Although as group data conversation shows higher percentage means of irregularity individually (in the dysphonics) subjects 24 and 21 (figure 12) show the highest percentage means of irregularity for both speech tasks whereas subject 9 show the lowest percentage of irregularity. Subject 21 is a physical teacher with regular pharyngitis and intermittent voice complaints. The ENT diagnosis is 'slightly assymetrical vibration'. Subject 24 is 49 years old, she is a cook, started the pre-menopause two months ago and the ENT diagnosis is 'right vocal fold nodule'.

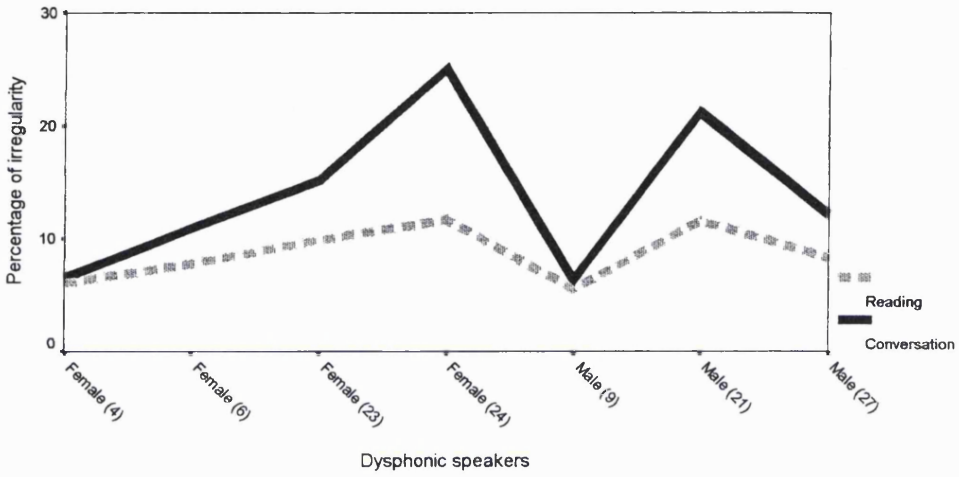


Figure 12 – Percentage of irregularity for the dysphonic speakers

In the controls (figure 13) subject 14 presents the highest percentage mean of irregularity for both speech tasks. Subject 14 is African and smokes 20 cigarettes a day for more than 20 years.

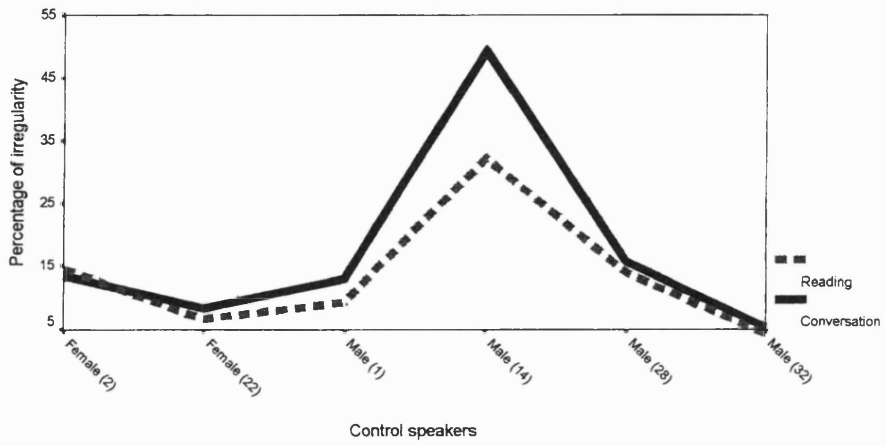
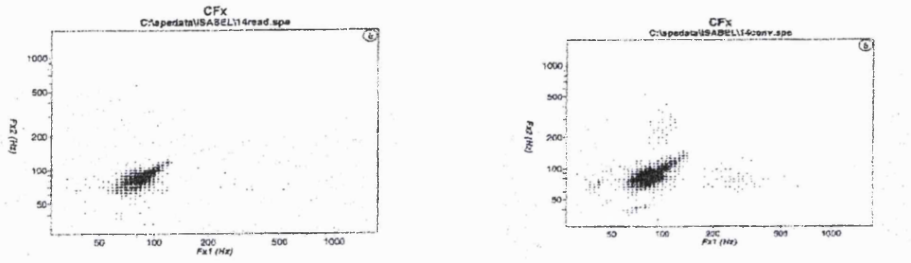


Figure 13 – Percentage of irregularity for the control speakers

Additionally an inspection of the DFx and CFx distribution of the subject 14 (as might be seen in the figure 14) shows that is similar to the shape for creaky voice (widening of the plot a low frequencies) as presented by Abberton & Fourcin (1984) and Carlson (1995a)



During reading

During conversation

Figure 14 – CFx distributions for the speaker 14

8.1.6 – Mean QFx

It is clear from table 7 that, firstly the mean contact closed quotient of the males and older adults is higher than that of the females and young adults. Secondly, no specific trend was found for dysphonics versus controls and for the different speech tasks.

Table 7 – The mean QFx (%) according to groups

		Vowel [a]		Vowel [i]		Vowel [u]		Reading	Conversation
Group	N		N		N		N		
Dysphonics	5	42.7	5	41.7	6	44.9	7	42.4	41.6
Controls	5	42.1	5	43.3	5	42.2	6	41.8	41
Gender									
Females	4	38.5	5	38.9	5	37.4	6	37	36.8
Males	6	44.9	5	46.1	6	49	7	46.5	45.2
Age									
Young adults	3	37	3	38	3	37	4	38.5	39
Older adults	7	44.7	7	44.4	8	46.3	9	43.7	42.4

Intersubject variability shown in the figure 15 and 16 is more evident in the controls than dysphonics however a wider mean contact quotient range is shown in the dysphonics (figure 15).

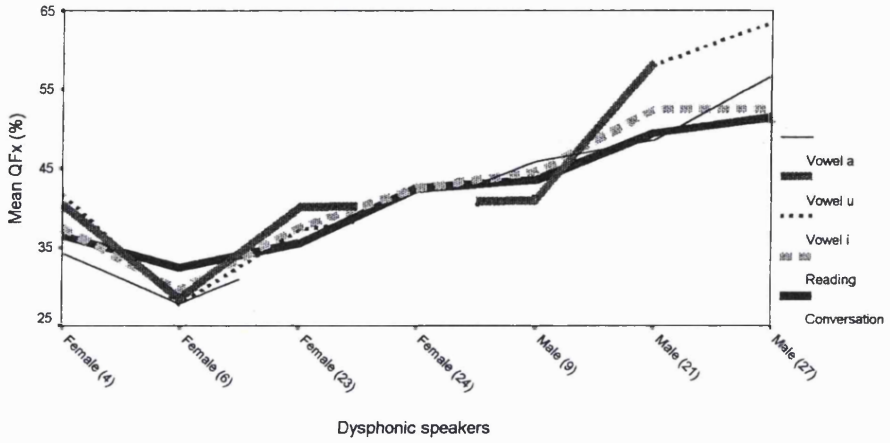


Figure 15 – Mean contact quotient for the dysphonic speakers

Overall, subject 6 show the lowest mean QFx for all speech tasks and the subject 27 the highest means. It seems obvious once subject 6 presents 'right vocal fold immobility in paramedian position with posterior glottal chink' and subject 27 presents laryngeal dysfunction.

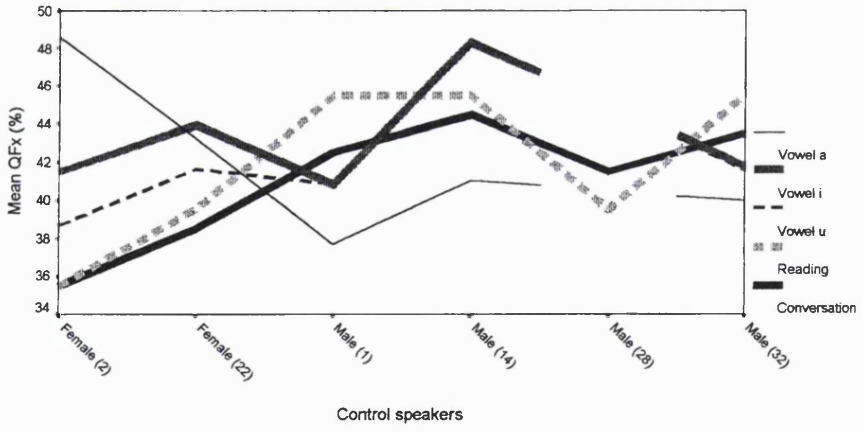


Figure 16 – Mean contact quotient for the control speakers

In the controls (figure 16) subject 14 shows the highest mean QFx for the vowel [i], reading and conversation whereas subject 1 shows the lowest mean QFx for the vowels [a] and [i].

8.1.7 – QFx standard deviation

The mean contact closed quotient standard deviation for all speech tasks for both groups (table 8) shows that: (1) the dysphonics present higher standard deviation values than the controls for almost all speech tasks except conversation; (2) females present higher standard deviation values than males for the vowel [a] and reading whereas for conversation both genders present the same mean values; (3) young adults show higher values than the old adults only for the vowels [a] and [i].

Table 8 – The mean QFx standard deviation (%) for all the speech tasks according to groups.

	Vowel [a]		Vowel [i]		Vowel [u]		Reading		Conversation
Group	N		N		N		N		N	
Dysphonics	5	.17	5	.09	6	.08	7	.08	7	.10
Controls	5	.12	5	.06	5	.04	6	.07	6	.10
Gender										
Females	4	.23	5	.07	5	.03	6	.08	6	.10
Males	6	.10	5	.08	6	.08	7	.07	7	.10
Age										
Young adults	3	.17	3	.08	3	.05	4	.06	4	.09
Older adults	7	.14	7	.07	8	.06	9	.08	9	.11

An individual inspection of the data shows that, overall, the QFx standard deviation intersubject variability is: (1) within the vowels, higher for the vowel [a] than for the vowel [i] followed by the vowel [u]; (2) within the connected speech, higher for conversation than reading.

8.2 – Data for the subjects' self-rated psychosocial voice impact

The mean Voice Handicap Index (VHI) for both groups presented in V.9 shows that the mean overall of the dysphonics, females and older adults is higher than that of the controls, males and young adults. The fact that dysphonics present higher scores of psychosocial voice impact is not surprising due to the aim of the type of measurement.

Table V.9 – The mean VHI scores according to groups.

		Total
Group	N	
Dysphonics	7	16.29 ± 7.7*
Controls	6	3.2 ± 3.6
Gender		
Females	6	14.3 ± 7
Males	7	6.7 ± 9.5
Age		
Young adults	4	8.8 ± 10.4
Older adults	9	10.9 ± 9

* M (mean) ± SD (standard deviation)

Intersubject variability in the results of the VHI is shown in figure 19 for the dysphonics and figure 20 for controls.

Overall, within the dysphonics (figure 19) the physical subscale shows higher scores than the other two subscales (emotional and functional). The only exception is in speaker 27 where the functional score is slightly higher than the physical one. Nevertheless, the scores for this speaker are among the lowest ones in the dysphonics, which does not seem surprising. In fact, speaker 27 is 44 years old, is a non vocal professional (gardener), has been smoking 20 cigarettes a day for more than 30 years and has a dysphonia for more than 10 years. The ENT examination for this speaker showed minor larynx abnormalities (dimished vocal fold amplitude and asymmetrical vibration).

Subjects 6 (female) and 9 (male) present the highest scores for the total and subtotal VHI scores (except in the emotional subscale score for subject 9). This result for

subject 6 seems obvious once she had been submitted to an aesthetical thyroidectomy and ended up with a larynx pathology (right vocal fold immobility in paramedian position). This condition resulted in unemployment once she was not able to talk in the noisy working environment as a barman.

Subject 9 was worried with his voice problem mainly because of its handicap in his voice hobbie performance (in the singing choir).

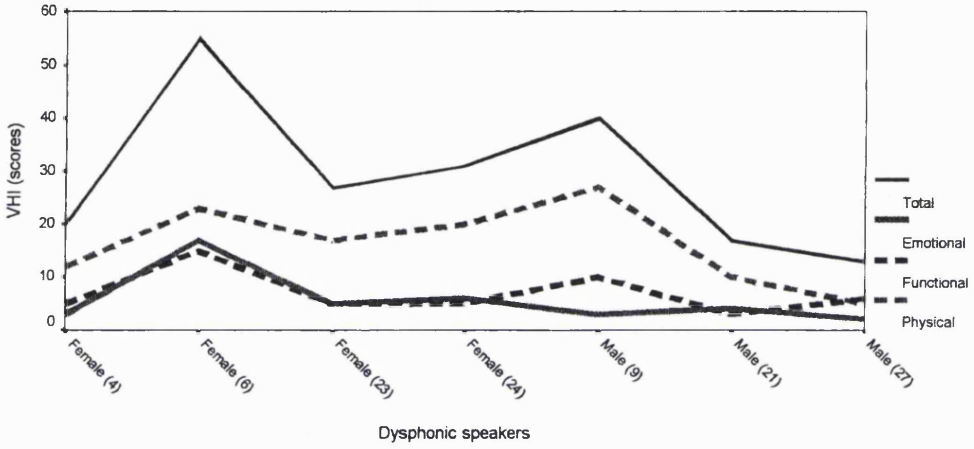


Figure 17 – VHI for the dysphonic speakers

No specific trend was found in the controls for the VHI scores (figure 20) although subject 22 showed the highest total and emotional scores. Nevertheless, the total maximum score (9) is not significant according to the total possible score (120).

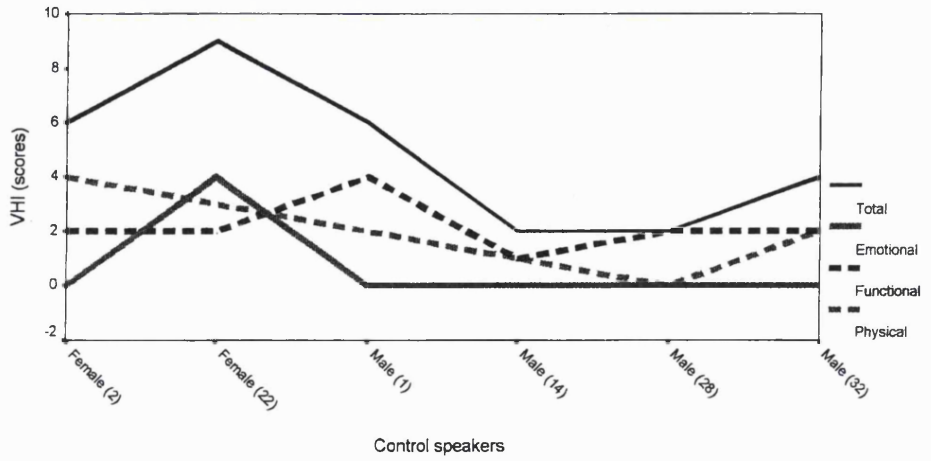


Figure 18 – VHI for the control speakers

8.3 – Limitations encountered in the pilot study

Several limitations were experienced in (1) sample gathering; (2) materials used and (3) environment and equipment.

Firstly, although the original intention was to have groups of equal size matched by age and gender, this was not possible in practice. It is important to mention that for this pilot study no dedicated sessions were pre-organised within the ENT unit for stroboscopy assessment. As a result, 'healthy subjects' appointments turned out to be more difficult to incorporate into the workload of the ENT surgeons. Also, another constraint was the time period and the season during which the pilot study took place. As it progressed the patients' failure to attend or cancellation were due to summer holidays. And, in addition, recruitment of 'healthy' subjects was found to be difficult (especially men) because of the submission to stroboscopy. Also, after stroboscopic examination some speakers complained about discomfort in the anaesthetised nostril and sneezing.

Secondly, caveats in the research design emerged in sustained vowels, pitch contours, reading and in the subjects' self-rated scales.

Data analyses of sustained vowels appeared to show some order effect in the quality of the obtained data.

During pitch contour production almost all subjects found it strange and asked for the demonstration to be repeated and although they were not confident, they mostly asked for confirmation "is it like this?". For that reason pitch contour production was not considered for the pilot study analysis.

Few subjects presented with reading difficulties (hesitations, part-word segmentations and repetitions and dysrhythmic flow of reading) and a male from the control group (subject 28) had a negative reaction reading aloud. Nevertheless, the inclusion of the reading data from three subjects with reading difficulties and another with a negative attitude towards reading was considered preferable to reducing the sample size (especially in the female controls who would be reduced to one subject).

The Voice Handicap Index questionnaire performed well in the subjects studied. However, a consistent difficulty was found in the sentence 'P13 – my voice is creaky and dry' because the translation used the perceptual meaning of the word 'creaky' ('estaladiça'). A synonym based on the physical sensation ('àspera') was given to the

subjects with success. Also, one subject commented on the ambiguity of the double statement 'creaky and dry'.

On the Social Readjustment Rating Scale it appeared from the pilot study that there were several content and phrasing simplicity limitations. The originally published SRRS used abbreviated wording (to save space) for the life events items. It also uses items referring to mortgage or loan greater than $ 10,000 or less than $ 10,000. The SRRS in use today has undergone only these slight modifications, more complete wording and 'mortgage or loan for a major purchase' or 'for a lesser purchase' Holmes and David (1989:39). Problems about the Social Readjustment Rating Scale (SRRS) were especially about the question 'Wife begins or stops work'. The Portuguese translation had the word 'Mulher' (Wife) that might mean wife or woman in Portuguese, instead of 'Esposa' (Spouse/ Wife). Subject 6 asked if they should mark it because she was unemployed. The version used was the original one with short sentences and without examples. From oral feedback it was felt that sometimes the subjects tried to confirm what was intended with some sentences as for instance 'minor violations of the law'.

It is also worth noting, however, that this pilot study was the first attempt at a multidisciplinary study of Portuguese voice quality. For this reason some instruments had to be created, translated and adapted to Portuguese especially those related to voice outcomes for dysphonia.

Thirdly, several other limitations were experienced during the procedures with the environment and equipment. The experimental environment (cubicle room in the Audiology unit) was not acoustically immune to noise (some background noises, e.g. telephone ringing, voices, door shutting, were heard inside the cubicle) and was too small for two persons and the equipment did not allow for face-to-face interaction; however, this was the best place in the ENT department for ELG data gathering. The same environment was used during the 'main study' regardless of the noise. However, acoustic data gathered was not considered for analyses, only Lx data that is immune to ambient noise.

The recording equipment was connected to the mains supply. Because, during data analysis, spurious noise was found in almost all the data, in the 'main study' the equipment was battery powered.

The microphone used was a tie clip attached to velcro and grouped with the leads from the Laryngograph at the level of the speaker's chest. Although an explanation was given before the recordings about the importance of not moving the head and the arms, this

was very difficult to achieve (most of them, naturally, moved the arms while talking and some interference with the electrode leads was seen).

8.4 – Suggestions for the main study

Due to the limitations experienced in the pilot study several minor modifications were performed for the main study: (1) the time period for data collection was extended for 8 months (from January to September 2000, without considering the summer holiday month); (2) sessions were programmed in the ENT unit agenda specifically for this study; (3) although stroboscopy examination was one of the major constraints for 'healthy subjects' data gathering it was decided to maintain this criterion for all subjects in the 'main study'; (4) all the recording equipment was battery powered; (5) the microphone was changed to a head-mounted type because the distance from the subject's mouth is more fixed as the subject moves the head. Also, the interference with leads by the natural moving of the subject's arms during conversation is not so likely because the leads are fixed by a clip over her/his shoulder; (6) pitch contours were omitted from the 'main study' for the reasons explained in the experimental procedure; (7) sustained sequence of vowels [i]; [a]; [i]; [u] instead of only the last three to avoid an order effect. The first vowel was not considered for data analyses; (8) VHI scale uses the new tested word 'àspera' instead of 'estaladiça' for 'creaky'; (9) a full version of the SRRS scale was used instead of the simplest version (appendix F); (10) only Lx data was used for analyses considering that acoustic recordings were inappropriate because of environment noise.

8.5 – Conclusions

The original purpose of this pilot study was to practise ELG recording and data collection to develop a cross sectional design in voice study. The main aim of the research design was to investigate the quality of Portuguese pathological voice using electrolaryngographic (ELG) measurement during the production of sustained vowels, pitch contours, reading and conversation. As has already been mentioned pitch contours analysis was not attempted. Other aims were to investigate the subjects' self-rating degree of psychosocial consequences of voice disorder using the Voice Handicap Index (VHI). And a final purpose was to investigate which variables contribute most

importantly to predict the speech tasks data results and subjects' self-reported voice handicap index (with was not attempted for the reasons already explained).

Although the limitations meant that it was impossible to have a group design result with statistical significance, the experience gained on ELG data gathering and analyses was valuable and represented a significant improvement in the main study (95% of the subjects' data is reliable) compared to this pilot study (36% of the subjects' data is reliable). Also, the results obtained represented an important step towards knowledge about Portuguese pathological voice. Hence, one needs to be cautious about the data and differences found compared with previous studies.

Appendix B – Consent form

(English and Portuguese version)

CONFIDENTIAL

Subject number

Informed consent

(Name)

I declare that I have been informed of the purpose of this study in which I agree to participate. I also accept the methods to be used, knowing that they will not harm my health.
Moreover, I authorise the recording of my voice, as well as the use of the data obtained for scientific and educational purposes.

Lisbon,
Signature

Study title – An electrolaryngographic study of dysphonic Portuguese speakers.
Principal investigator -Isabel Guimarães
Supervisor- Evelyn Abberton

UNIVERSITY COLLEGE LONDON - Department of Phonetics and Linguistics
Wolfson House - 4 Stephenson Way - London NW1 2HE - Tel: (0171) 380 7401 - Fax: (0171) 383 0752

CONFIDENCIAL

Sujeito número:

Consentimento Informado

(Nome)

Declaro ter sido informado dos objectivos do estudo em que aceito participar. Aceito igualmente os métodos utilizados, sabendo que não prejudicam a minha saúde.
Autorizo pois, a realização das gravações bem como a utilização dos dados obtidos para efeitos científicos e educacionais.

Lisboa,
Assinatura

Título do estudo – Estudo electrolaringográfico de indivíduos portugueses com disfonia.
Investigador principal - Isabel Guimarães
Supervisor – Evelyn Abberton

UNIVERSITY COLLEGE LONDON - Department of Phonetics and Linguistics
Wolfson House - 4 Stephenson Way - London NW1 2HE - Tel: (0171) 380 7401 - Fax: (0171) 383 0752

Appendix C – ENT rating form

(English and Portuguese version)

ENT examination – Larynx-stroboscopy

Name:_____ Age:_____ Date:__/__/__ ENT Surgeon:_____
Clinical history:_____
Personal history:_____
Family history:_____

ENT OBSERVATION
Nose:_____
Rhinopharynx:_____
Hypopharynx:_____
Larynx:_____
Epiglottis:_____ A.E. Folds:_____
Arytenoids:_____ Com. Post.:_____
Right vocal fold:_____
Left vocal fold :_____
Glottic gap:_____
Ventricular Folds:_____
Sub-Glottis:_____
STROBOSCOPY:
F0=_____ Normal vibrations:_____ Asymmetrical:_____ Symmetrical:_____
Glottic closure: Complete:_____ Incomplete:_____
Regularity: Regular:_____ Irregular:_____

	Right Vocal Fold	Left Vocal fold
Amplitude		
Increased:		
Normal:		
Dimished:		
Absent:		
Mucosal wave		
Increased:		
Normal:		
Dimished:		
Absent:		
Non vibrating portion		
None:		
Partial:		
Total:		
Occasional		
Always:		

Subjective assessment:
Normal voice:_____ Dysphonic: _____
Comments:_____

Avaliação O.R.L. – Laringo-estroboscopia

Nome:_____Idade_____Data:__/__/__Médico:_____
História Clínica:_____
Antecedentes Pessoais:_____
Antecedentes Familiares:_____

OBSERVAÇÃO O.R.L.
Fossas Nasais:_____
Rinofaringe:_____
Hipofaringe:_____
Laringe:_____
Epiglote:_____Pregas A.E.:_____
Aritenoides:_____Com. Post.:_____
Corda Vocal Drtª:_____
Corda Vocal Esqª:_____
Escape Glótico:_____
Bandas Ventriculares:_____
Sub-Glote:_____
ESTROBOSCOPIA:
F0=_____Vibrações normais:_____Assimétricas:_____Simétricas:_____
Encerramento glótico: Completo:_____Incompleto:_____
Regularidade: Regular:_____Irregular:_____

	Corda Vocal Drtª	Corda Vocal Esqª
Amplitude		
Aumentada:		
Normal:		
Diminuida:		
Ausente:		
Onda da Mucosa		
Aumentada:		
Normal:		
Diminuida:		
Ausente:		
Porção não vibrátil		
Nenhuma:		
Parcial:		
Total:		
Ocasional		
Sempre:		

Avaliação subjectiva:
Voz normal:_____Disfónica: _____
Comentários:_____

Appendix D – Questionnaire

Questionnaire

Date: Subject n°:

Identification

Name: Male/Female:

Date of birth: Age: Place of birth:

Address :

Education: Occupation:

Social history

Who do you live with? (For how long?)

Do you have children? What are their ages?

When you are not working, what do you enjoy doing?

Do you play any sport regularly (if yes specify)

Do you play any wind instrument? Which? How often?

Are you involved in any hobbies or activities where you are in contact with dust, fumes, chemicals or paints?

General health history

Do you have or have had previous major problem/illness?

(Heart disease; High blood pressure; Diabetes; Gastrointestinal problems; Epilepsy, Asthma; Respiratory disease (emphysema; bronchiolitis); Allergies; Sinusitis; Rhinitis; Cancer; Pancreatic dysfunction; Hypothyroidism; Hyperthyroidism; Rheumatic fever; Kidney disease; Psychiatric problems, Psychological problems; Depression; Alcoholism; Anorexia; Bulimia; Drug abuse)

(if yes, when was it; what was the length of each illness, extent of incapacity, types and effects of treatments, intubation, medication)

Do you have or have had any other salient medical problems?

Are you undergoing any surgery?

(what?, when? was anaesthetic/intubation/tracheostomy required? Where there any surgical complications (e.g. recurrent laryngeal nerve section, airway obstruction, hemangioma)

Do you have or have had any serious accident or injury?

(none, occurred with head or neck injury, loss of consciousness, or whiplash; occurred without head or neck injury, loss of consciousness, or whiplash; foreign body aspiration (e.g., choking on foods or small objects); ingestion of caustic agents (e.g., acids, alkalines); inhalation injuries or burns; iatrogenic injuries (from irradiation, prolonged intubation, or tracheostomy)

Have you had x-ray treatments to your head or neck?

List all current prescribed medication and doses

(Aspirin; Vitamins; Cold tablets; Sore throat (spray, tablets), Antihistamines; Nasal steroid spray; Oral sprays, Bronchial dilator; Mucolitics; Sleeping pills; Anxiolytics; Antidepressives; Antacids; Antibiotic, Hypertension medication; Laxatives; Diuretics; Others).

For females

At what age did you reach puberty?

Are you pregnant?

Are your menstrual periods regular?

What is your menstrual cycle at the moment?

Do you take oral contraceptives? (since when)

Have you gone through the menopause? (since when, prescribed medication)

Have you undergone hysterectomy? (when, prescribed medication)

Where your ovaries removed? (when, prescribed medication, HRT (hormonal replace therapy)

Do you experience significant voice changes or problems prior to or during menstruation?

ENT history

Have you been evaluated by an ENT?

(no, yes, If yes when was it, reason for that, examinations done like laryngoscopy)

Do you have or have had any of the following problems?

(Nasal obstruction; Nose deformity; Difficult nasal breathing; Purulent nasal discharge; Facial pain; Generalized dental pain; Watery rhinorrhea, Sneezing; Frequent nose bleeding, Frequent mouth breathing, Itchy nose; Recent cold; Persistent cough; Chest pain during exercise; Dyspnoea, Wheezing respiratory sounds; Snoring; Frequent oral mucus; Chronic sore throat, Pharyngeal infection; Neck pain; Muffled ears, Otalgia; Itchy ears; Hearing problems; Itchy eyes, Occasional headache, Bruxism, Migraines).

(No, Unsure, Yes; if yes when was it and specify, assessment done, prescribed treatment,)

Have you been evaluated by an allergist?

(no, yes, if yes, what allergies do you have (dust, mold, trees, cats, dogs, foods, medication, other)

Voice history

Voice usage

Do you usually speak with a loud voice? (No, yes, If yes, specify)

Do you frequently yell? (No, yes, If yes, specify)

Do you usually speak fast? (No, yes, If yes, specify)

While speaking, do you usually "run out of air" or lose your voice at the end of a sentence? (No, yes, If yes, specify)

Do you usually find yourself doing most of the talking when conversing with a friend or in a social gathering? (No, yes, If yes, specify)

Do you usually speak in a voice that is too low in pitch? (No, yes, If yes, specify)

Do you frequently speak in a whisper? (No, yes, If yes, specify)

Do you have hobbies that involve your voice? What are the conditions?

Voice usage professionally

In what capacity do you use your voice professionally?

(teacher, coach, salesperson, actor, singer, telephone operator, receptionist, announcer (television/radio/sports arena), attorney, clergy, politician, other);

since when interruptions (when)?

How much are you speaking at present (average hours a day)?

What are the conditions of speaking at work?

1.irritant airborne substances (dusty ambience, smoky environment, chalk, toxic and chemical exposition, air-conditioning);

2.noisy/non-noisy ambience (music, machines, people; library ambience);

3.temperature variations/ humidity-dry (fridges; swimming pool; labs)

4.outdoors (gym)

Have you ever had training for your speaking/singing voice?

When? For how long?

Voice Problems

Do you have or have you had a voice problem?

If yes, when, who notices it?

Do you know what caused it?

(work/home conditions; hobbies;stress;personal injury; health problem, surgery, specify)

Was the onset sudden or gradual in nature?

Is the problem acute (days or weeks) or chronic (months or years)?

Is it getting worse, better,same?

If the problem varies, what factors seem to aggravate or relieve the symptoms (e.g., weather, amount of speaking, noisy environment; coughing episodes, time of day)?

Which symptoms do you have or have had?

Tickling or choking sensation while speaking;

Dryness in the throat;

Voice fatigue (voice tires or changes quality after speaking for a short period of time;

Pain in throat while speaking)

Burning sensation;

Volume disturbance (trouble speaking softly or loudly);

Loss of range (high or low pitch);

Hoarseness (coarse or scratchy sound);

Breathiness

Other (specify)

Have there been previous evaluations? (ENT/SLT; when; type of evaluation)

Have there been previous treatments? (surgery; voice therapy; radiation therapy;other)

Smoking history

Do you smoke? Never;

Quit. (When/Why? For how long did you smoke? How many packs a day did you smoke?)

Yes What kind – cigarettes (with or without filter, name), cigar, pipe; How many packs/ or cigarettes per day? For how long (Years) Since when?

Continuosly? Interruptions (when)

Do you live in a smoky environment? (if yes, specify)

Do you work in a smoky environment?

Is there past or present use of any recreational drugs (street drugs)? such as pot (marijuana), speed (amphetamines) Dexetrine, Ritalin, LSD (acid), PCP (angel dust), cocaine, heroin, mushrooms, barbiturates, other)

Drinking history

Do you frequently have thirsty? (if yes, specify)

How much water do you drink a day? (half a litre , 1 litres, 1,5 litres, 2 litres , or more).

What kind of water do you usually drink (sparkling or still)?

At what temperature do you usually drink (natural or cold)?

Do you usually drink milk?

If yes; How much milk do you drink a day? (cups)

Do you usually drink tea?

If yes; How much tea do you drink a day? (cups)

What kind of tea do you usually drink (black or herbal)?

At what temperature do you usually drink (cold, hot or too hot)?

Do you take sugar? How many spoons per cup?

Do you drink coffee?

What kind of coffee do you usually drink (espresso, instant coffee, other)?

How much coffee do you drink a day? (cups)

Do you take sugar? How many spoons per cup?

Do you drink coffee and milk?

How much do you drink a day? (cups)

Do you take sugar? How many spoons per cup?

Do you drink soft drinks?

If yes, What kind of soft drinks do you usually drink?

(coca-cola, natural or sparkling citric juice, others)

How much do you drink a day?

Do you drink alcohol?

No; Rarely; A few times per week; Daily

If daily or a few times a week, on average how much alcohol do you consume? (glasses per day or week of beer, wine or liquor)

Did you formerly drink more heavily?

Food history

Height: Weight: Sudden variations (reason)

Do you frequently eat:

(salted food; spicy food; high-fat foods; dairy products (yoghurt, cheese); chocolate or other sweets (between or after meals).

Is there a special or restricted diet? (If yes, why; specify)

Do you eat late at night?

Do you have or have ever had any of the following problems?

(Excessive dry mouth; Swallowing difficulties; Odynophagia; "Wet-gurgly vocal quality"; TMJ pain, "Heartburn"; Halitosis, Frequent vomiting; Sour taste; Throat clearing from excessive phlegm; Microstomia; Frequent nausea; Throat globus sensation.)

Do you ever used antacids?

Appendix E – Voice Handicap Index

(English and Portuguese version)

CONFIDENTIAL

Subject number

Instructions:
These are the statements that many people have used to describe the effects of their voices on their lives. Check the response that indicates how frequently you have the same experience. (Never=0 points; Almost never=1 point; Sometimes=2 points; Almost always= 3 points; Always= 4 points)

	Never	Almost never	Sometimes	Almost always	Always
F1 My voice makes difficult for people to hear me.					
P2 I run out of air when I talk.					
F3 People have difficulty understanding me in a noisy room.					
P4 The sound of my voice varies throughout the day.					
F5 My family has difficulty hearing me when I call them throughout the house.					
F6 I use the phone less often than I would like.					
E7 I'm tense when talking with others because of my voice.					
F8 I tend to avoid groups of people because of my voice.					
E9 People seem irritated with my voice.					
P10 People ask, "What's wrong with my voice?".					
F11 I speak with friends, neighbours, or relatives less often becuse of my voice.					
F12 People ask me to repeat myself when speaking face-to-face.					

	Never	Almost never	Sometimes	Almost always	Always
P13 My voice sounds creaky and dry.					
P14 I feel as though I have to strain to produce voice.					
E15 I find other people don't understand my voice problem.					
F16 My voice difficulties restrict my personal and social life.					
P17 The clarity of my voice is unpredictable.					
P18 I try to change my voice to sound different.					
F19 I feel left out of conversations because of my voice.					
P20 I use a great deal of effort to speak.					
P21 My voice is worse in the evening.					
F22 My voice problem causes me loose income.					
F23 My voice problem upsets me.					
E24 I am less outgoing because of my voice problem.					
E25 My voice makes me feel handicapped.					
P26 My voice "gives out" on me in the middle of speaking.					
E27 I feel annoyed when people ask me to repeat					
E28 I feel embarrased when people ask me to repeat.					
E29 My voice makes me feel incompetent.					
E30 I'm ashamed of my voice problem.					

CONFIDENCIAL

Sujeito nº

Instruções
Estas são declarações que muitas pessoas usaram para descrever os efeitos das suas vozes, nas suas vidas. Assinale a resposta que indica com que frequência teve a mesma experiência (Nunca= 0 pontos; Quase nunca= 1 ponto; Ás vezes= 2 pontos; Quase sempre= 3 pontos; Sempre= 4 pontos).

	Nunca	Quase nunca	Ás vezes	Quase sempre	Sempre
F1 A minha voz faz com que seja difícil os outros ouvirem-me.					
Fi2 Falta-me o ar quando falo.					
F3 As pessoas têm dificuldade em me compreender num local ruidoso.					
Fi4 O som da minha voz varia ao longo do dia.					
F5 A minha família tem dificuldade em me ouvir quando os chamo dentro de casa.					
F6 Uso menos o telefone do que gostaria.					
E7 Fico tenso quando falo com os outros por causa da minha voz.					
F8 Costumo evitar grupos de pessoas por causa da minha voz.					
E9 As pessoas parecem irritadas por causa da minha voz.					
Fi10 As pessoas perguntam "O que se passa com a minha voz?".					
F11 Falo menos com amigos, vizinhos, ou familiares por causa da minha voz.					
F12 As pessoas pedem-me para repetir quando falamos cara a cara.					
Fi13 A minha voz é àspera e seca.					

	Nunca	Quase nunca	Às vezes	Quase sempre	Sempre
Fi14 Sinto como se tivesse de me esforçar para produzir voz.					
E15 Sinto que as outras pessoas não compreedem o meu problema de voz.					
F16 As minhas dificuldades com a voz limitam-me a minha vida pessoal e social.					
Fí17 A clareza da minha voz é imprevísivel.					
Fí18 Tento modificar a minha voz de modo a soar diferente.					
F19 Sinto-me fora das conversas por causa da minha voz.					
Fí20 Faço muito esforço para falar.					
Fí21 A minha voz está pior à noite.					
F22 O meu problema de voz causa-me problemas económicos.					
E23 O meu problema de voz preocupa-me.					
E24 Saio menos por causa do meu problema de voz.					
E25 A minha voz faz-me sentir deficiente.					
Fí26 A minha voz "falha" quando estou no meio de uma conversa.					
E27 Sinto-me irritado quando as pessoas me pedem para repetir.					
E28 Sinto-me embaraçado quando as pessoas me pedem para repetir.					
E29 A minha voz faz-me sentir incompetente.					
E30 Tenho vergonha do meu problema de voz.					

Appendix F – Social Readjustment Rating Scale

(English and Portuguese version)

CONFIDENTIAL

Subject number

CONFIDENCIAL

Subject number

	Life event	Mean value
1	Death of spouse	100
2	Divorce	73
3	Marital separation from mate	65
4	Detention in jail or other institution	63
5	Death of a close family member	63
6	Major personal injury or illness	53
7	Marriage	50
8	Being fired at work	47
9	Marital reconciliation with mate	45
10	Retirement from work	45
11	Major change in the health or behavior of a family member	44
12	Pregnancy	40
13	Sexual difficulties	39
14	Gaining a new family member (e.g., through birth, adoption, oldster moving in, etc.)	39
15	Major business readjustment (e.g., merger, reorganization, bankruptcy, etc.)	39
16	Major change in financial state (e.g., a lot worse off or a lot better off than usual)	38
17	Death of a close friend	37
18	Changing to a different line of work	36
19	Major change in the number of arguments with spouse (e.g., either a lot more or a lot less than usual regarding child-rearing, personal habits, etc.)	35
20	Taking out a mortgage or loan for a major purchase (e.g., for a home, business, etc.)	31
21	Foreclosure on a mortgage or loan	30
22	Major change in responsibilities at work (e.g., promotion, demotion, later transfer)	29
23	Son or daughter leaving home (e.g., marriage, attending college, etc.)	29
24	Trouble with in-laws	29
25	Outstanding personal achievement	28
26	Wife beginning or ceasing work outside the home	26
27	Beginning or ceasing formal schooling	26
28	Major change in living conditions (e.g., building a new home, remodeling, deterioration of home or neighborhood)	25
29	Revision of personal habits (dress, manners, associations, etc.)	24
30	Trouble with boss	23
31	Major change in working hours or conditions	20
32	Change in residence	20
33	Change to a new school	20
34	Major change in usual type and/or amount of recreation	19
35	Major change in church activities (e.g., a lot more or a lot less than usual)	19
36	Major change in social activities (e.g., clubs, dancing, movies, visiting, etc)	18
37	Taking out a mortgage or a loan for a lesser purchase (e.g., for a car, TV, freezer, etc.)	17
38	Major change in sleeping habits (a lot more or a lot less sleep, or change in part of day when asleep)	16
39	Major change in number of family get-togethers (e.g., a lot more or a lot less than usual)	15
40	Major change in eating habits (a lot more or a lot less food intake, or very different meal hours or sourroundings)	15
41	Vacation	13
42	Christmas	12
43	Minor violations of the law (e.g., traffic tickets, jaywalking, disturbing the peace, etc.)	11

CONFIDENCIAL

Sujeito número

	Ocorrências da vida	Valor do Stress
1	Morte do cônjuge	100
2	Divórcio	73
3	Separação matrimonial	65
4	Período na prisão	63
5	Morte de familiar chegado	63
6	Ferimentos pessoais ou doença	53
7	Casamento	50
8	Despedimento do emprego	47
9	Reconciliação conjugal	45
10	Reforma	45
11	Alteração na saúde de um familiar	44
12	Gravidez	40
13	Dificuldades sexuais	39
14	Novo membro na família (ex. nascimento, adopção, familiar mais velho que vem viver com a família, etc.)	39
15	Reajustamento profissional (ex. fusão de duas empresas, reorganização, falência, etc.)	39
16	Mudança na situação económica (ex. muito melhor ou pior que o habitual).	38
17	Morte de amigo íntimo	37
18	Mudança para diferente ramo de emprego	36
19	Alteração no número de discussões com o cônjuge (ex. muito mais ou menos do que o habitual por causa da educação dos filhos, hábitos pessoais, etc).	35
20	Hipoteca ou empréstimo para uma grande compra (ex. casa, negócio, etc)	31
21	Execução de hipoteca ou empréstimos	30
22	Alteração das responsabilidades no trabalho (ex. promoção, despromoção, transferência)	29
23	Filho ou filha que sai de casa (ex. casamento, para estudar, etc.)	29
24	Problemas com familiares do cônjuge	29
25	Realização pessoal excelente	28
26	Esposa que começa ou deixa de trabalhar fora de casa.	26
27	Começo ou fim de escola	26
28	Alteração nas condições de vida (ex.construção de uma casa, obras de remodelação, deterioração da casa ou da zona).	25
29	Revisão de hábitos pessoais (vestir, modos, associações)	24
30	Problemas com o patrão	23
31	Alteração no horário ou condições de trabalho	20
32	Mudança de casa	20
33	Mudança de escola	20
34	Mudança de tipo e quantidade de passatempos	19
35	Mudança nas actividades religiosas (ex. muito mais ou muito menos que o habitual)	19
36	Mudança nas actividades sociais (ex. clubes, dançar, cinema, visitas, etc.)	18
37	Hipoteca ou empréstimo para uma pequena compra (ex. carro, TV, frigorífico, etc.)	17
38	Mudança nos hábitos de sono (ex. muito mais ou muito menos sono ou mudança relativamente à parte do dia em que se sente ensonado).	16
39	Mudança no número de encontros familiares (muito mais ou muito menos que o habitual)	15
40	Mudança nos hábitos alimentares (ex. muito mais ou muito menos quantidade, horários de refeição ou ambientes diferentes)	15
41	Férias	13
42	Natal	12
43	Pequenas violações à lei (ex. multas de trânsito, distúrbios da paz, etc)	11

Appendix G – The Story of Arthur the Rat

(English and Portuguese)

The Story of Arthur the Rat

There was once a young rat named Arthur who would never take the trouble to make up his mind. Whenever his friends asked him if he would like to go out with them, he would only answer, "I don't know." He wouldn't say, "Yes" and he wouldn't say "No" either. He could never learn to make a choice.

His Aunt Helen said to him "No-one will ever care for you if you carry on like this. You have no more mind than a blade of grass. " Arthur looked wise but said nothing.

One rainy day the rats heard a great noise in the loft where they lived. The pine rafters were all rotten, and at last one of the joists had given way and fallen to the ground. The walls shook and the rats' hair stood on end with fear and horror. " This won't do," said the old rat who was chief, "I'll send out scouts to search for a new home."

Three hours later the seven scouts came back and said, "We've found a stone house which is just what we wanted. There's room and good food for us all. There's a kindly horse named Nelly, a cow, a calf and a garden with an elm tree." Just then the old rat caught sight of young Arthur. "Are you coming with us?" he asked. "I don't know," Arthur, sighed, "The roof may not come down just yet." "Well," said the old rat angrily, "We can't wait all day for you to make up your mind. Right about face! March!" And they went off.

Arthur stood and watched the other rats hurry away. The idea of an immediate decision was too much for him. "I'll go back to my hole for a bit," he said to himself, "just to make up my mind."

That night there was a great crash that shook the earth, and down came the whole roof. Next day some men rode up and looked at the ruins. One of them moved a board, and under it they saw a young rat lying on his side, quite dead, half in and half out of his hole.

In Abercrombie, D. (1964). **Elements of general phonetics**. Edinburgh: U.P. Edinburgh.

A história do rato Artur

Havia uma vez um rato chamado Artur, que era muito indeciso. Sempre que os seus amigos lhe perguntavam se queria sair com eles, ele apenas respondia:

- Não sei.

Ele nunca dizia que "sim", nem que "não". Não era capaz de tomar decisões.

A sua tia Helena disse-lhe:

- Ninguém se interessará por ti se continuares a ser assim. Tu não tens mais cabeça do que uma folha de erva. Artur olhou atento mas não disse nada.

Num dia chuvoso os ratos ouviram um grande barulho no sotão onde viviam. As traves de pinho estavam todas estragadas e por fim uma das juntas cedeu e caiu ao chão. As paredes tremiam e os ratos ficaram com os pelos em pé, cheios de medo e terror.

-Isto não dá, disse o rato velho que era o chefe,

-Vou mandar escuteiros lá para fora à procura de uma nova casa.

Três horas mais tarde os sete escuteiros voltaram e disseram:

-Encontrámos uma casa de pedra que é mesmo o que nós queríamos. Há espaço e boa comida para todos nós. Há lá um cavalo, simpático, chamado Nelly, uma vaca, um novilho e um jardim com uma árvore.

De repente o rato velho lembrou-se do jovem Artur.

- Vens connosco? perguntou.

- Não sei, suspirou Artur, "Talvez o telhado não venha abaixo já".

- Bom, disse o rato velho zangado, "Não podemos esperar todo o dia até que te decidas. Toca a andar! Marchar!" e foram-se embora.

Artur ficou e viu os outros ratos a irem embora cheios de pressa. A ideia de uma decisão imediata era de mais para ele.

-Vou voltar para o meu buraco por um momento, disse para ele mesmo, "só para me decidir".

Nessa noite houve uma grande tempestade que fez tremer a terra, e o telhado veio todo abaixo. No dia seguinte, vieram uns homens ver as ruínas. Um deles moveu uma tábua, e debaixo dela viram um rato jovem deitado de lado, morto, meio dentro e meio fora do seu buraco.

Appendix H – Frequence distribution of VHI ratings

Table - Frequencies of responses in the VHI subscales

	Dysphonics					Controls				
Emotional subscale	Never	Almost never	Some times	Almost always	Always	Never	Almost never	Some times	Almost always	Always
E7 I'm tense when talking with others because of my voice.	45.1	13.7	21.6	7.8	11.8	80.4	3.6	12.5	1.8	1.8
E9 People seem irritated with my voice.	56.9	19.6	19.6	3.9	0	89.3	3.6	7.1	0	0
E15 I find other people don't understand my voice problem.	45.1	15.7	29.4	5.9	3.9	92.9	1.8	1.8	0	3.6
E23 My voice problem upsets me.	19.6	2	33.3	11.8	33.3	76.8	8.9	8.9	0	5.4
E24 I am less outgoing because of my voice problem.	76.5	9.8	3.9	3.9	5.9	94.6	1.8	0	0	3.6
E25 My voice makes me feel handicapped.	80.4	0	11.8	3.9	3.9	96.4	0	1.8	0	1.8
E27 I feel annoyed when people ask me to repeat	54.9	25.5	15.7	2	2	76.8	16.1	5.4	1.8	0
E28 I feel embarrased when people ask me to repeat.	54.7	17.6	11.8	2	3.9	75	21.4	1.8	1.8	0
E29 My voice makes me feel incompetent.	72.5	9.8	9.8	0	7.8	89.3	7.1	0	1.8	1.8
E30 I'm ashamed of my voice problem	80.4	9.8	7.8	0	2	98.2	0	0	1.8	0
Functional subscale	Never	Almost never	Some times	Almost always	Always	Never	Almost never	Some times	Almost always	Always
F1 My voice makes difficult for people to hear me.	41.2	15.7	27.5	11.8	3.9	64.3	17.9	16.1	1.8	0
F3 People have difficulty understanding me in a noisy room.	33.3	11.8	25.5	15.7	13.7	41.1	21.4	26.8	5.4	5.4
F5 My family has difficulty hearing me when I call them throughout the house.	43.1	21.6	17.6	11.8	5.9	71.4	21.4	7.1	0	0
F6 I use the phone less often than I would like.	70.6	9.8	13.7	3.9	2	82.1	5.4	8.9	1.8	1.8
F8 I tend to avoid groups of people because of my voice.	80.4	5.9	7.8	3.9	2	92.9	3.6	3.6	0	0
F11 I speak with friends, neighbours, or relatives less often because of my voice.	23.5	11.8	37.3	9.8	17.6	91.2	5.3	0	1.8	1.8
F12 People ask me to repeat myself when speaking face-to-face.	56.9	15.7	13.7	9.8	3.9	61.4	22.8	14	0	1.8
F16 My voice difficulties restrict my personal and social life.	45.1	11.8	25.5	3.9	13.7	83.9	7.1	5.4	0	3.6
F19 I feel left out of conversations because of my voice	66.7	9.8	15.7	2	5.9	89.3	5.4	3.6	0	1.8
F22 My voice problem causes me loose income.	86.3	0	7.8	0	5.9	96.4	0	0	0	3.6

Table – Frequencies of responses (%) in the VHI subscales (continuation)

	Dysphonics					Controls				
Physical subscale	**Never**	**Almost never**	**Sometimes**	**Almost always**	**Always**	**Never**	**Almost never**	**Sometimes**	**Almost always**	**Always**
P2 I run out of air when I talk	36	16	40	8	0	64.3	25.0	7.1	1.8	1.8
P4 The sound of my voice varies throughout the day.	7.8	13.7	43.1	19.6	15.7	46.4	33.9	10.7	3.6	5.4
P10 People ask "What's wrong with my voice?"	23.5	11.8	37.3	9.8	17.6	80.4	10.7	5.4	0	3.6
P13 My voice sounds creaky and dry	26	8	36	16	14	58.9	26.8	12.5	1.8	0
P14 I feel as though I have to strain to produce voice.	15.7	15.7	35.3	11.8	21.6	64.3	21.4	8.9	3.6	1.8
P17 The clarity of my voice is unpredictable.	17.6	21.6	33.3	15.7	11.8	53.6	32.1	10.7	0	3.6
P18 I try to change my voice to sound different.	33.3	9.8	35.3	7.8	13.7	62.5	21.4	12.5	1.8	1.8
P20 I use a great deal of effort to speak	23.5	15.7	35.3	13.7	11.8	76.8	12.5	7.1	0	3.6
P21 My voice is worse in the evening.	22	20	30	14	14	62.5	23.2	10.7	1.8	1.8
P26 My voice "gives out" on me in the middle of speaking.	27.5	17.6	37.3	7.8	9.8	60.7	23.2	12.5	0	3.6

Biographical sketch

Isabel Cristina Ramos Peixoto Guimarães was born in Oporto, Portugal on May 1962.

She is a lecturer in the Speech Therapy Department at the Higher School of Health of Alcoitão, Portugal since 1988. She received her bachelor's (1985) degree in speech therapy and a higher specialized studies (1993) degree in teaching and administration from the Higher School of Health of Alcoitão. In October (1998) she was awarded a scholarship from the Calouste Gulbenkian Foundation to start Ph.D. studies (Experimental Phonetics) in the Phonetics and Linguistics department at University College London.

Since the beginning of her career she has received additional training in voice, speech and swallowing from national and internationally recognized leaders in the field. She also worked in the education field for three years (1985-1988), did private clinical work for 8 years (1989-1998) working with children and adults, published articles and a book in the voice, speech and swallowing areas and has presented workshops for speech therapists and other professionals since 1989.

She was for five consecutive years (1987-1992) president of the Portuguese Association of Speech Therapists and participated in the foundation (1989) and the following seven years of activity of the Permanent Liasion Committee of EU Speech Therapists and Logopedes.

Her research and clinical interests have focused on voice, speech and swallowing.