# THE DEVELOPMENT OF VOICE DISORDER EVALUATION SYSTEM BASED ON DYSPHONIA SEVERITY INDEX

MOHD REDZUAN BIN JAMALUDIN

A thesis submitted in fulfillment of the
requirements for the award of
the Degree of Master of Engineering (Biomedical)

Faculty of Health Science and Biomedical Engineering
Universiti Teknologi Malaysia

JANUARY 2012

Dedicated to Abah, Umi, Akak, Abang Shahriz, Abangah, Kak Ida, Eg, Hanah, Hariz, Hadif, Qaseh, and Siti Nur Shazleena.

## ACKNOWLEDGEMENT

In the name of Allah the Most Gracious, the Most Merciful, praised to Him for the opportunity that He has given to me to keep on living, giving me strength, and pursuing my study up to this level. Without His blessings, I would not have been able to come this far.

First and foremost, I would like to wish my deepest gratitude to my supervisor Prof. Ir. Dr. Sheikh Hussain Shaikh Salleh for his endless guidance and support throughout the development of this thesis. With his enthusiastic support, wisdom, and advice, I am motivated to continue researching new knowledge for the benefits of mankind.

I would also like to thank my co-supervisor, Dr. Tan Tian Swee for the continuous sharing of knowledge and experience. Anytime that I would face a problem regardless the matter, he would always help me in any way that he can. My appreciation also goes to my second co-supervisor Prof. Madya. Dr. Kartini Ahmad for her trusts in my ability and strength. She would also share her knowledge without any constrain whenever I seek upon her guidance.

I am also thankful to the staffs of the Centre for Biomedical Engineering for their help and assistance to provide me resource and pleasant environment to make my research a success. Special thanks to Amar, Kamarul, and Ting for their willingness to share their knowledge and technical support. I also wish to thank my parents, family and friends for their support and encouragement that would uplift my spirit and strength whenever I am feeling down and discouraged. Thank you everyone that has directly and indirectly supports me in any ways to make this thesis a success.

# ABSTRACT

Voice disorder is dramatically increasing due to the unhealthy social habits such as smoking and alcohol consumption, voice abuse, and the most importantly the lack of awareness among the general public and from the health care provider. Objective non-invasive multiparameter voice assessment is seen as a way to improve the voice rehabilitation process by allowing home care at own responsibility. The purpose of the research is to develop an automatic voice diagnostic system based on objective non-invasive multiparameter method known as Dysphonia Severity Index (DSI). DSI consists of four parameters which are the highest pitch, jitter percentage, lowest intensity, and maximum phonation time. They are combined into a linear regression equation that will give values from -5 to +5 indicating severely dysphonic voice or normal voice respectively. The proposed system is named as Automatic Dysphonia Evaluation System (ADES). It integrates a new proposed pitch detection algorithm (PDA), start/end point detection algorithm, jitter equation, and intensity equation to obtain the four DSI parameters allowing the system to be used by patient at home to monitor their voices. The proposed PDA was proven more accurate by having no error detected for normal voice while only one pathological voice was detected with doubling error. The modified start/end point detection algorithm is proven better with silence detection error rate of 0.0752. ADES was tested with KayPENTAX voice database and had 55.6054% sensitivity and 50% specificity when -6.7249 is used as the cutoff value. Different sets of database consisted of trained and untrained vocalists, and also teachers and non-teachers were also used to evaluate ADES' performance. The results of ADES show that it is able to get the DSI values for different voices from different types of groups.

# ABSTRAK

Kecacatan suara merupakan masalah yang semakin meningkat kini disebabkan oleh aktiviti yang tidak sihat seperti merokok, meminum minuman beralkohol, penyalahgunaan suara, dan kurang kesedaran diberikan kepada orang awam dari pihak kesihatan. Kaedah objektif yang “non-invasive” dan “multiparameter” merupakan kaedah yang dilihat boleh menambahbaik proses rehabilitasi suara dengan membenarkan pesakit untuk menjalani rehabilitasi di rumah. Tujuan kajian ini dijalankan adalah untuk membangunkan sistem yang boleh menganalisa suara secara automatik dengan menggunakan kaedah “Dysphonia Severity Index” (DSI). DSI terdiri daripada pengiraan empat parameter suara iaitu nada tertinggi, peratusan “jitter”, intensiti terendah, dan “Maximum Phonation Time” (MPT). Parameter-parameter ini akan digabung dalam satu persamaan regresi linar yang akan memberi nilai dari -5(suara yang bermasalah) hingga +5(suara normal). Sistem yang dicadangkan dinamakan “Automatic Dysphonia Evaluation System” (ADES). Sistem ini menggabungkan algoritma pengesan nada (PDA) yang baru dicadangkan dalam tesis ini, algoritma pengesan permulaan dan pengakhiran pembunyian yang dimodifikasi, persamaan “jitter” dan persamaan keamatan untuk memperoleh keempat-empat parameter DSI supaya para pesakit boleh memantau suara mereka. PDA yang dicadangkan ternyata lebih baik berbanding algoritma-algoritma yang dibandingkan dengan memperoleh ralat sifar untuk suara normal dan satu suara bermasalah dikesan dengan “doubling error”. Algoritma pengesan permulaan dan pengakhiran pembunyian yang dimodifikasi juga ternyata lebih efektif dengan ralat “silence detection” sebanyak 0.0752. ADES telah diuji dengan data suara daripada KayPENTAX dan memperoleh 55.6054% kepekaan dan 50% kekhususan. Data suara daripada vokalis terlatih dan tidak terlatih, serta guru dan bukan guru turut digunakan untuk menguji prestasi ADES. Keputusannya, ADES berupaya untuk mendapatkan nilai-nilai DSI untuk kumpulan suara yang berbeza.

# TABLE OF CONTENT


| CHAPTER | TITLE | | PAGE |
|---|---|---|---|
| | **TITLE PAGE** | | i |
| | **DECLARATION PAGE** | | ii |
| | **DEDICATION** | | iii |
| | **ACKNOWLEDGEMENT** | | iv |
| | **ABSTRACT** | | v |
| | **ABSTRAK** | | vi |
| | **TABLE OF CONTENTS** | | vii |
| | **LIST OF TABLES** | | xi |
| | **LIST OF FIGURES** | | xiii |
| | **LIST OF SYMBOLS** | | xx |
| | **LIST OF APPENDICES** | | xxvi |
| **1** | **INTRODUCTION** | | 1 |
| | 1.1 | Background of Research | 1 |
| | 1.2 | Problem Statement | 2 |
| | 1.3 | Objectives | 4 |
| | 1.4 | Scope of the Research | 5 |
| | 1.5 | Thesis Layout | 5 |
| | 1.6 | Contribution of Thesis | 6 |

| | | | | | |
|---|---|---|---|---|---|
| **2** | **LITERATURE REVIEW** | | | | 8 |
| | 2.1 | Reviews on Voice Measurement Techniques | | | 8 |
| | | 2.1.1 | The Human Speech Production System and Acoustic Wave Generation | | 8 |
| | | 2.1.2 | Types of Speech Disorders | | 11 |
| | | | 2.1.2.1 Articulation Disorders | | 12 |
| | | | 2.1.2.2 Fluency Disorders | | 12 |
| | | | 2.1.2.3 Voice Disorders | | 13 |
| | | | 2.1.2.4 Language Disorders | | 15 |
| | | 2.1.3 | The Assessment for Voice Disorders | | 15 |
| | | | 2.1.3.1 Invasive Approach | | 16 |
| | | | 2.1.3.2 Non-invasive Approach | | 17 |
| | | | | 2.1.3.2.1 History of Patient | 18 |
| | | | | 2.1.3.2.2 Auditory Perceptual Analysis | 20 |
| | | | | 2.1.3.2.3 Instrumental Analysis | 21 |
| | | 2.1.4 | Summary | | 34 |
| | 2.2 | Reviews on Technical Aspects of Dysphonia Severity Index Parameters | | | 35 |
| | | 2.2.1 | Highest pitch | | 35 |
| | | | 2.2.1.1 Introduction on Pitch | | 35 |
| | | | 2.2.1.2 Difficulties in Pitch Detection | | 38 |
| | | | 2.2.1.3 Pitch Detection Algorithm | | 39 |
| | | | | 2.2.1.3.1 Time Domain Based Pitch Detection Algorithm | 40 |
| | | | | 2.2.1.3.2 Frequency | 44 |

| | | | | | |
|---|---|---|---|---|---|
| | | | | Domain Based Pitch Detection Algorithm | |
| | | | 2.2.1.3.3 | Alternative Transformation Domain Based Pitch Detection Algorithm | 45 |
| | | 2.2.1.4 | Reviews on Pitch Detection Algorithm | | 46 |
| | 2.2.2 | Jitter | | | 60 |
| | 2.2.3 | Maximum Phonation Time | | | 64 |
| | | 2.2.3.1 | Start/End Point Detection | | 64 |
| | 2.2.4 | Lowest Intensity | | | 73 |
| | 2.2.5 | Summary | | | 74 |

| | | | | |
|---|---|---|---|---|
| **3** | **METHODOLOGY** | | | 75 |
| | 3.1 | Introduction | | 75 |
| | 3.2 | Research Methodology | | 75 |
| | 3.3 | Voice Disorder Evaluation System Based on Dysphonia Severity Index | | 77 |
| | 3.4 | Start/End Point Detection | | 78 |
| | 3.5 | Pitch Detection Algorithm | | 84 |
| | 3.6 | Jitter Calculation | | 95 |
| | 3.7 | Intensity | | 95 |
| | 3.8 | Graphical User Interface for ADES | | 96 |
| | | 3.8.1 | Highest Pitch Measurement | 98 |
| | | 3.8.2 | Jitter Measurement | 99 |
| | | 3.8.3 | Lowest Intensity Measurement | 101 |
| | | 3.8.4 | Maximum Phonation Time Measurement | 102 |
| | | 3.8.5 | Dysphonia Severity Index | 104 |

| | | | | |
|---|---|---|---|---|
| | | | Measurement | |
| | 3.9 | Subjects and Experiment Setup | | 104 |
| | | 3.9.1 | Subjects | 105 |
| | | | 3.9.1.1 Subjects for Pitch Detection Algorithm Accuracy Test | 105 |
| | | | 3.9.1.2 Subjects for Start/End Point Detection Algorithm Accuracy Test | 105 |
| | | | 3.9.1.3 Subjects for Sound Intensity Accuracy Test | 106 |
| | | | 3.9.1.4 Subjects for ADES Accuracy Test | 107 |
| | | 3.9.2 | Experiments | 108 |
| | | | 3.9.2.1 Accuracy Test for Pitch Detection Algorithm | 108 |
| | | | 3.9.2.2 Accuracy Test for Start/End Point Detection | 112 |
| | | | 3.9.2.3 Accuracy Test for Sound Intensity Calculation | 114 |
| | | | 3.9.2.4 Accuracy Test for Voice Disorder Evaluation System Based on Dysphonia Severity Index | 116 |
| | 3.10 | Summary | | 121 |
| **4** | **RESULTS AND DISCUSSION** | | | 123 |
| | 4.1 | Introduction | | 123 |
| | 4.2 | Accuracy Test for Pitch Detection Algorithm | | 123 |
| | 4.3 | Accuracy Test for Start/End Point Detection | | 129 |
| | 4.4 | Accuracy Test for Sound Intensity Calculation | | 135 |
| | 4.5 | Accuracy Test for Voice Disorder Evaluation System Based on Dysphonia Severity Index | | 137 |

| | | | |
|---|---|---|---|
| **5** | **CONCLUSION AND FUTURE WORK** | | 147 |
| | 5.1 | Voice Disorder Evaluation System Based on Dysphonia Severity Index | 147 |
| | 5.2 | Future Work | 150 |

**REFERENCES** 151

Appendices A - D 159

# LIST OF TABLES


| TABLE NO. | TITLE | PAGE |
|---|---|---|
| 2.1 | A summary of the studies on objective multiparameter or multivariate model | 22 |
| 2.2 | Researches that used DSI | 27 |
| 2.3 | Results of Poincaré PDA and Praat PDA | 58 |
| 3.1 | Example of the marked peaks before rearrangement and after rearrangement of ACF | 91 |
| 4.1 | Performance of developed PDA with different $K_P$ on 21 normal voices | 124 |
| 4.2 | Performance of developed PDA with different $K_P$ on 51 pathological voices | 124 |
| 4.3 | Halving, doubling, and gross errors for the three PDA’s on 21 normal voices | 125 |
| 4.4 | Halving, doubling, and gross errors for the three PDA’s | 125 |

| | | |
|---|---|---|
| | on 51 pathological voices | |
| 4.5 | Transcription of the word ‘Enam’ obtained from Wavesurfer | 129 |
| 4.6 | Transcription of the word ‘Enam’ obtained from the modified start/end point algorithm | 130 |
| 4.7 | Comparison of sound level meter, ADES, and Praat software intensity readings | 136 |
| 4.8 | Comparison of sound level meter, and Praat software intensity readings | 137 |
| 4.9 | Mean, highest value, lowest value, and standard deviation of DSI parameters for 25 normal voices | 138 |
| 4.10 | Mean, highest value, lowest value, and standard deviation of DSI parameters for 200 pathological voices | 138 |
| 4.11 | Sensitivity and specificity of the DSI | 139 |
| 4.12 | Sensitivity and specificity of the DSI | 140 |
| 4.13 | Experiment results on karaoke sessions on participants’ voice DSI’s | 141 |
| 4.14 | The comparison of DSI between teacher and non-teacher groups | 144 |

## LIST OF FIGURES


| FIGURE NO. | TITLE | PAGE |
|---|---|---|
| 2.1 | The diagram of respiration and parts that are related to voice production (Stevens, 1999) | 9 |
| 2.2 | The vocal cord a) Illustration of the vocal cord with labels (Quatieri, 2002), b) Image of the vocal cord via video laryngoscope (Verikas *et al*., 2006) | 9 |
| 2.3 | Autoregressive modeling of the speech signal (Tan, 2009) | 10 |
| 2.4 | Types of speech disorders (Charles and Robert, 1996) | 11 |
| 2.5 | Images of vocal cords a) Healthy, b) Nodular class, c) Diffuse class (Verikas *et al*., 2006) | 14 |
| 2.6 | The fractions of voice assessment approaches | 16 |
| 2.7 | The EVA® software interface. Arrow B indicates the length of 2000 ms window selected for analysis which started after the voice onset indicated by arrow A with the length of 200 ms (Yu *et al*., 2001) | 31 |

| | | |
|---|---|---|
| 2.8 | The open phase, close phase, and return phase of glottal waveform (Quatieri, 2002) | 36 |
| 2.9 | The vocal cords cycle of vibration (Kotnik *et al*., 2009) | 36 |
| 2.10 | Acoustic waveform with its corresponding ACF's a) Acoustic waveform of an /a/ utterance, b) The corresponding ACF according to equation (2.5), c) The corresponding ACF according to equation (2.6) | 42 |
| 2.11 | A voice signal with its corresponding AMDF. Top: A frame of a /a/ utterance. Bottom: AMDF of the /a/ utterance | 43 |
| 2.12 | The blue and red lines indicate each pitch period for a perfectly periodic speech signal. Blue lines are the pitch period found by using waveform matching technique, while the red lines are the marked peaks indicating each pitch period by using peak picking technique. | 62 |
| 2.13 | The blue and red lines indicate each pitch period for a perfectly speech waveform added with 1% noise. Blue lines are the pitch period found by using waveform matching technique, while the red lines are the marked peaks indicating each pitch period by using peak picking technique. | 63 |
| 2.14 | The accurate percentile of Three End-point algorithms. Dash-Line: Algorithm based on Short-Time Energy. Line: Algorithm based on Zero-Crossing Rate. Dot-Line: Algorithm based on Four- | 66 |

| | | |
|---|---|---|
| | Order Cumulants (Chen and Wang, 2005). | |
| 2.15 | a) Noise free utterance 'six' waveform, b) NACC peaks values by using 3-level binary HOS NACC algorithm (Alkulaibi *et al*., 1997). | 69 |
| 2.16 | a) Utterance 'six' waveform added with coloured noise of 5 dB SNR, b) NACC peaks values by using 3-level binary HOS NACC algorithm, c) NACC peaks values by using second order statistics (Alkulaibi *et al*., 1997) | 70 |
| 2.17 | a) Utterance 'six' waveform added with coloured noise of 0 dB SNR, b) NACC peaks values by using 3-level binary HOS NACC algorithm, c) NACC peaks values by using second order statistics (Alkulaibi *et al*., 1997). | 71 |
| 3.1 | The five phases involved in designing Voice Disorder Evaluation System Based on Dysphonia Severity Index. | 76 |
| 3.2 | The Architecture of ADES. | 77 |
| 3.3 | The flowchart of start/endpoint detection. | 78 |
| 3.4 | The flowchart of start/endpoint detection. | 79 |
| 3.5 | The flowchart of start/endpoint detection. | 80 |
| 3.6 | The recorded low intensity /a/ utterance. | 82 |
| 3.7 | Start and end point marked with two red lines of an /a/ utterance. | 84 |

| | | |
|---|---|---|
| 3.8 | The top figure is the corresponding pitch plot of the /a/ utterance shown in the bottom figure. | 85 |
| 3.9 | The speech signal of a frame separated to three sections indicated by two vertical red lines to obtain $c_{11}$, $c_{12}$, $c_{21}$, and $c_{22}$. | 87 |
| 3.10 | The process to obtain the centre-clipping threshold, $P_1$, and $P_2$. | 87 |
| 3.11 | Top: Speech waveform of an /a/ utterance. Bottom: Speech waveform after centre-clipping. | 88 |
| 3.12 | Top: A frame of speech signal of /a/ utterance. Middle: Signal after centre-clipping. Bottom: ACF of the centre-clipped signal. | 89 |
| 3.13 | Top: ACF of the centre-clipped speech. Bottom: ACF with marked peaks indicated by red lines. | 90 |
| 3.14 | The pitch marks in Figure 3.13 have been reduced into two dominant peaks. | 92 |
| 3.15 | The red lines indicate the new smaller specific frame to find the pitch period. | 93 |
| 3.16 | The maximum ACF within the new frame is the pitch period indicated by blue line. | 94 |
| 3.17 | Top: Intensity plot of vowel /a/. Bottom: Speech signal of vowel /a/. | 96 |
| 3.18 | The user interface of ADES. | 97 |

| | | |
|---|---|---|
| 3.19 | The user interface of the developed system. The vertical lines in the pitch figure indicate: Blue – Start and end point of the utterance, Red – Start point + 200 ms and end point – 200 ms, Green – The location of the highest pitch of the utterance. The highest pitch of the participant is 338.028 Hz. | 99 |
| 3.20 | The user interface of the developed system. The vertical lines in the pitch figure indicate: Blue – Start and end point of the utterance, Red – Start point + 200 ms and one second after start point + 200 ms. The jitter percentage of the participant is 0.13086%. | 100 |
| 3.21 | The user interface of the developed system. The vertical lines in the intensity figure indicate: Blue – Start and end point of the utterance, Red – Start point + 200 ms and end point – 200 ms, Green – The location of the minimum intensity of the utterance. The lowest intensity of the participant is 60.8675 dB SPL. | 102 |
| 3.22 | The user interface of the developed system. The vertical lines in the bottom figure of the speech waveform indicate: Red – Start point and end point. The maximum phonation time of the participant is 28.8993 s. | 103 |
| 3.23 | The user interface of the developed system. The value of DSI is shown once all the four parameters are obtained. The DSI of this participant is 1.9685. | 104 |
| 3.24 | Synthetic speech with $F_0$ of 140 Hz and a large formant of 280 Hz. | 107 |

| | | |
|---|---|---|
| 3.25 | A segment of speech waveform of /a/ utterance (red line) marked with the place of pitch marks (blue lines). | 109 |
| 3.26 | A segment of speech waveform of /a/ utterance (red line) marked with the place of pitch marks (blue lines). | 110 |
| 3.27 | Wavesurfer software interface. The displayed signal is the word 'Enam'. On top of the waveform circled with red line is the transcription of the signal which consists of silence (sil) and the transcript of the waveform itself 'Enam' (enam). | 113 |
| 3.28 | Experiment setup a) Researchers conducting the intensity experiment, b) Zoom in image of the apparatus. | 115 |
| 3.29 | The flow of getting DSI from a voice sample of KayPENTAX database. | 117 |
| 3.30 | The flow of getting actual DSI. | 117 |
| 4.1 | Example of autocorrelation function (Boersma, 1993). | 126 |
| 4.2 | A frame of AMDF by Manfredi *et al*. (2000) PDA. The horizontal red line indicates the threshold of pitch period selection and the circle marks the AMDF value that exceeds the threshold thus selected as the pitch mark. | 127 |
| 4.3 | A frame of AMDF by Manfredi *et al*. (2000) PDA. Horizontal red line indicates the threshold of pitch period selection and the circle marks the AMDF value | 128 |

| | | |
|---|---|---|
| | that exceeds the threshold therefore selected as the pitch mark. | |
| 4.4 | Modified start/end point algorithm labeled the start (green line) and the end point (purple line) of the 'Enam' utterance. | 130 |
| 4.5 | The silence detection error rates for five different start/end point detection algorithms. | 131 |
| 4.6 | The start and end point detected by Rabiner and Sambur (1975) algorithm. Top: Low intensity /a/ utterance marked with start and end points (red lines). Middle: ZCR of the utterance. Horizontal red line indicates the ZCR threshold, IZCT. Bottom: Short-time energy (STE) of the utterance. The two horizontal red lines indicate the lower STE threshold (ITL) and the upper STE threshold (ITU). | 133 |
| 4.7 | /a/ utterance with normal pitch and normal loudness. Wrong classification of start and end points by Rabiner and Sambur (1975) algorithm. | 134 |

## LIST OF ABBREVIATIONS

| | | |
|---|---|---|
| GRBAS | - | Grade, Roughness, Asthenia, and Strain |
| AVQI | - | Acoustic Voice Quality Index |
| VHI | - | Voice Handicap Index |
| DSI | - | Dysphonia Severity Index |
| ADES | - | Automatic Dysphonia Evaluation System |
| PDA | - | Pitch detection algorithm |
| PSOLA | - | Pitch Synchronize Overlap and Add |
| VFN | - | vocal fold nodules |
| SVM | - | support vector machine |
| HNR | - | Harmonics-to-noise ratio |
| Fmin | - | minimum frequency |
| Fmax | - | maximum frequency |
| Imin | - | minimal intensity |
| MPT | - | maximum phonation time |
| MF | - | mean flow |
| SGP | - | subglottal pressure |
| F | - | functional domain of VHI |
| E | - | emotional domain of VHI |
| P | - | physical domain of VHI |
| Hz | - | Hertz |
| Slope | - | slope of the long-term average spectrum |
| Tilt | - | tilt of the trend line through the long-term average |
| rap | - | relative average pertubation |
| ppq | - | pitch perturbation quotient |

| | | |
|---|---|---|
| dB | - | decibel |
| apq | - | amplitude perturbation quotient |
| mACF | - | mean autocorrelation |
| NHR | - | noise-to-harmonics ratio |
| CPP | - | cepstral peak prominence |
| CSL | - | Computerized Speech Lab |
| SPL | - | Sound pressure level |
| MDVP | - | Multi Dimensional Voice Program |
| MCT | - | Manual circumlaryngeal therapy |
| MTD | - | Muscle tension dysphonia |
| SLP | - | Speech language pathologist |
| LC | - | Lyapunov Coefficient |
| ESGP | - | Estimated subglottic pressure |
| EVA | - | Evaluation Vocal Assistee |
| k-NN | - | k neural network |
| VASS | - | Voice Analysis and Screening System |
| TNI | - | Turbulent Noise Index |
| NFHE | - | Normalized First Harmonic Energy |
| $HNR_Y$ | - | Harmonics-to-noise ratio in time domain |
| $HNR_Q$ | - | Harmonics-to-noise ratio in spectral domain |
| DH | - | Degree of hoarseness |
| NNE | - | Normalized noise energy |
| TNI | - | Turbulent noise index |
| NFHE | - | Ratio of the first harmonic energy to the rest of harmonics |
| LDA | - | Linear discriminant analysis |
| MFCC | - | Mel-frequency cepstral coefficients |
| MLP | - | Multilayer perceptron |
| LVQ | - | Learning vector quantization |

| | | |
|---|---|---|
| GCI | - | Glottal closure instant |
| ACF | - | Autocorrelation function |
| AMDF | - | Average magnitude difference function |
| LPC | - | Linear predictive coefficient |
| WT | - | Wavelet transform |
| HHT | - | Hilbert-Huang transform |
| EMD | - | Empirical mode decomposition |
| IMF | - | Intrinsic mode function |
| AC | - | Autocorrelation |
| YIN AC | - | YIN autocorrelation |
| SIFT | - | Simple inverse filtering tracking |
| CEP | - | Cepstrum |
| WAV | - | Wavelet |
| CHAOS | - | Chaotic time series |
| EGG | - | Electroglottographic |
| U/V | - | Unvoiced/Voiced |
| V/U | - | Voiced/Unvoiced |
| FFT | - | Fast Fourier transform |
| SR | - | Sampling rate |
| AR | - | Autoregressive |
| DME | - | Dynamic Mean Evaluation |
| LP | - | Linear predictive |
| PSD | - | Power spectral density |
| SVD | - | Singular value decomposition |
| STFT | - | Short Time Fourier Transform |
| CWT | - | Continuous wavelet transform |
| DWT | - | Discrete wavelet transform |
| MPPD | - | Modified pitch period detection |
| PPD | - | Pitch period detection |

| | | |
|---|---|---|
| IMF | - | Intrinsic mode functions |
| FIE | - | Instantaneous energy density level |
| HOS | - | Higher order statistic |
| SNR | - | Signal-to-noise ratio |
| NACC | - | Normalized autocorrelation function of the 1 dimensional slice of the third order cumulants |
| FOC | - | Fourth-order cumulants |
| EF | - | Energy |
| CF | - | Cumulants |
| STE | - | Short time energy |
| RMSE | - | Root mean square energy |
| Pa | - | Pascal |
| rms | - | Root mean square |
| *MAX_PER* | - | Maximum pitch period of human speech |
| ZCR | - | Zero crossing rate |
| MSF | - | Magnitude sum function |
| dB A | - | The curve A that represents the characteristic hearing curve of the human ear. Lower sounds seem quieter than middle or high frequency sounds to the human ear |
| VTI | - | Voice turbulence index |
| CBE | - | Centre for Biomedical Engineering Universiti Teknologi Malaysia |

# LIST OF SYMBOLS

| | | |
|---|---|---|
| $F_0$ | - | Fundamental frequency/pitch |
| $f_0$ | - | Fundamental frequency/pitch |
| $T_0$ | - | Fundamental period/pitch period |
| $R_x$ | - | Autocorrelation value |
| $s[n]$ | - | Signal value at sample $n$ |
| $N$ | - | Frame size |
| $l$ | - | Lag |
| $R_y$ | - | Average magnitude difference function value |
| X2 | - | Doubling error |
| /2 | - | Halving error |
| G error | - | Gross error |
| F error | - | Fine error |
| S error | - | Standard deviation error |
| $f_H$ | - | Highest frequency |
| $F_H$ | - | Real value for highest frequency |
| $N_H$ | - | Consecutive number of the highest harmonic peak |
| $Q_N$ | - | the $K^{th}$ root of the multiplication of spectral coefficients (amplitudes) |
| L | - | Number of points used for the FFT |
| X(f) | - | Log spectrum |
| S(f) | - | Cepstral smoothed log spectrum |
| $\eta_{min}$ | - | Lowest value of the AMDF within the frame |

| | | |
|---|---|---|
| $\sigma$ | - | Standard deviation of AMDF of the frame |
| $p$ | - | Order of AR models |
| $CO_2$ | - | Carbon dioxide |
| $E(n)$ | - | Energy at time $n$ |
| *ITL* | - | Lower threshold |
| *ITU* | - | Upper threshold |
| *IZCT* | - | Zero crossing threshold |
| $I$ | - | Intensity |
| $x$ | - | Pressure units of Pascal |
| $p_{rms}$ | - | root mean square (rms) pressure |
| $p_{ref}$ | - | Reference sound pressure in air considered as threshold of human hearing (0dB) at 1 kHz |
| *IMX* | - | Maximum energy |
| *IMN* | - | Minimum energy |
| $f_s$ | - | Sampling frequency |
| sgn | - | Sign |

# LIST OF APPENDICES


| APPENDIX | TITLE | PAGE |
|---|---|---|
| A | Malaysian Statistics on Medical Devices 2007 | 159 |
| B | Voice Handicap Index | 160 |
| C | KayPENTAX Voice Database | 163 |
| D | Microphone for Data Collection | 174 |

# CHAPTER 1

# INTRODUCTION

## 1.1 Background of Research

Maier *et al*. (2009) stated that 87.5% of people who live in urban area need communication for their daily work. These people are dependent upon their voices for their career such as teachers, lawyers and other professionals. A loss of $154 billion to $186 billion per year to the United States economy is due to the cost of care and the degradation of the employment opportunities for the people with communication disorder. Because of the communication disabilities, they do not have many opportunities for employment.

Voice disorder is one of the speech disorders, which is increasing dramatically nowadays (Maier *et al*. 2009). Voice disorder can be caused by unhealthy social habits such as smoking and alcohol consumption, voice abuse for example shouting inappropriately, and most importantly the lack of awareness among the general public and the health care provider (Godino Llorente *et al*., 2006). Surveys made by Behrman (2005) to 53 speech therapists with at least three years experience in using stroboscopy and acoustic instrumentation discovered that voice therapy are performed as much as 4.9 cases per week on the average. Since the voice disorders can occur anytime due to various causes, prevention measures need to be taken because the problem can become permanent if not treated (Van Lierde *et al.*, 2009c).

Throughout the years, different types of clinical procedures, objective and subjective measurement devices have been developed for voice assessment and therapy in order to prevent, to reduce, and perhaps to cure the severity and the occurrence of voice problems. Different assessment techniques provide different information for the speech therapists. Among them are video laryngoscopy, GRBAS (a short for Grade, Roughness, Breathiness, Asthenia, and Strain parameters), Acoustic Voice Quality Index (AVQI), Voice Handicap Index (VHI), and Dysphonia Severity Index (DSI).

This research will focus on the development of a single platform that can perform DSI to evaluate the degree of voice disorder. A single DSI evaluation platform can be built by the integration of pitch detection algorithm (PDA), start/endpoint detection algorithm, jitter equation, and intensity equation. A new PDA is developed to detect pitch for both normal and pathological voices with higher accuracy than the existing algorithms. Experimental results are evaluated on each of the mentioned algorithm, and also on the overall developed system.

## 1.2 Problem Statement

Voice assessment generally includes subjective and objective approaches. Subjective measurement such as auditory perceptual analysis is one of the methods of diagnosing the existence or the severity of a dysphonia (Johnson, 2007). But study shows that this approach will cause variety in the results and the assessment can be questionable if the clinicians did not come out with a standardized procedure

(Tomblin *et al*., 2000) (Johnson, 2007). This problem has ignited the researchers to develop objective measurement techniques or system in order to assist the speech pathologists assessing the voice disorders.

A survey was made by National Medical Device Survey published in the publication of Engineering Services Division, Medical Device Bureau and the Clinical Research Centre, Ministry of Health Malaysia has come out with the table as presented in Appendix A (Ariza *et al*., 2007). It shows that the amount of devices for otorhinolaryngology and audiology in Malaysia are not capable to cover the overall population in Malaysia. Moreover, the access to these devices is concentrated in Selangor and Kuala Lumpur. Details of the survey are included in the Appendix A.

Until today, most of the acoustic analysis systems are owned by the speech pathology clinics and by speech therapists (Godino-Llorente *et al*., 2006) (Johnson, 2007). General public do not have the access to these systems and they cannot monitor their progress or improvement during their rehabilitation at home and even the public or the patients do not have access to the system. The patients will find difficulties to process their own voices as the system is complicated with everything operated manually and usually conducted by the specialists or speech therapists. This situation has lead to the development of an automatic system to diagnose voice disorders that can be used easily by the patients themselves. The automatic approach in diagnosing voice requires a non-invasive and objective technique so that the patient can use it with ease at home during rehabilitation.

The method that is used in this thesis is based on an objective multiparametric approach known as Dysphonia Severity Index (DSI). Until now, this technique is widely used for various types of voice research but none of them used a single platform to obtain the DSI. Therefore, the result obtained by a research is incomparable with the result by other researches because of the non-standardized platforms used by each research. The developed system presented by this thesis is aimed to overcome:

i) the problems by subjective voice assessment approach
ii) the lack of assessment tools in Malaysia
iii) the inefficiency of home rehabilitation at own responsibility
iv) the non-standardized platforms used to calculate DSI

## 1.3 Objectives

The thesis aims to achieve the three objectives outlined in this section.

Firstly, the main objective of this project is to develop a voice disorder evaluation system based on Dysphonia Severity Index (DSI) named as Automatic Dysphonia Evaluation System (ADES) throughout this thesis.

Secondly, to find and to develop the best and accurate way of Pitch Detection Algorithm (PDA) to detect the voice pitches especially to detect disordered voice for diagnosing purposes.

The final objective is to indentify, to modify and to assemble appropriate algorithms so that system will be able to diagnose the voice disorder objectively and automatically based on DSI.

### 1.4 Scope of the Research

1. The evaluation system will implement Dysphonia Severity Index (DSI) that requires the combination of pitch detection algorithm (PDA), an algorithm that can find the voice intensity, an effective start/end point detection algorithm and an algorithm to calculate the percentage jitter.
2. The type of speech disorder that will be focused in the thesis is the voice disorder.
3. The PDA accuracy will be tested using KayPENTAX database for benchmarking with the existing time domain PDA's such as MATLAB Mathworks' Pitch Synchronize Overlap and Add (PSOLA) PDA, Praat software's autocorrelation and crosscorrelation, and average magnitude difference function by Manfredi *et al*. (2000).
4. KayPENTAX normal and pathological voices are used for ADES' accuracy test.
5. The start/endpoint detection algorithm will only concern the finding of the start and the end points of the vowel /a/ utterance (voiced) with silence as the background and will not concern the voiced/unvoiced classification.

### 1.5 Thesis Outline

This thesis is divided into six chapters. Chapter 1 includes the introduction, the background, the objective and the scope of the thesis. The main aim is to show how this research will overcome the problem described in the problem statement. It will also discuss the scope of the research and its feasibility.

Chapter 2 presents the background and the literature of speech therapy, the types of speech disorders and how they are treated. The main interest which is voice disorder will be explained further. Chapter 2 also describes several techniques of objective multiparameter that have been developed in the literatures and how finally DSI was chosen as the objective multiparameter technique to be used for the development of the research. It also provides the comprehensive study on the technical aspects and the literatures on PDA, speech start/endpoint detection and how these algorithms are combined with jitter and intensity algorithm to form the DSI assessment platform.

Chapter 3 describes the research methodology and the design of the developed PDA and modified start/endpoint detection for ADES. This chapter shows how jitter and intensity equations are integrated to obtain the maximum phonation time, jitter percentage, lowest intensity and highest frequency for the DSI calculation. This chapter also describes the experiments involved to test the accuracy of the developed system.

Chapter 4 presents results and the discussion of all different algorithms for PDA and start/endpoint detection in comparison to the ones that have been designed and developed for ADES. This chapter also includes the evaluation of ADES by using the voices from KayPENTAX database, the collected database of trained and untrained vocalists, and teachers and non-teachers.

Chapter 5 concludes the works and findings of this thesis. Some suggestions for future approach and enhancement are also given.

## 1.6 Contribution of Thesis

The development of a single platform to evaluate the DSI itself is a new invention for voice assessment system. Until today, each of the DSI parameter is evaluated by using different devices which are mostly from KayPENTAX, a world leading company for speech, voice, and swallowing instrumentation.

Another contribution is the new developed PDA which can detect the pitch without fixing the frame length according to a person's voice with high accuracy. Even the PDA developed for KayPENTAX's Multi-Dimensional Voice Program requires the user to select the pitch searching range for pitch detection (Deliyski, 1993). This feature is usually included in many PDA so that the algorithm will not wrongly detect the pitch. However, the proposed PDA in this thesis performs pitch detection within the pitch searching range of 50 Hz to 1000 Hz without requiring the user to limit the pitch searching range.

**REFERENCES**

Abdullah-Al-Mamun, K., Sarker, F., and Muhammad, G. (2009). A High Resolution Pitch Detection Algorithm Based on AMDF and ACF. Journal of Scientific Research. 1(3). 508-515.

Ahmadi, S., and Spanias, A. S. (1999). Cepstrum-Based Pitch Detection Using a New Statistical V/UV Classification Algorithm. Speech and Audio Processing, IEEE Transactions. 7(3). 333-338.

Ai, O. C., and Yunus, J. (2007). Computer-based System to Assess Efficacy of Stuttering Therapy Techniques. *3rd Kuala Lumpur International Conference on Biomedical Engineering 2006, IFMBE Proceedings*. 10 (15). 374-377.

Alkulaibi, A., Soraghan, J. J., and Durrani T. S. (1997). Fast 3-level Binary Higher Order Statistics for Simultaneous Voiced/Unvoiced and Pitch Detection of a Speech Signal. *Signal Processing*. 63(2), 133-140.

Amado, R. G., and Filho, J. V. (2008). Pitch Detection Algorithms Based on Zero-Cross Rate and Autocorrelation Function for Musical Notes. International Conference on Audio, Language and Image Processing. 449-454.

Amir, O., Amir N., and Michaeli, O. (2005). Evaluating the Influence of Warmup on Singing Voice Quality Using Acoustic Measures. *Journal of Voice*. 19 (2), 252-260.

Awan, S. N. and Ensslen, A. J. (2010). A Comparison of Trained and Untrained Vocalists on the Dysphonia Severity Index. *Journal of Voice*. 24 (6), 661-666.

Beech, J. R., Harding, L., and Hilton-Jones, D. (1993). Assessment in Speech and Language Therapy. New Fetter Lane, London: Routledge

Behrman, A. (2005). Common Practices of Voice Therapists in the Evaluation of Patients. *Journal of Voice*. 19 (3), 454-469.

Boersma, P. (2009). Should Jitter be Measured by Peak Picking or by Waveform Matching. Folia Phoniatrica et Logopaedica. 61. 305-308.

Boersma, P., & Weenink, D. (2011). Praat: doing phonetics by computer [Computer program]. Version 5.2.10, retrieved 11 January 2011 from http://www.praat.org/

Charles, V. R., and Robert, L. E. (1996). Speech Correction: An Introduction to Speech Pathology and Audiology. (9th ed.) Needham Heights: Ally & Bacon.

Chu, W. C. (2003). Speech Coding Algorithms: Foundation and Evolution of Standardized Coders. Hoboken, New Jersey: John Wiley and Sons, Inc.

De Bodt, M. S., Ketelslagers, K., Peeters, T., Wuyts, F. L., Mertens, F., Pattyn, J., Heylen, L., Peeters, A., Boudewyns, A. and Van de Heyning, P. (2007). Evolution of Vocal Fold Nodules from Childhood to Adolescence. *Journal of Voice*. 21 (2), 151-156.

De Cheveigne, A. and Kawahara, H. (2002). YIN, a Fundamental Frequency Estimator for Speech and Music. *The Journal of the Acoustical Society of America*. 111(4). 1917-1930.

Deliyski, D. D. (1993). Acoustic Model and Evaluation of Pathological Voice Production. *3rd Conference on Speech Communication and Technology Eurospeech '93*, 1993. 3. 1969-1972.

Dong-Yan, H., Weisi, L. and Susanto, R. (2004). Speech pitch detection in noisy environment using multi-rate adaptive lossless FIR filters. *Circuits and Systems, 2004. ISCAS '04. Proceedings of the 2004 International Symposium on*. 23-26 May 2004. III-429-432 Vol.423.

Dueñas, W. R. R., Vaquero, C., Saz, O., and Lleida, E. (2008). Speech Technology Applied to Children with Speech Disorders. *4th Kuala Lumpur International Conference on Biomedical Engineering*. 21. 247–250.

Duffy, O. M. and Hazlett, D. E. (2004). The impact of preventive voice care programs for training teachers: A longitudinal study. *Journal of Voice*. 18 (1), 63-70.

Gelfer, M. P. (1996). Survey of Communication Disorders: A Social and Behavioral Perspective, McGraw-Hill Companies, Inc.

Gelzinis, A., Verikas, A. and Bacauskiene, M. (2008). Automated speech analysis applied to laryngeal disease categorization. *Computer Methods and Programs in Biomedicine*. 91 (1), 36-47.

Godino-Llorente, J. I. and Gomez-Vilda, P. (2004). Automatic detection of voice impairments by means of short-term cepstral parameters and neural network based detectors. *Biomedical Engineering, IEEE Transactions on*. 51 (2), 380-384.

Godino-Llorente, J. I., Sáenz-Lechón, N., Osma-Ruiz, V., Aguilera-Navarro, S. and Gómez-Vilda, P. (2006). An integrated tool for the diagnosis of voice disorders. *Medical Engineering & Physics*. 28 (3), 276-289.

Hadjitodorov, S. and Mitev, P. (2002). A computer system for acoustic analysis of pathological voices and laryngeal diseases screening. *Medical Engineering & Physics*. 24 (6), 419-429.

Hagmüller, M. and Kubin, G. (2006). Poincaré pitch marks. *Speech Communication*. 48 (12), 1650-1665.

Hakkesteegt, M. M., Brocaar, M. P. and Wieringa, M. H. (2010). The Applicability of the Dysphonia Severity Index and the Voice Handicap Index in Evaluating Effects of Voice Therapy and Phonosurgery. *Journal of Voice*. 24 (2), 199-205.

Hakkesteegt, M. M., Brocaar, M. P., Wieringa, M. H. and Feenstra, L. (2008). The Relationship Between Perceptual Evaluation and Objective Multiparametric Evaluation of Dysphonia Severity. *Journal of Voice*. 22 (2), 138-145.

Haynes, W. O., Pindzola, R. H. and Emerick, L. L. (1992). Diagnosis and Evaluation in Speech Pathology. Englewood Cliffs, New Jersey, Prentice-Hall, Inc.

Henry, L. R., Helou, L. B., Solomon, N. P., Howard, R. S., Gurevich-Uvena, J., Coppit, G., and Stojadinovic, A. (2010). Functional Voice Outcomes After Thryroidectomy: An Assessment of the Dysphonia Severity Index (DSI) After Thyroidectomy. Surgery.

Hu, W., Wang, X., Liang, Y. and Du, M. (2005). A Novel Pitch Period Detection Algorithm Bases on HHT with Application to Normal and Pathological Voice. *Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the*. 17-18 Jan. 2006. 4541-4544.

Huang, H. and Pan, J. (2006). Speech pitch determination based on Hilbert-Huang transform. *Signal Processing*. 86 (4), 792-803.

Jinhai, C. and Zhi-Qiang, L. (1997). Robust pitch detection of speech signals using steerable filters. *Acoustics, Speech, and Signal Processing, 1997. ICASSP-97., 1997 IEEE International Conference on*. 21-24 April 1997. 1427-1430 vol.1422.

John, R. D., John, G. P., John, H. L. H. (1993). Discrete-Time Processing of Speech Signal. US: Macmillan, Inc.

Johnson, A. F. (2007). *Medical speech-language pathology: a practitioner's guide*. (2nd ed.) Thieme.

Kang, G. and Guo, S. (2009). Improving AMDF for pitch period detection. *Electronic Measurement & Instruments, 2009. ICEMI '09. 9th International Conference on*. 16-19 Aug. 2009. 4-283-284-286.

Kotnik, B., Höge, H. and Kacic, Z. (2009). Noise robust F0 determination and epoch-marking algorithms. *Signal Processing*. 89 (12), 2555-2569.

Ma, E. P. M. and Yiu, E. M. L. (2006). Multiparametric Evaluation of Dysphonic Severity. *Journal of Voice*. 20 (3), 380-390.

Maier, A., Haderlein, T., Eysholdt, U., Rosanowski, F., Batliner, A., Schuster, M. and Nöth, E. (2009). PEAKS - A system for the automatic evaluation of voice and speech disorders. *Speech Communication*. 51 (5), 425-437.

Manfredi, C., D'Aniello, M., Bruscaglioni, P. and Ismaelli, A. (2000). A comparative analysis of fundamental frequency estimation methods with application to pathological voices. *Medical Engineering & Physics*. 22 (2), 135-147.

Mangayyagari, S. and Sankar, R. (2007). Pitch conversion based on pitch mark mapping. *SoutheastCon, 2007. Proceedings. IEEE*. 22-25 March 2007. 8-13.

Marks, J. A. (1988). Real time speech classification and pitch detection. *Communications and Signal Processing, 1988. Proceedings., COMSIG 88. Southern African Conference on*. 24 Jun 1988. 1-6.

Maryn, Y., De Bodt, M. and Roy, N. (2009). The Acoustic Voice Quality Index: Toward improved treatment outcomes assessment in voice disorders. *Journal of Communication Disorders*. 43 (3), 161-174.

Mitev, P. and Hadjitodorov, S. (2003). Fundamental frequency estimation of voice of patients with laryngeal disorders. *Information Sciences*. 156 (1-2), 3-19.

Moran, R. J., Reilly, R. B., de Chazal, P. and Lacy, P. D. (2006). Telephony-based voice pathology assessment using automated speech analysis. *Biomedical Engineering, IEEE Transactions on*. 53 (3), 468-477.

Omar, R., Nasir, A. M. M., Sabirin, J., Awang, M. A. and Zawawi, N. S. M. (2008). Othorhinolaryngology and Audiology Facilities and Devices. Malaysian Statistic on Medical Devices 2007. A. Zakaria, F. A. M. Yusof and L. T. O. Kuala Lumpur**:** 61.

Paul Christopher Bagshaw. Automatic Prosodic Analysis for Computer Aided Pronunciation Teaching. Ph. D. Thesis. The University of Edinburgh; 1994

Quatieri, T. F. (2002). Discrete-Time Speech Signal Processing Principles and Practice. (1st ed.) Upper Sadle River, N. J.: Prentice-Hall.

Rabiner, L. R. and Sambur, M. R. (1975). An Algorithm for Determining the Endpoints of Isolated Utterances. *The Bell System Technical Journal*. 54(2). 297-315.

Rabiner, L., Cheng, M., Rosenberg, A. and McGonegal, C. (1976). A comparative performance study of several pitch detection algorithms. *Acoustics, Speech and Signal Processing, IEEE Transactions on*. 24 (5), 399-418.

Rodríguez-Parra, M. J., Adrián, J. A. and Casado, J. C. (2009). Voice Therapy Used to Test a Basic Protocol for Multidimensional Assessment of Dysphonia. *Journal of Voice*. 23 (3), 304-318.

Rosemary, O. (2005). Methods of Voice Analysis for Estimating the Robustness of the Student Teacher's Voice. Ph. D. Thesis. Trinity College, Dublin, Ireland.

Ru-wei, L., Chang-chun, B. and Hui-jing, D. (2008). Pitch detection method for noisy speech signals based on pre-filter and weighted wavelet coefficients. *Signal Processing, 2008. ICSP 2008. 9th International Conference on*. 26-29 Oct. 2008. 530-533.

Schoentgen, J. (2003). Decomposition of Vocal Cycle Length Perturbations into Vocal Jitter and Vocal Microtremor, and Comparison of Their Size in Normophonic Speakers. *Journal of Voice*. 17 (2), 114-125.

Schoentgen, J. and De Guchteneere, R. (1997). Predictable and random components of jitter. *Speech Communication*. 21 (4), 255-272.

Seung-Jin, J., Seong-Hee, C., Hyo-Min, K., Hong-Shik, C. and Young-Ro, Y. (2007). Evaluation of Performance of Several Established Pitch Detection Algorithms in Pathological Voices. *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*. 22-26 Aug. 2007. 620-623.

Shahnaz, C., Zhu, W. P. and Ahmad, M. O. (2007). A Pitch Detection Method for Speech Signals with Low Signal-to-Noise Ratio. *Signals, Systems and Electronics, 2007. ISSSE '07. International Symposium on*. July 30 2007-Aug. 2 2007. 399-402.

SM48 Cardioid Dynamic Vocal Microphone. http://www.shureasia.com/products/microphones/sm48 (Accessed on 19 November 2011).

Stevens, K. N. (1999). Acoustic Phonetics. Massachusetts Institute of Technology, United States of America

Tabrikian, J., Dubnov, S. and Dickalov, Y. (2004). Maximum a-posteriori probability pitch tracking in noisy environments using harmonic model. *Speech and Audio Processing, IEEE Transactions on*. 12 (1), 76-87.

Tan Tian Swee. Corpus-based Malay Text to Speech Synthesis System. Ph. D. Thesis. Universiti Teknologi Malaysia; 2009

Tan Tian Swee. The Design and Verification of Malay Text to Speech Synthesis System. Masters Thesis. Universiti Teknologi Malaysia; 2004

Tan, T. S., Salleh, S. H. S. and Jamaludin, M. R. (2010). Speech pitch detection using short-time energy. *Computer and Communication Engineering (ICCCE), 2010 International Conference on*. 11-12 May 2010. 1-6.

Tian-Swee, T., Helbin, L., Ariff, A. K., Chee-Ming, T. and Salleh, S. H. (2007). Application of Malay speech technology in Malay Speech Therapy Assistance Tools. *Intelligent and Advanced Systems, 2007. ICIAS 2007. International Conference on*. 25-28 Nov. 2007. 330-334.

Timmermans, B., De Bodt, M. S., Wuyts, F. L., Boudewijns, A., Clement, G., Peeters, A. and Van de Heyning, P. H. (2002). Poor Voice Quality in Future Elite Vocal Performers and Professional Voice Users. *Journal of Voice*. 16 (3), 372-382.

Timmermans, B., De Bodt, M. S., Wuyts, F. L., Boudewijns, A., Clement, G., Peeters, A. and Van de Heyning, P. H. (2002). Poor Voice Quality in Future Elite Vocal Performers and Professional Voice Users. *Journal of Voice*. 16 (3), 372-382.

Ting, H. N., Yunus, J., Vandort, S., and Salleh, S. H. S. (2004). Computer-based Speech Training for Children. *International Journal Computer Applications in Technology*. 21. 52-57.

Tomblin, J. B., Morris, H. L., and Spriestersbach, D. C. (1994). Diagnosis in Speech-Language Pathology. (1st ed.) San Diego, California: Singular Publishing Group, Inc.

Tomblin, J. B., Morris, H. L., and Spriestersbach, D. C. (2000). Diagnosis in Speech-Language Pathology. (2nd ed.) San Diego, California: Singular Publishing Group, Inc.

Uloza, V., Saferis, V. and Uloziene, I. (2005). Perceptual and Acoustic Assessment of Voice Pathology and the Efficacy of Endolaryngeal Phonomicrosurgery. *Journal of Voice*. 19 (1), 138-145.

Van Lierde, K. M., Bodt, M. D., Dhaeseleer, E., Wuyts, F. and Claeys, S. (2010a). The Treatment of Muscle Tension Dysphonia: A Comparison of Two Treatment Techniques by Means of an Objective Multiparameter Approach. *Journal of Voice*. 24 (3), 294-301.

Van Lierde, K. M., Claeys, S., De Bodt, M. and van Cauwenberge, P. (2007). Long-Term Outcome of Hyperfunctional Voice Disorders Based on a Multiparameter Approach. *Journal of Voice*. 21 (2), 179-188.

Van Lierde, K. M., Claeys, S., De Bodt, M. and Van Cauwenberge, P. (2004). Vocal quality characteristics in children with cleft palate: a multiparameter approach. *Journal of Voice*. 18 (3), 354-362.

Van Lierde, K. M., Claeys, S., Dhaeseleer, E., Deley, S., Derde, K., Herregods, I., Strybol, I. and Wuyts, F. (2010b). The Vocal Quality in Female Student Teachers During the 3 Years of Study. *Journal of Voice*. 24 (5), 599-605.

Van Lierde, K. M., De Ley, S., Clement, G., De Bodt, M. and Van Cauwenberge, P. (2004). Outcome of laryngeal manual therapy in four Dutch adults with persistent moderate-to-severe vocal hyperfunction: a pilot study. *Journal of Voice*. 18 (4), 467-474.

Van Lierde, K. M., D'Haeseleer, E., Wuyts, F. L., De Ley, S., Geldof, R., De Vuyst, J. and Sofie, C. (2010c). The Objective Vocal Quality, Vocal Risk Factors, Vocal Complaints, and Corporal Pain in Dutch Female Students Training to be Speech-Language Pathologists During the 4 Years of Study. *Journal of Voice*. 24 (5), 592-598.

Van Lierde, K. M., Vinck, B., De Ley, S., Clement, G. and Van Cauwenberge, P. (2005). Genetics of Vocal Quality Characteristics in Monozygotic Twins: A Multiparameter Approach. *Journal of Voice*. 19 (4), 511-518.

Van Riper, C., and Erickson, R. L. (1996). Speech Correction: An Introduction to Speech Pathology and Audiology. (9th ed.) Needham Heights: Ally & Bacon.

Vasilakis, M. and Stylianou, Y. (2009). Spectral jitter modeling and estimation. *Biomedical Signal Processing and Control*. 4 (3), 183-193.

Verikas, A., Gelzinis, A., Bacauskiene, M. and Uloza, V. (2006). Towards a computer-aided diagnosis system for vocal cord diseases. *Artificial Intelligence in Medicine*. 36 (1), 71-84.

Verikas, A., Gelzinis, A., Bacauskiene, M. and Uloza, V. (2010). Towards noninvasive screening for malignant tumours in human larynx. *Medical Engineering & Physics*. 32 (1), 83-89.

Verikas, A., Gelzinis, A., Valincius, D., Bacauskiene, M. and Uloza, V. (2007). Multiple feature sets based categorization of laryngeal images. *Computer Methods and Programs in Biomedicine*. 85 (3), 257-266.

Voice Handicap Index. http://www.phillyent.com/appointments/pdf/VoiceHandicapIndex.pdf (Accessed on 31 May 2011).

Wanda, W. Z., and Tokunbo, O. (1999). Formant and Pitch Detection Using Time-Frequency Distribution. *International Journal of Speech Technology*. 3. 35-49.

Wikipedia. http://en.wikipedia.org/wiki/Sound_pressure (Accessed on 31 May 2011).

Woisard, V., Bodin, S., Yardeni, E. and Puech, M. (2007). The Voice Handicap Index: Correlation Between Subjective Patient Response and Quantitative Assessment of Voice. *Journal of Voice*. 21 (5), 623-631.

Xi, S., Changsheng, X. and Kankanhalli, M. S. (2006). Predominant Vocal Pitch Detection in Polyphonic Music. *Multimedia and Expo, 2006 IEEE International Conference on*. 9-12 July 2006. 897-900.

Ying, G. S., Mitchell, C. D. and Jamieson, L. H. (1993). Endpoint detection of isolated utterances based on a modified Teager energy measurement. *Acoustics, Speech, and Signal Processing, 1993. ICASSP-93., 1993 IEEE International Conference on*. 27-30 April 1993. 732-735 vol.732.

Yu, P., Ouaknine, M., Revis, J. and Giovanni, A. (2001). Objective Voice Analysis for Dysphonic Patients: A Multiparametric Protocol Including Acoustic and Aerodynamic Measurements. *Journal of Voice*. 15 (4), 529-542.

Zhen-Dong, Z., Xi-Mei, H. and Jing-Feng, T. (2008). An effective pitch detection method for speech signals with low signal-to-noise ratio. *Machine Learning and Cybernetics, 2008 International Conference on*. 12-15 July 2008. 2775-2778.

⇒ **“vAm”** - vAm /%/ - Peak Amplitude Variation represents the relative standard deviation of the period-to-period calculated peak-to-peak amplitude. It reflects the very long-term amplitude variations within the analyzed voice sample.

⇒ **“NHR”** - NHR - Noise-to-Harmonic Ratio is an average ratio of energy of the in-harmonic components in the range 1500-4500 Hz to the harmonic components energy in the range 70-4500 Hz. It is a general evaluation of the noise present in the vocalization.

⇒ **“VTI”** - VTI - Voice Turbulence Index is an average ratio of the spectral in-harmonic high-frequency energy to the spectral harmonic energy in stable phonation areas. VTI measures the relative energy level of high-frequency noise, such as turbulence.

⇒ **“SPI”** - SPI - Soft Phonation Index is an average ratio of the lower- frequency to the higher-frequency harmonic energy. This index is not a measurement of abnormality but rather a measurement of the spectral “type” of the vocalization.

⇒ **“FTRI”** - FTRI /%/ - Fo-Tremor Intensity Index shows (in percent) the ratio of the frequency magnitude of the most intensive low-frequency modulating component (Fo-tremor) to the total frequency magnitude of the analyzed voice signal.

⇒ **“ATRI”** - ATRI /%/ - Amplitude Tremor Intensity Index shows (in percent) the ratio of the amplitude of the most intensive low-frequency amplitude-modulating component (amplitude tremor) to the total amplitude of the analyzed voice signal.

⇒ **“DVB”** - DVB /%/ - Degree of Voice Breaks shows (in percent) the ratio of the total length of areas representing voice breaks to the time of the complete voice sample.

⇒ **“DSH”** - DSH /%/ - Degree of sub-harmonics is an estimated relative evaluation of sub-harmonic to Fo components in the voice sample.

⇒ **“DUV”** - DUV /%/ - Degree of Voiceless is an estimated relative evaluation of non-harmonic areas (where Fo can not be detected) in the voice sample. DUV considers as voiceless all pauses either before, after, and/or between the voiced areas.

⇒ **“NVB”** - NVB - Number of Voice Breaks shows how many times the generated Fo was interrupted from the beginning of the first until the end of the last voiced area.

⇒ **“NSH”** - NSH- Number of Sub-Harmonic Segments found during analysis.

⇒ **“NUV”** - NUV - Number of Unvoiced Segments detected during the autocorrelation analysis.

⇒ **“SEG”** - SEG - Total number of segments computed during the MDVP-autocorrelation analysis.

⇒ **“PER”** - PER - Pitch Periods detected during the period-to-period pitch extraction using MDVP.

**3.2. Patient' Information Spreadsheet** - provides patient' clinical information for each subject. It is the same as section 3.1 but does not include the MDVP-results. This spreadsheet can be used for quick access or processing of subject information part of the database when MDVP results are not needed. It contains 1689 rows, but only the following 11 columns are included: **"PAT_ID"**, **"VISITDATE"**, **"FILE VOWEL 'ah'"**, **"AGE"**, **"SEX"**, **"#"**, **"DIAGNOSIS"**, **"LOCATION"**, **"SMOKE"**, **"NATLANG"**, **"ORIGIN".**

**3.3. Pathological Voice Spreadsheet** - provides MDVP-analysis results from each pathological voice recording of sustained vowel 'ah' included in the database. It is similar to 3.1 but does not include some of the clinical information such as diagnoses, site of disorder, patient' ID, etc. There are 655 rows in the spreadsheet, each one corresponding to an existing ***.nsp*** file from the CD-ROM. There is no repetition of filenames, as in section 3.1 and 3.2, caused by multiple diagnoses for the same subject. This spreadsheet can be used for quick access or processing of the MDVP-results. The following 39 columns are included: **"FILE VOWEL 'ah'"**, **"AGE"**, **"SEX"**, **"SMOKE"**, **"NATLANG"**, **"ORIGIN"**, **"Fo"**, **"To"**, **"Fhi"**, **"Flo"**, **"STD"**, **"PFR"**, **"Fftr"**, **"Fatr"**, **"Tsam"**, **"Jita"**, **"Jitt"**, **"RAP"**, **"PPQ"**, **"sPPQ"**, **"vFo"**, **"ShdB"**, **"Shim"**, **"APQ"**, **"sAPQ"**, **"vAm"**, **"NHR"**, **"VTI"**, **"SPI"**, **"FTRI"**, **"ATRI"**, **"DVB"**, **"DSH"**, **"DUV"**, **"NVB"**, **"NSH"**, **"NUV"**, **"SEG"**, **"PER".**

**3.4. Normal Voice Spreadsheet** - provides MDVP-analysis results from each normal voice recording of sustained vowel 'ah' included in the database. It is similar to section 3.1, but it does not include some of the clinical information such as diagnoses, site of disorder, patient' ID, etc. The spreadsheet contains 3 tables. The first table has 53 rows representing every existing normal voice ***.nsp*** file from the CD-ROM. The additional two rows at the end of the table give the average value and standard deviation for every MDVP parameter, calculated from the table using Excel commands. The other two tables are similar - one includes only the female subject MDVP-results from the first table, their average values, and standard deviations; the other table - only the male subject results. This spreadsheet can be used for quick access or processing of MDVP-results of normal voices. The following 39 columns are included: **"FILE VOWEL 'ah'"**, **"AGE"**, **"SEX"**, **"SMOKE"**, **"NATLANG"**, **"ORIGIN"**, **"Fo"**, **"To"**, **"Fhi"**, **"Flo"**, **"STD"**, **"PFR"**, **"Fftr"**, **"Fatr"**, **"Tsam"**, **"Jita"**, **"Jitt"**, **"RAP"**, **"PPQ"**, **"sPPQ"**, **"vFo"**, **"ShdB"**, **"Shim"**, **"APQ"**, **"sAPQ"**, **"vAm"**, **"NHR"**, **"VTI"**, **"SPI"**, **"FTRI"**, **"ATRI"**, **"DVB"**, **"DSH"**, **"DUV"**, **"NVB"**, **"NSH"**, **"NUV"**, **"SEG"**, **"PER".**

Press the 'Note' buttons to obtain on-screen information about every field in the spreadsheets.

Caution: If you are not using this file directly from the CD-ROM, you may wish to store a copy as backup before manipulating it (especially when using sorting functions).

The file ***readme.xls*** is an Excel 5.0 version of the current file.

## 4. EXCEL 3.0 FILES:

Kay Elemetrics Corp. also provides an Excel 3.0 version of the database.

Sub-directory **\EXCEL50\EXCEL30** contains the files ***kaycdall.xls*** ***kaycdinf.xls*** ***kaycdnor.xls***, ***kaycdpat.xls***, and ***readme30.xls*** which represent Excel 3.0 spreadsheets almost identical with the spreadsheets from the Excel 5.0 workbook ***kaycd_db.xls*** and the file ***readme.xls*** described in section 3.

The file ***kaycdall.xls*** includes the **Full Database Spreadsheet** described in section 3.1**.**, ***kaycdinf.xls*** includes the **Patient' Information Spreadsheet** described in section 3.2**.**, ***kaycdpat.xls*** includes the **Pathological Voice Spreadsheet** described in section 3.3**.**, and ***kaycdnor.xls*** includes the **Normal Voice Spreadsheet** described in section 3.4**.**

Cells with red dots in the upper right corner include Notes giving information on the current field.

Caution: If you are not using these files directly from the CD-ROM, you may wish to store a copy as backup before manipulating it (especially when using sorting functions).

The file ***readme30.xls*** is an Excel 3.0 version of the current file .

## 5. MICROSOFT WORD 6.0 FILES:

The CD-ROM database is also available in Microsoft Word 6.0 format. MS Word offers sorting, calculating and selecting tools for analysis of tables.

Directory **\EXCEL50\WORD60** contains the files ***kaycdinf.doc***, ***kaycdnor.doc***, ***kaycdpat.doc***, and ***readme.doc***. The first three files represent Microsoft Word 6.0 tables containing the Disordered Voice Database.

The file ***kaycdinf.doc*** includes a **Patient' Information Table** containing information identical to the Patient' Information Spreadsheet described in section 3.2**.** but in MS Word 6.0 format.

The file ***kaycdpat.xls*** includes a **Pathological Voice Table** similar to the Pathological Voice Spreadsheet described in section 3.3**.** The table includes the following fields: **"FILE VOWEL 'ah'"**, **"Fo"**, **"To"**, **"Fhi"**, **"Flo"**, **"STD"**, **"PFR"**, **"Fftr"**, **"Fatr"**, **"Jita"**, **"Jitt"**, **"RAP"**, **"PPQ"**, **"sPPQ"**, **"vFo"**, **"ShdB"**, **"Shim"**, **"APQ"**, **"sAPQ"**, **"vAm"**, **"NHR"**, **"VTI"**, **"SPI"**, **"FTRI"**, **"ATRI"**, **"DVB"**, **"DSH"**, **"DUV"**, **"NVB"**, **"NSH"**, **"NUV".**

The file ***kaycdnor.xls*** includes a **Normal Voice Table** similar to the Normal Voice Spreadsheet described in section 3.4**.** The table includes the following fields: **"FILE VOWEL 'ah'"**, **"Fo"**, **"To"**, **"Fhi"**, **"Flo"**, **"STD"**, **"PFR"**, **"Fftr"**,

**“Fatr”, “Jita”, “Jitt”, “RAP”, “PPQ”, “sPPQ”, “vFo”, “ShdB”, “Shim”, “APQ”, “sAPQ”, “vAm”, “NHR”, “VTI”, “SPI”, “FTRI”, “ATRI”, “DVB”, “DSH”, “DUV”, “NVB”, “NSH”, “NUV”.**

Caution: If you are not using this file directly from the CD-ROM, you may wish to store a copy as backup before manipulating it (especially when using sorting functions).

The file ***readme.doc*** is the current file.

## 6. TEXT (ASCII) FILES:

Kay Elemetrics Corp. also provides a Text (ASCII) version of the database.

Sub-directory **\EXCEL50\TEXT** contains the files ***kaycdall.txt***, ***kaycdinf.txt***, ***kaycdnor.txt***, ***kaycdpat.txt***, and ***readme.txt*** which represent ASCII tab-delimited versions of the spreadsheets from the Excel 5.0 workbook ***kaycd_db.xls*** and ***readme.xls*** described in section 3.

The file ***kaycdall.txt*** includes the **Full Database Data** described in section 3.1**.**, ***kaycdinf.txt*** includes the **Patient’ Information Data** described in section 3.2**.**, ***kaycdpat.txt*** includes the **Pathological Voice Data** described in section 3.3**.**, and ***kaycdnor.txt*** includes the **Normal Voice Data** described in section 3.4**.**

The file ***readme.txt*** is a Text (ASCII) version of the current file.

## 7. ACOUSTIC VOICE SAMPLE FILES (VOWEL ‘AH’)

Along with the spreadsheets and tables allowing easy visual access, sorting and statistical analysis, the CD-ROM contains the actual Acoustic Voice Sample Files - sustained vowel ‘ah’.

Directory **\PATHOL\AH** contains 657 files with extension ***.nsp***. They represent Kay Elemetrics CSL format recordings of one second sustained vowel ‘ah’ from patients with a wide variety of organic, neurological, traumatic, and psychogenic voice disorders described in the database. These files were collected at the Massachusetts Eye and Ear Infirmary (MEEI) Voice and Speech Lab, Boston, MA.

Directory **\NORM\AH** contains 53 files with extension ***.nsp***. They represent Kay Elemetrics CSL-format recordings of one-second sustained vowel ‘ah’ from normal subjects. These files were collected at both Kay Elemetrics Corp. and Massachusetts Eye and Ear Infirmary (MEEI) Voice and Speech Lab.

The ***.nsp*** files can be played, edited and analyzed using MDVP or CSL with a 4300B, 4300, or 3300 hardware platform.

## 8. MDVP Parameter Files (of Vowel 'ah')

Directory **\PATHOL\AH\RESULTS** contains 635 files with extension ***.res***. They represent Kay Elemetrics MDVP format (version 1.34 or above) recordings of the acoustic parameter results deriving from analysis of ***.nsp*** files in directory **\PATHOL\AH**. They keep the same file names, but different extensions. The format is a comma-delimited ASCII text. Some of the original ***.nsp*** files do not have corresponding ***.res*** files in this directory because the acoustical manifestation of the pathology is sometimes too severe to allow MDVP evaluation.

Directory **\NORM\AH\RESULTS** contains 53 files with extension ***.res*** deriving from MDVP analysis of the ***.nsp*** files in directory **\NORM\AH**.

The ***.res*** files can be accessed and edited using any text editor or software accepting comma-delimited format.

## 9. Acoustic Speech Sample Files (Rainbow Passage)

Along with the voice sample files allowing acoustic evaluation of the vocal function, the CD-ROM also contains Acoustic Speech Sample Files - readings of the "Rainbow Passage",

Directory **\PATHOL\RAINBOW** contains 662 files with extension ***.nsp***. They represent Kay Elemetrics CSL format recordings of up to 12-second readings of the "Rainbow Passage" from the same patients included in **\PATHOL\AH** and described in the database. These files were collected at the Massachusetts Eye and Ear Infirmary (MEEI) Voice and Speech Lab, Boston, MA.

Directory **\NORM\RAINBOW** contains 53 files with extension ***.nsp*** - recordings of up to 12-second readings of the "Rainbow Passage" from the same normal subjects included in **\NORM\AH** and described in the database. These files were collected at both Kay Elemetrics Corp. and Massachusetts Eye and Ear Infirmary (MEEI) Voice and Speech Lab.

The ***.nsp*** files can be played, edited, and analyzed using MDVP or CSL with a 4300B, 4300, or 3300 hardware platform. The speech files can be very useful for perceptual evaluation along with the automatic acoustic assessment of the sustained vowels.

## 10. Method of recording

All acoustic files have been recorded using the same methodology which was the following:

A condenser microphone in a sound-proof booth has been used. The distance from the mouth to the microphone was 15 cm. All recordings have been done on a DAT-recorder at sampling rate 44.1 kHz. Uniform SPL calibration has been used for all recordings.

From the DAT-tape the recordings have been converted into an analog signal an acquired into a CSL system model 4300 at sampling rates 25 kHz (with 12 kHz anti-aliasing filtering), or 50 kHz (with 24 kHz anti-aliasing filtering). They have been saved as CSL DOS-files (Kay Elemetrics ***.nsp*** format). The names have been formed as described above.

# APPENDIX D

## MICROPHONE FOR DATA COLLECTION

Models SM48 and SM48S User Guide

### MODELS SM48 AND SM48S UNIDIRECTIONAL DYNAMIC MICROPHONES

Shure Models SM48 and SM48S unidirectional dynamic microphones are designed for professional sound reinforcement, studio recording, and broadcasting applications. They maintain a true cardioid pattern throughout their frequency range, ensuring high gain before feedback and rejection of off-axis sound. The tailored frequency response is ideal for vocals. A presence rise brightens mid-range performance, and a low frequency rolloff controls proximity effect.

Both models include a shock mounted cartridge, a steel mesh grille, and an integral "pop" filter. The SM48S also includes a lockable On/Off switch.

Each microphone is supplied with an adjustable stand adapter and a zippered carrying/storage bag. Cables are available separately.

### FEATURES

- *Cardioid pickup pattern rejects off-axis sound and provides superior gain before feedback*
- *Frequency response tailored for vocals, with brightened mid-range and bass rolloff to control proximity effect*
- *Shock-mounted cartridge for exceptional ruggedness and reduced handling noise*
- *Built-in "pop" filter that reduces explosive breath sounds and wind noise*
- *Supplied stand adapter and carrying/storage bag*
- *Lockable On/Off switch (SM48S only)*
- *Legendary Shure quality and ruggedness*

### VARIATIONS

SM48

SM48S (with lockable On/Off switch)

### SPECIFICATIONS

**Type**

Dynamic

**Frequency Response**

55 to 14,000 Hz (see Figure 1)

**Polar Pattern**

Cardioid (unidirectional) rotationally symmetrical about microphone axis, uniform with frequency (see Figure 2)

**Impedance**

Microphone rating impedance is 150 Ω (270 Ω actual) for connection to microphone inputs rated at 19 to 300 Ω

**Output Level (at 1,000 Hz)**

Open Circuit Voltage: * ............ −57.5 dBV/Pa (1.3 mV)

*1 Pa = 94 dB SPL

**Connector**

XLR professional audio

**Polarity**

Positive pressure on diaphragm produces positive voltage on pin 2 with respect to pin 3. See Figure 3.

**Dimensions**

See Figure 4

**Net Weight**

370 grams (13.1 oz)

**Packaged Weight**

672 grams (1 lb, 8.5 oz)

### CERTIFICATIONS

Eligible to bear CE Marking. Conforms to European EMC Directive 89/336/EEC. Meets applicable tests and performance criteria in European Standard EN55103 (1996) parts 1 and 2, for residential (E1) and light industrial (E2) environments.

### FURNISHED ACCESSORIES

Swivel Adapter ............ A25D
Carrying/Storage Bag ............ 26A13

### OPTIONAL ACCESSORIES

Windscreen ............ A58WS
Desk Stand ............ S37A, S39A
Isolation Mount ............ A55M, A55HM
Dual Mount ............ A26M
7.6 m Cable (25 ft) ............ C25F

### REPLACEMENT PARTS

Cartridge ............ R136
Screen and Grille Assembly ............ RK248G
Plug Assembly ............ RK40P
On/Off Switch ............ RK81S

### LOCKING THE ON/OFF SWITCH (SM48S ONLY)

To lock the On/Off switch in On position, remove the screw from the lockplate and rotate the lockplate 180°. Then reinstall the screw.

For additional service or parts information, please contact the Shure Service Department at 1-800-516-2525. Outside the United States, please contact your authorized Shure Service Center.



Printed in U.S.A.

## MICROPHONES DYNAMIQUES UNIDIRECTIONNELS MODÈLES SM48 ET SM48S

### CARACTÉRISTIQUES

**Type**
Dynamique

**Réponse en frequence**
55 à 14.000 Hz (voir la figure 1)

**Dispérsion**
Cardioïde (unidirectionnel) symétrique circulairement sur axe, uniforme avec la fréquence (voir la figure 2)

**Impédance**
L'impédance nominale est 150 Ω. (effective 270 Ω) pour entrées microphones données pour 19 à 300 Ω.

**Niveau de sortie (à 1 kHz)**
En circuit ouvert (voltage):* . . . . . . .–57.5 dBV/Pa (1.3 mV)
*1 Pa = 94 dB SPL

**Polarité**
Une pression positive produit une tension positive sur la broche 2 par rapport a la broche 3 (voir la figure 1)

**Dimensions hors tout**
Voir la figure 4

**Poids net**
370 grammes

### VARIATIONS

SM48
SM48S (avec interupteur d'alimentation electrique)

### HOMOLOGATIONS

Autorisé à porter la marque CE. Conforme à la directive CEM européenne 89/336/CEE. Conforme aux critères applicables de test et de performances de la norme européenne EN 55103 (1996) parties 1 et 2 pour les environnements résidentiels (E1) et d'industrie légère (E2).

### ACCESSOIRES FOURNIS

Adaptateur pivotant ........ A25D
Sac de rangement ........ 26A13

### ACCESSOIRES EN OPTION

Bonnet ........ A58WS
Pied de table ........ S37A
Pied a montage anti-vibrations ........ S39A
Câble (7.6 m) ........ C25F

### PIÈCES DE RECHANGE

Cartouche ........ R136
Prise ........ RK40P
Grille avec écran ........ RK248G
Interrupteur Marche/Arrêt (SM48S) ........ RK81S

Pour tout renseignement complémentaire, prière de prendre, contact avec le service Entretien Shure au 1/800-516-2525. En dehors des États-Unis, prière de prendre contact avec le centre d'entretien agréé Shure local.

## DYNAMISCH RICHTMIKROFON MODELL SM48 UND SM48S

### TECHNISCHE DATEN

**Wandlerprinzip**
Dynamisch, Tauchspule

**Frequenzgang**
55...14.000 Hz (siehe Abbildung 1)

**Richtcharakteristik**
Nierenförmig, frequenz-unabhäningig und symmetrisch zur Mikrofonachse (siehe Abbildung 2)

**Impedanz**
Die Nenn-Impedanzen betragen 150 Ω. Vorgesehen für den Anschluss an niederohmige Mikrofoneingänge, die für Quellimpe-danzen von 150 Ω bemessen sind.

**Ubertragungsfaktor (bei 1 kHz)**
Feldleerlaufbertragungsfaktor*........–57,5 dBV/Pa (1,3 mV)
*1 Pa = 94 dB SPL

**Polarität**
Positiver Schalldruck ergibt positive Spannung am Anschlußstift 2, gemessen tum Anschlußstift 3 (siehe Abbildung 3)

**Gesamtabmessungen**
Siehe Abbildung 4

**Nettogewicht**
370 g

**Schalter**
SM48
SM48S (mit netz ein/aus-kopf)

### ZULASSUNG

Zur CE-Kennzeichnung berechtigt. Entspricht der EU-Richtlinie über elektromagnetische Verträglichkeit 89/336/EEC. Erfüllt die Prüfungs- und Leistungskriterien der europäischen Norm EN 55103 (1996) Teil 1 und 2 für Wohngebiete (E1) und Leichtindustriegebiete (E2).

### MITGELIEFERTES ZUBEHÖR

Stativhalterung........ A25D
Tragetasche........ 26A13

### SONDERZUBEHÖR

Windschutz ........ A58WS
Tischständer........ S37A
Schwingungsisolierter Tischständer........ S39A
Kabel ........ C25F

### ERSATZTEILE

Kapsel ........ R136
Anschlusskupplung ........ RK40P
Einsprechkorb ........ RK248G
Ein-/Ausschalter (SM48S) ........ RK81S

Weitere Informationen hinsichtlich Service oder Ersatzteile erhalten Sie vom Shure-Zentral-Kundendienst unter der Nummer 1-800-516-2525. Außerhalb der Vereinigten Staaten von Amerika wenden Sie sich Bitte an das entsprechende autorisierte Service-Center Ihres Landes.

**MICROFONO DINAMICO UNIDIRECCIONAL**
**MODELOS SM48 Y SM48S**

## CARACTERISTICAS

**Tipo**
Dinámico

**Respuesta a Frecuencias**
55 a 14.000 Hz (vea la Figura1)

**Dispersión Polar**
Cardioide (unidireccional), uniforme respecto a la frecuencia y simétrica respecto al eje (vea la Figura 2)

**Impedancia**
La impedancia especificada es de 150 Ω (270 Ω efectivos) para su conexión a entradas de micróno especificadas entre 19 y 300 Ω..

**Nivel de salida (a 1 kHz)**
Tensión en circuito abierto* .............. −57,5 dBV/Pa (1,3 mV)

*1 Pa = 94 dB SPL

**Polaridad**
La presión positiva produce un voltaje positivo en el pin 2 respecto al 3 (vea la Figura 3)

**Dimensiones totales**
Vea la Figura 4

**Peso neto**
370 g

**Variaciones**
SM48
SM48S (con interruptor de encidido)

**CERTIFICACIONES**

Califica para llevar las marcas CE. Cumple la directiva europea 89/336/EEC de compatibilidad electromagnética. Se ajusta a los criterios correspondientes de verificación y funcionamiento establecidos en la norma europea EN 55103 (1996), partes 1 y 2, para zonas residenciales (E1) y zonas de industria ligera (E2).

## ACCESORIOS SUMINISTRADOS

Abrazadera .................................................................. A25D
Bolsa de transporte.........................................................26A13

## ACCESORIOS OPCIONALES

Pantalla antiviento .............................................. A58WS
Pié de sobremesa ................................................... S37A
Pié con aislador de vibraciones ............................ S39A
Cable (7.6 m)............................................................ C25F

## REPUESTOS

Cápsula.......................................................................R136
Enchufe ......................................................................RK40P
Conjunto de protector y rejilla.................................. RK248G
Interruptor (SM48S)....................................................RK81S

Para información adicional acerca del servicio o de partes, llame al Departamento de Servicio Shure a 1-800-516-2525. Fuera de los EE.UU., llame al servicentro autorizado de productos Shure.

**MICROFONO DINAMICO UNIDIREZIONALE**
**MODELLO SM48 Y SM48S**

## CARATTERISTICA

**Tipo**
Dinamico

**Risposta in frequenza**
Da 55 a 14.000 Hz (vedi Figura 1)

**Diagrammi Polare**
Cardioide (unidirezionale), uniforme rispetto all'asse del microfono, uniforme con la frequenza (vedi Figura 2)

**Impedenza**
Valore nominale: 150 Ω (270 Ω effettivi) per il collegamento a ingressi microfonici con bassi valori nominali di impedenza (19-300 Ω).

**Livelli de uscita (a 1.000 Hz)**
Tensione a circuito aperto*: ............... −57,5 dBV/Pa (1,3 mV)

*1 Pa = 94 dB SPL

**Polarità**
Una pressione positiva sul diaframma produce una tensione positiva al piedino 2 rispetto al piedino 3 (vedi Figura 3)

**Dimensioni totali**
Vedi Figura 4

**Peso netto**
370 g

**Variazioni**
SM48
SM48S (con interruttore di alimentazione ON/OFF)

**OMOLOGAZIONI**

Contrassegnabile con il marchio CE. Conforme alla direttiva europea sulla compatibilità elettromagnetica 89/336/CEE. Conforme ai criteri sulle prestazioni e alle prove pertinenti specificati nella norma europea EN 55103 (1996) parti 1 e 2, per ambienti residenziali (E1) e industriali leggeri (E2).

## ACCESSORI IN DOTAZIONE

Adattore regolabile ................................................. A25D
Borsa ..........................................................................26A13

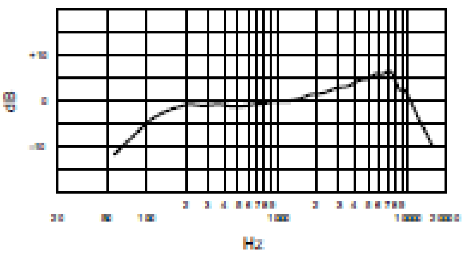
## ACCESSORI OPTIONAL

Schemo antivento.......................................... A58WS
Base da tavolo .............................................S37A
Adattore antivibrazioni ..................................S39A
Cavo (7.6 m)....................................................C25F

## RECAMBI

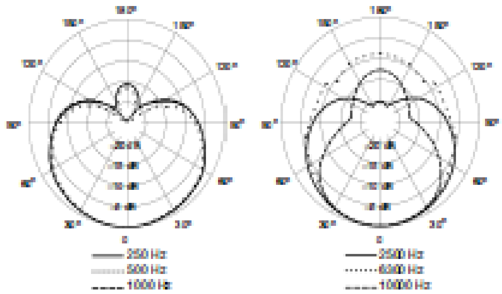
Cartuccia ......................................................R136
Connettore....................................................RK40P
Gruppo griila e schemo .................................RK248G
Interruttore (SM48S)......................................RK81S

Per ulteriori informazioni sui ricambi o per assistenza, chiamare l'assistenza clienti della Shure al numero verde 1-880-516-2525 (solo negli stati Uniti). Fuori dagli Stati Uniti rivolgersi al rivenditore o ad un centro di assistenza autorizzato.



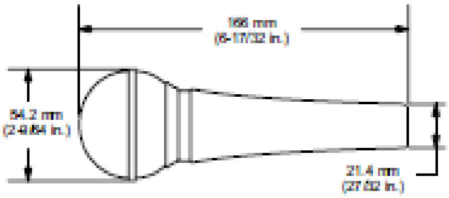

TYPICAL FREQUENCY RESPONSE
COURBE DE REPONSE TYPIQUE
TYPISCHES FREQUENZVERHALTEN
RESPUESTA DE FRECUENCIA TIPICA
TIPICA RISPOSTA IN FREQUENZA

FIGURE 1 – ABBILDUNG 1 – FIGURA 1



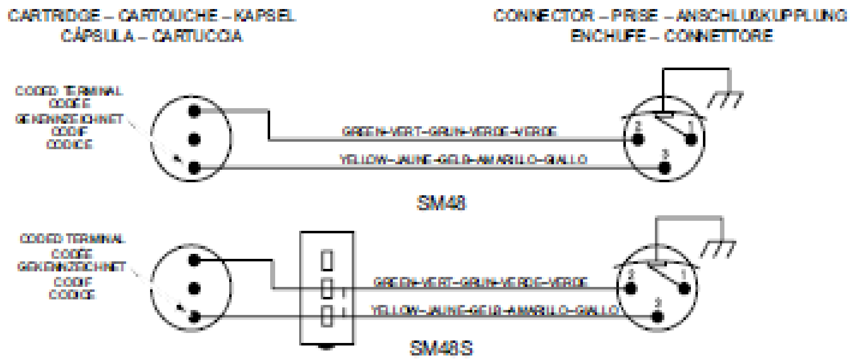

TYPICAL POLAR PATTERNS
COURBE DE DIRECTIVITE TYPIQUES
TYPISCHE POLARCHARAKTERISTIK
PATRONES DE CAPTACION POLAR TIPICOS
TIPICI DIAGRAMMI POLARI

FIGURE 2 – ABBILDUNG 2 – FIGURA 2




OVERALL DIMENSIONS – DIMENSIONS HORS TOUT
GESAMTABMESSUNGEN – DIMENSIONES TOTALES – DIMENSIONI TOTALI

FIGURE 3 – ABBILDUNG 3 – FIGURA 3




INTERNAL CONNECTIONS – CONNEXIONES INTERNES
INTERNE SCHALTUNGEN – CONEXIONES INTERNAS
COLLEGIAMENTI INTERNI

FIGURE 4 – ABBILDUNG 4 – FIGURA 4

SHURE

*SHURE Incorporated http://www.shure.com*
*United States, Canada, Latin America, Caribbean:*
*5800 W. Touhy Avenue, Niles, IL 60714-4608, U.S.A.*
*Phone: 847-600-2000 U.S. Fax: 847-600-1212 Intl Fax: 847-600-6446*
*Europe, Middle East, Africa:*
*Shure Europe GmbH, Phone: 49-7131-72140 Fax: 49-7131-721414*
*Asia, Pacific:*
*Shure Asia Limited, Phone: 852-2893-4290 Fax: 852-2893-4055*