AUTOMATIC PHONETOGRAPHIC RECORDING OF NORMAL VOICE

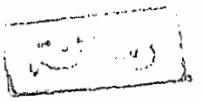
Thesis

Submitted for Partial Fulfillment of the Master Degree in

PHONIATRICS

Presented by

Ahmad Abdel Aziz Orabi



Supervisors

PROF. DR. MOHAMMAD NASSER KOTBY

Professor and Head of Phoniatric Unit

Ain Shams University - Egypt

ASS. PROF. DR. MAHMOUD Y. ABUL ELA

Assistant Professor of Phoniatrics

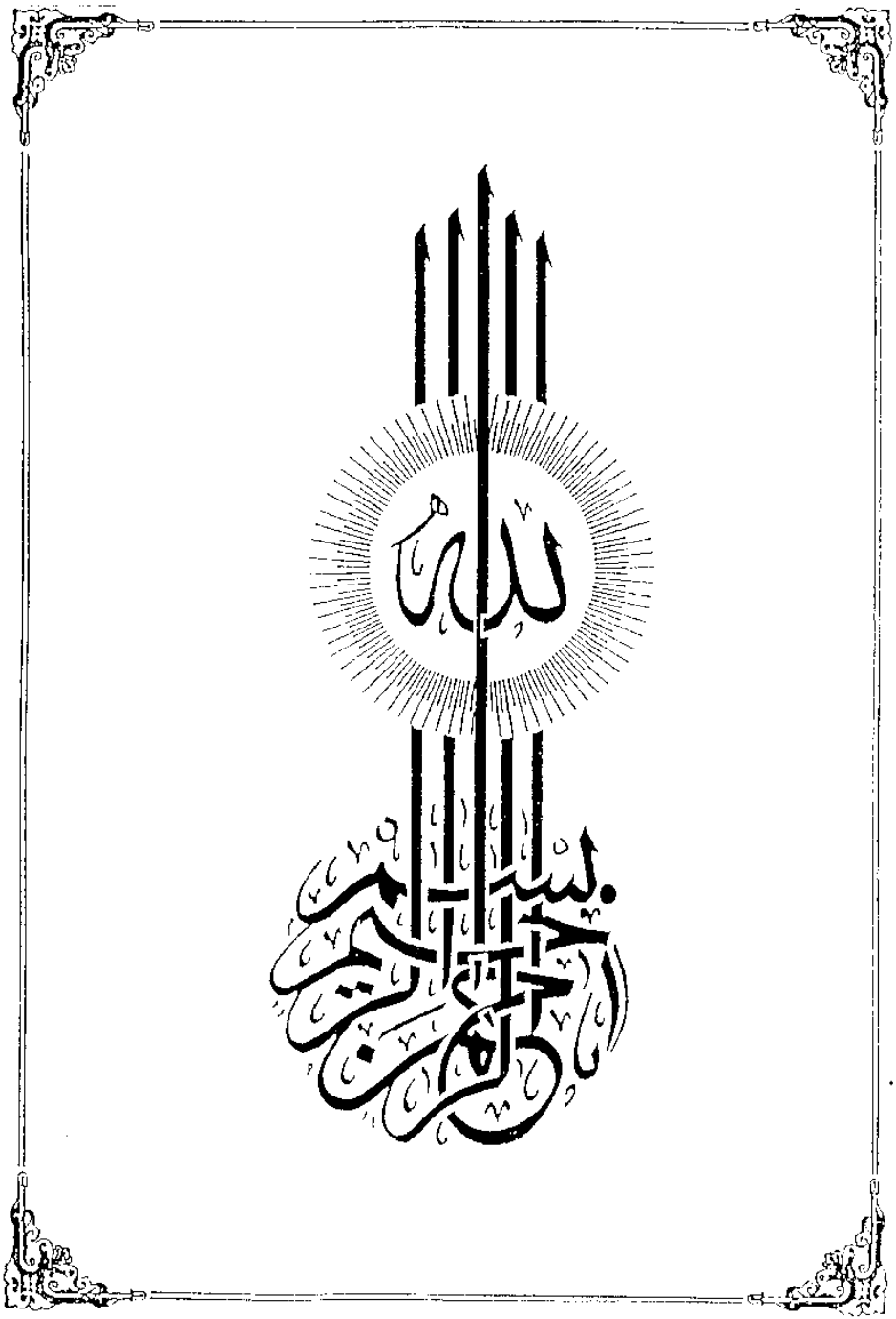
Ain Shams University - Egypt

FACULTY OF MEDICINE

AIN SHAMS UNIVERSITY

1994

In the name of Allah, the Compassionate, the Merciful



ACKNOWLEDGMENT

Praise be to Allah, who had guided us this; never could we have found guidance, had it is not been for the guidance of Allah.

Then I Would like to express my sincere thanks and gratitude to PROF. DR. NASSER KOTBY, Head of the Phoniatric Unit, Ain Shams University, for his expert guidance, endless persistent support and constructive criticism throughout this work.

I am delighted to express my deepest gratitude to Ass. PROF. DR MAHMOUD YOSSEF ABUL ELLA for his advice and valuable time he spent in revising this work. I would never forget the friendliness and good- natured assistance I received from him. Thanks to everyone who offered me assistance or encouragement along the course of this work.

I do hope that my Lord will show me the smooth and straight Path.

A.A.Orabi

To My Mother

with appreciation

Contents

1

Introduction and Aim of the Work

Introduction

Aim of the work

Voice is the carrier wave of verbal communication. It is classified as a musical sound. The propagation of a musical sound requires an excitor, a vibrator and a resonator. The vibrator has to be set in motion or excited by some source of energy; a blow or blast of air that creates subglottic pressure. This energy is converted by the vocal folds into acoustic energy expressed as sound pressure level (SPL). The number at which the vocal folds vibrate per second is called the fundamental frequency (Fo). Fo reflects the biochemical properties of the vocal folds which are determined by the laryngeal structure and applied muscle forces which in turn are adjusted by reflexive and learned voluntary behaviors. The ability to change Fo demonstrates a great deal about the mechanical adequacy of laryngeal structure and precision and extent of laryngeal control (Baken, 1987).

As sound waves travel across the vocal tract, it is differentially emphasized by the resonator. This makes the SPL of any vocalization reflecting both voice source and vocal tract properties, and it is regarded as an interesting voice trait from a voice function point of view (Gramming, 1988a).

Thus , Fo and SPL of the resultant acoustic signal are the most common control parameters in speech and singing voice synthesis (Coleman, 1993). It is well known that voice Fo and vocal intensity often co- vary in speech and singing. People raise their voices in pitch when they raise their voices in

loudness and vice versa (Titze, 1992). Thus the co-variation between Fo and SPL is found in most aspects of phonation (Coleman, 1993).

This intrigued many scientists to record such a relationship. Although results were published as early as 1938 by Stout, only after re-introduction by Calvet and Malhiac (1952) and Damste' (1970) did the method known more widely under the name of *PHONETOGRAPHY*.

Voice range profile (Phonetograph) is a plot of the dynamic range of the voice as a function of Fo. Traditional (manual) phonetogram was defined as the contour of the minimum and maximum SPL value over the patient's Fo range. But, after the invention of the automatic recording technique, it extends to cover all possible Fo/ SPL combinations within one's range (Pabon and Plomp, 1988).

The phonetogram contour outlines the physiological limits of sustained phonation (Pabon, 1991). As it maps what is vocally possible, it delineates the extremes of vocal effort; a good voice often covers a wide phonetographic area (Coleman, 1993).

It should be recognized that the phonetogram represents the output of the entire phonatory mechanism not just laryngeal function. The inter-actions of voice source and vocal tract make

activity difficult without concurrent monitoring by some other means. So, the present clinical usefulness of the phonetogram is more as a descriptor of the total vocal tract output rather than a definitive measure of laryngeal function (Coleman, 1993).

Recent work in automatic phonetogram recording permits faster sampling and a more comprehensive mapping of an individual's entire range of vocal frequency and SPL function. This may produce a more realistic performance profile than the traditional maximum- minimum contours (Pabon, 1991).

The phonetogram is a potentially valuable tool in the description of vocal performance in health and disease. Documentation of vocal pitch and loudness co-variance may be a means to track vocal changes as a function of maturation, differentiating trained from untrained subjects or changes in pathological conditions of the vocal system. It provides indications of a damage- risk criterion. Extensive sampling within an individual's profile may lead to discovery of areas of SPL or Fo discontinuity, indicating either lack of control (as in "register break") or the presence of a physical problem such as a vocal polyp (Coleman, 1993). Furthermore, it is felt that identification of the Fo/ intensity ranges of speech within phonetograms may provide important information dealing with the use and development of the speaking voice (Awan, 1993a).

Aim of the work:-

The aim of this work is to establish quantitative data for automatic phonetograms for normal male and female voices in order to present data base for normative values to be a reference for pathological vocal tract conditions and examine the relationship of age and sex to these phonetographic configurations and results. Furthermore, this work aims at examining vocal behavior in speech within phonetograms and relating both for better understanding of the voice dynamics.

2

Review of Literature

Highlights of vocal physiology

Vocal registers

Influence of the vocal tract on the glottal source

Control of pitch, loudness and quality

Phonetogram

Historical backgrounds

Definition and methodology

Possible variables in obtaining phonetograms

Data and characteristics

Interpretation

Applications

Highlights of vocal physiology

Vocal registers

Human beings, especially singers, in contrast to many musical instruments can produce great varieties of the vibratory patterns of the vocal folds- described as *vocal registers*- and subsequent tonal variations with the use of only one sound generator, that is a pair of vocal folds. A piano, violin or guitar, for instance, requires multiple sound generators in order to produce variations in tone. In other words, the human vocal folds can become vibrators which have variant mechanical properties and subsequent vibratory characteristics. This remarkable versatility in sound production is possible because of two properties of the human vocal folds; they are subject to fine grained and delicate muscular control, and their structure is ideally suited for the task because of being layered and physically pliable (Hirano, 1988).

Kotby *(in personal communication)* defines *Register* as a group of notes that have general acoustically and physiologically related characteristics. Vocal registers remain a controversial and hotly debated subject among voice scientists and teachers of singing (Coleman, 1993). Three main registers have been identified describing three types of vocal activities in the larynx:-

1. The vocal fry (Pulse register): Which has a relatively low frequency and a very long closed phase **(Fig. 2.1a)**. It is characterized by excessive lateral compression forces resulting in a large glottal resistance. In vocal fry, vocalis (VOC) activity is minimum. Cricothyroid (CT) activity is presumably minimum too, because the vocal folds are very short, massive and loose along their borders, so the vocal fold cover vibrates very slowly, only 30-80 times/ second **(Painter, 1986)**. Therefore, all the layers are supposed to be slackened and pliant **(Hirano, 1987)**.

2. The modal register: The vocal fold vibrations seen in this register are characterized by being clearly made of a vertical and horizontal element (glottal mucosal wave). The amplitude of the wave depends on the frequency and intensity levels. The upper and lower lips of the medial surface of the vocal folds move at different phases. The lower lips of this surface separate earlier in the opening phase and are approximated (sucked) earlier during the closing phase of the vibrating cycle. This register may be subdivided into 3 main divisions (head, mid and chest) depending on the frequency level. These differences are supposed to depend chiefly on the delicate balance between thyroarytenoid (TA) (shortening and rounding) and CT (lengthening and thinning) of the vocal fold. In this register, VOC activity is dominated over CT activity. As Fo increases, both VOC and CT activities increase,

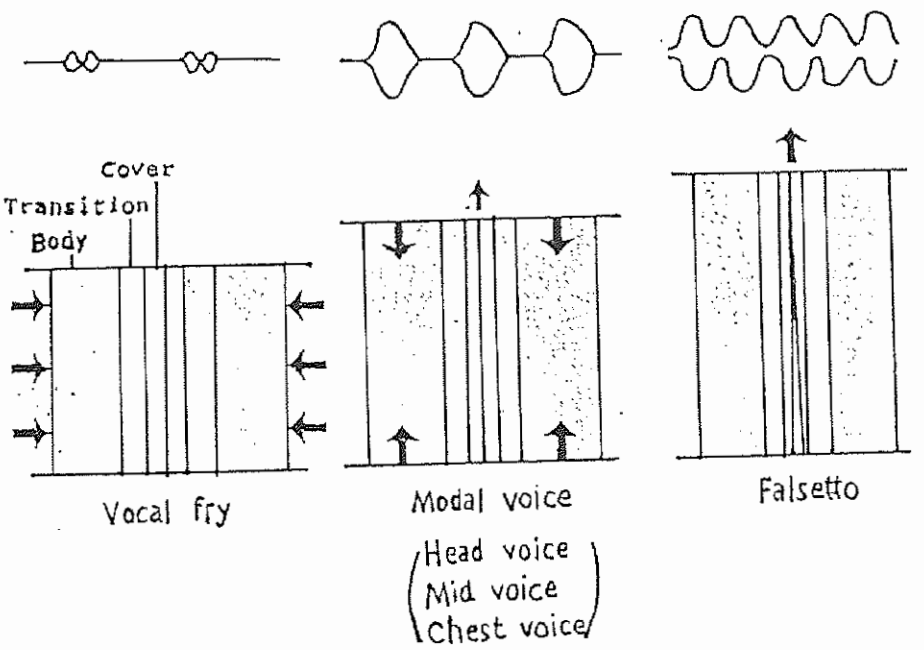


Fig. 2.1. Vocal registers: a) vocal fry b) modal c) falsetto (after Hirano, 1981)